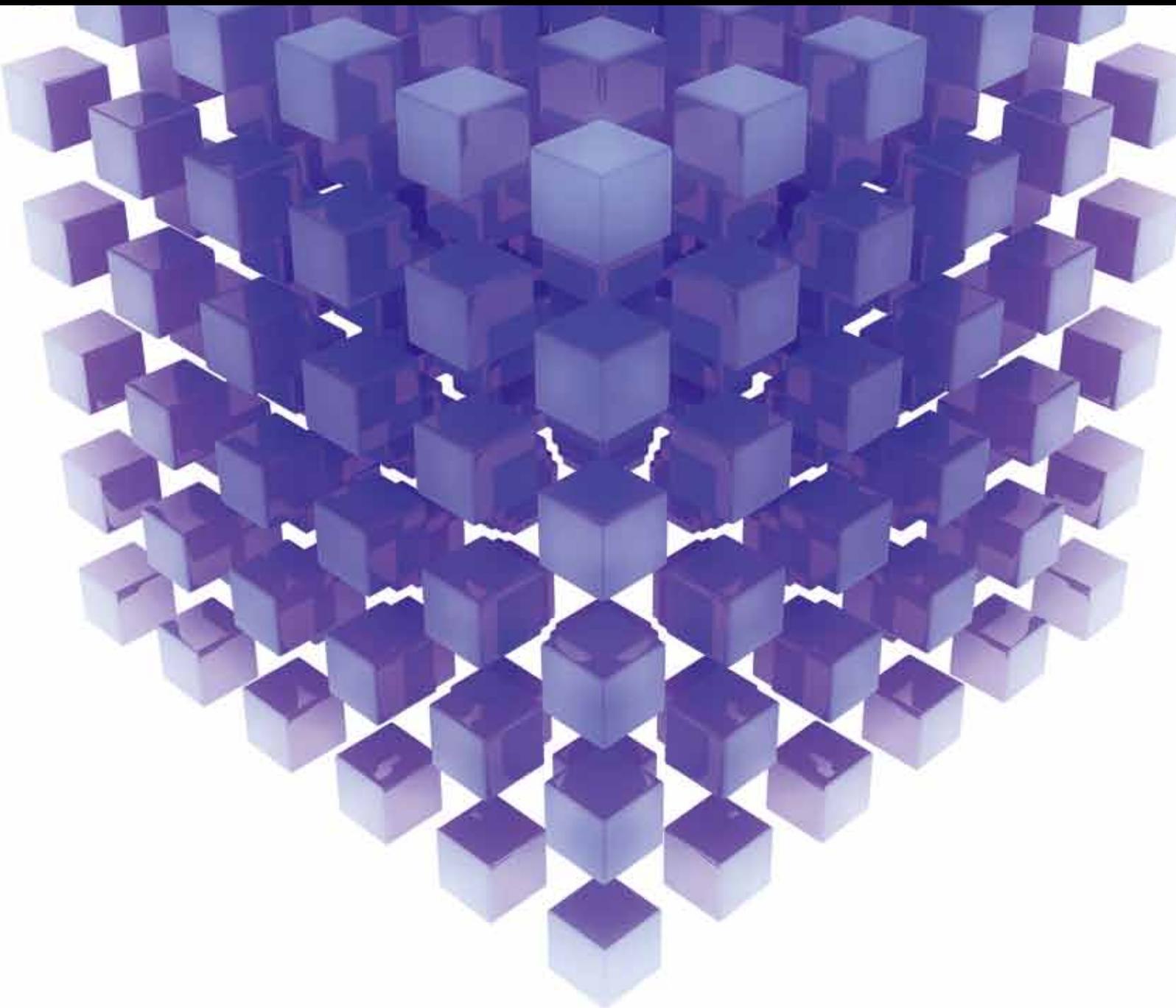


MATHEMATICAL PROBLEMS IN ENGINEERING

GREEN TRANSPORTATION SYSTEM AND SAFETY

GUEST EDITORS: WUHONG WANG, GEERT WETS, AND HEINER BUBB





Green Transportation System and Safety

Mathematical Problems in Engineering

Green Transportation System and Safety

Guest Editors: Wuhong Wang, Geert Wets, and Heiner Bubb



Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “Mathematical Problems in Engineering.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- M. Abd El Aziz, Egypt
Sarp Adali, South Africa
Ricardo Aguilar-López, Mexico
Salvatore Alfonzetti, Italy
Tofiqh Allahviranloo, Iran
Christian B. Allen, UK
Lionel Amodeo, France
Igor Andrianov, Germany
Sebastian Anita, Romania
Sabri Arik, Turkey
Fumihiro Ashida, Japan
W. Assawinchaichote, Thailand
Seungik Baek, USA
Ezzat G. Bakhoun, USA
Stefan Balint, Romania
José M. Balthazar, Brazil
Alfonso Banos, Spain
Rasajit K. Bera, India
Jonathan N. Blakely, USA
Stephane P.A. Bordas, USA
Daniela Boso, Italy
Francesco Braghin, Italy
Reyolando M. Brasil, Brazil
Michael J. Brennan, UK
Salvatore Caddemi, Italy
Jose E. Capilla, Spain
Carlo Cattani, Italy
Marcelo M. Cavalcanti, Brazil
M. Chadli, France
Yong-Kui Chang, China
Shuenn-Yih Chang, Taiwan
Ching-Ter Chang, Taiwan
Michael J. Chappell, UK
Kui Fu Chen, China
Kue-Hong Chen, Taiwan
Xinkai Chen, Japan
Jianbing Chen, China
Xuefeng Chen, China
Zhang Chen, China
Chunlin Chen, China
Singa W. Chiu, Taiwan
Jyh-Hong Chou, Taiwan
Slim Choura, Tunisia
Hung-Yuan Chung, Taiwan
Marcelo J. Colao, Brazil
- Carlo Cosentino, Italy
C. Cruz-Hernandez, Mexico
Erik Cuevas, Mexico
Weizhong Dai, USA
Binxiang Dai, China
P. Damodaran, USA
Swagatam Das, India
John M. Davis, USA
Filippo de Monte, Italy
Y. Dimakopoulos, Greece
Baocang Ding, China
Zhengtao Ding, UK
Mohamed Djemai, France
Joao B. R. Do Val, Brazil
Alexandre B. Dolgui, France
Rodrigo W. dos Santos, Brazil
Alexander N. Dudin, Belarus
George S. Dulikravich, USA
Horst Ecker, Austria
M. Onder Efe, Turkey
Elmetwally Elabbasy, Egypt
A. Elías-Zúñiga, Mexico
Fouad Erchiqui, Canada
Anders Eriksson, Sweden
Vedat S. Erturk, Turkey
Hua Fan, China
Ricardo Femat, Mexico
Thierry Floquet, France
C. R. Fuerte-Esquivel, Mexico
Tomonari Furukawa, USA
Zoran Gajic, USA
Ugo Galvanetto, Italy
Xin-Lin Gao, USA
Furong Gao, Hong Kong
Zhong-Ke Gao, China
Oleg V. Gendelman, Israel
Paulo B. Gonçalves, Brazil
Oded Gottlieb, Israel
Quang Phuc Ha, Australia
Masoud Hajarian, Iran
Zhenlai Han, China
Thomas Hanne, Switzerland
Xiao-Qiao He, China
K. R. (Stevanovic) Hedrih, Serbia
M.I. Herreros, Spain
- Wei-Chiang Hong, Taiwan
J. Horacek, Czech Republic
Feng-Hsiang Hsiao, Taiwan
Fu-Shiung Hsieh, Taiwan
Changchun Hua, China
Chiung-Shiann Huang, Taiwan
Zhenkun Huang, China
Gordon Huang, Canada
Huabing Huang, China
Chuangxia Huang, China
Yi Feng Hung, Taiwan
Hai-Feng Huo, China
Asier Ibeas, Spain
Anuar Ishak, Malaysia
Reza Jazar, Australia
Khalide Jbilou, France
Linni Jian, China
Bin Jiang, China
Zhongping Jiang, USA
Jun Jiang, China
Jianjun Jiao, China
J. J. Judice, Portugal
Tadeusz Kaczorek, Poland
Tamas Kalmar-Nagy, USA
Tomasz Kapitaniak, Poland
Haranath Kar, India
Hamid Reza Karimi, Norway
Metin O. Kaya, Turkey
Chaudry M. Khalique, South Africa
Farzad Khani, Iran
Kyung Y. Kim, Republic of Korea
Nam-Il Kim, Korea
Do W. Kim, Korea
Manfred Krafczyk, Germany
Vladislav Kravchenko, Mexico
Ren-Jieh Kuo, Taiwan
Jurgen Kurths, Germany
Hak-Keung Lam, UK
Usik Lee, Korea
Wen-Chiung Lee, Taiwan
Marek Lefik, Poland
Valter J. S. Leite, Brazil
Stefano Lenci, Italy
Roman Lewandowski, Poland
Shihua Li, China

Ming Li, China
Jian Li, China
Qing Q. Liang, Australia
Yan Liang, China
Teh-Lu Liao, Taiwan
Panos Liatsis, UK
Kim M. Liew, Hong Kong
Yi-Kuei Lin, Taiwan
Jui-Sheng Lin, Taiwan
Shueei M. Lin, Taiwan
Wanquan Liu, Australia
Yuji Liu, China
Xian Liu, China
Yan-Jun Liu, China
Peter Liu, Taiwan
Paolo Lonetti, Italy
Vassilios C. Loukopoulos, Greece
Jianquan Lu, China
Chien-Yu Lu, Taiwan
Junguo Lu, China
Jinhu Lü, China
Tiedong Ma, China
Nazim I. Mahmudov, Turkey
Fazal M. Mahomed, South Africa
Alexei Mailybaev, Brazil
Manoranjan K. Maiti, India
Oluwole D. Makinde, South Africa
Rafael Martínez-Guerra, Mexico
Driss Mehdi, France
Kh. S. Mekheimer, Egypt
Xiangyu Meng, Canada
Xinzhu Meng, China
Jose Merodio, Spain
Yuri V. Mikhlin, Ukraine
Gradimir Milovanović, Serbia
Hiroyuki Mino, Japan
Pablo Mira, Spain
Nenad Mladenovic, UK
Ebrahim Momoniat, South Africa
Gisele Mophou, France
Giuseppe Muscolino, Italy
N. B. Naduvinamani, India
Patrick Nelson, USA
Trung Nguyen Thoi, Vietnam
Hung Nguyen-Xuan, Vietnam
Ben T. Nohara, Japan
Sotiris K. Ntouyas, Greece
Gerard Olivar, Colombia
Maxim A. Olshanskii, Russia
Alkis Paipetis, Greece
Alessandro Palmeri, UK
B. K. Panigrahi, India
Manuel Pastor, Spain
Francesco Pellicano, Italy
Haipeng Peng, China
Mingshu Peng, China
Zhike Peng, China
Matjaz Perc, Slovenia
Vu Ngoc Phat, Vietnam
M. do Rosário Pinho, Portugal
José R. C. Piqueira, Brazil
Antonina Pirrotta, Italy
Stanislav Potapenko, Canada
Sergio Preidikman, USA
Carsten Proppe, Germany
Hector Puebla, Mexico
Yuming Qin, China
Dane Quinn, USA
Kumbakonam R. Rajagopal, USA
Sellakkutti Rajendran, Singapore
Gianluca Ranzi, Australia
Bhairavavajjula N. Rao, India
Sivaguru Ravindran, USA
Alessandro Reali, Italy
Giuseppe Rega, Italy
M. A. Reniers, The Netherlands
Ricardo Riaza, Spain
J. Rodellar, Spain
Rosana Rodriguez-Lopez, Spain
Carla Roque, Portugal
Debasish Roy, India
Rubén R. García, Spain
Antonio Ruiz-Cortes, Spain
Abbas Saadatmandi, Iran
Kishin Sadarangani, Spain
Roque J. Salterén, Spain
Miguel A. F. Sanjuan, Spain
Ilmar Ferreira Santos, Denmark
Nickolas S. Sapidis, Greece
E. J. Sapountzakis, Greece
Themistoklis P. Sapsis, USA
Valery Sbitnev, Russia
Massimo Scalia, Italy
Mohammed Seaid, UK
Mohamed A. Seddeek, Egypt
Fiorella Sgallari, Italy
Leonid Shaikhet, Ukraine
Cheng Shao, China
Sanjay K. Sharma, India
Bo Shen, Germany
Jian-Jun Shu, Singapore
Zhan Shu, UK
Dan Simon, USA
Luciano Simoni, Italy
Indra Vir Singh, India
C.H. Skiadas, Greece
Delfim Soares Jr., Brazil
Davide Spinello, Canada
Victor Sreeram, Australia
Sri Sridharan, USA
H. M. Srivastava, Canada
Ivanka Stamova, USA
Valder Steffen Jr, Brazil
Rolf Stenberg, Finland
Xi-Ming Sun, China
Jitao Sun, China
Yuangong Sun, China
Zhongkui Sun, China
Andrzej Swierniak, Poland
W.Y. Szeto, Hong Kong
Yang Tang, Germany
Alexander Timokha, Norway
Cristian Toma, Romania
Hien T. Tran, USA
Irina N. Trendafilova, UK
Jung-Fa Tsai, Taiwan
Chia-Cheng Tsai, Taiwan
George Tsiatas, Greece
E. Tzirtzilakis, Greece
F. Ubertini, Italy
K. Vajravelu, USA
Robertt A. Valente, Portugal
Victoria Vampa, Argentina
J. Van Dommelen, The Netherlands
Alain V. Wouwer, Belgium
Pandian Vasant, Malaysia
Josep Vehi, Spain
Kalyana C. Veluvolu, Korea
Stefano Vidoli, Italy
Michael Vynnycky, Ireland
Yan-Wu Wang, China
Junwu Wang, China
Shuming Wang, Singapore
Cheng C. Wang, Taiwan

Youqing Wang, China
Yongqi Wang, Germany
Moran Wang, China
Yijing Wang, China
Dan Wang, China
Xiaojun Wang, China
Qing-Wen Wang, China
Gerhard-Wilhelm Weber, Turkey
Hung-Yu Wei, Taiwan
J. Witteveen, The Netherlands
Kwok-Wo Wong, Hong Kong
Ligang Wu, China
Zhengguang Wu, China
Yuqiang Wu, China
Desheng D. Wu, Iceland
Huang Xia, China

Gongnan Xie, China
Xuejun Xie, China
Guangming Xie, China
Wang Xing-yuan, China
Lianglin Xiong, China
Xi Frank Xu, China
Gen-QI Xu, China
Hang Xu, China
Xing-Gang Yan, UK
Jun-Juh Yan, Taiwan
Suh-Yuh Yang, Taiwan
Chunyu Yang, China
Dan Ye, China
Mohammad I. Younis, USA
Simin Yu, China
Bo Yu, China

Huang Yuan, Germany
A. M. Zenkour, Saudi Arabia
Jianming Zhan, China
Huaguang Zhang, China
Hongbin Zhang, China
Yingwei Zhang, China
Xu Zhang, China
Hong Zhang, China
Hongyong Zhao, China
Lu Zhen, China
Liancun Zheng, China
Jian Guo Zhou, UK
Zexuan Zhu, China
Quanxin Zhu, China
Mustapha Zidi, France

Contents

Green Transportation System and Safety, Wuhong Wang, Geert Wets, and Heiner Bubb
Volume 2014, Article ID 631701, 1 page

A Simulation-Based Framework for the Cooperation of VMS Travel Guidance and Traffic Signal Control, Meng Li, Han Jiang, Zhen Zhang, Wei Ni, Pinchao Zhang, and Jingyan Song
Volume 2014, Article ID 803647, 13 pages

Research on the Safety Audit Methods for Two-Lane Highway Based on HRV, Zhan-dong Zhu, Ying-chih Lu, Cheng-hong Fu, and Ting Xu
Volume 2014, Article ID 308028, 6 pages

The Theory of Dynamic Public Transit Priority with Dynamic Stochastic Park and Ride, Chengming Zhu, Yanyan Chen, and Changxi Ma
Volume 2014, Article ID 525460, 7 pages

Pedestrian Evacuation Time Model for Urban Metro Hubs Based on Multiple Video Sequences Data, Ji-biao Zhou, Hong Chen, Jing Yang, and Jiao Yan
Volume 2014, Article ID 843096, 11 pages

Networked Timetable Stability Improvement Based on a Bilevel Optimization Programming Model, Xuelei Meng, Bingmou Cui, Limin Jia, Yong Qin, and Jie Xu
Volume 2014, Article ID 290937, 10 pages

A Novel Fast and Robust Binary Affine Invariant Descriptor for Image Matching, Xiujie Qu, Fei Zhao, Mengzhe Zhou, and Haili Huo
Volume 2014, Article ID 129230, 7 pages

Comparing the State-of-the-Art Efficient Stated Choice Designs Based on Empirical Analysis, Li Tang, Xia Luo, Yang Cheng, Fei Yang, and Bin Ran
Volume 2014, Article ID 740612, 8 pages

Decision of National and Provincial Highway Asphalt Pavement Structure Based on Value Engineering, Yingwei Ren and Jingsong Shan
Volume 2014, Article ID 369248, 8 pages

Impact of Road Bends on Traffic Flow in a Single-Lane Traffic System, Zeng Junwei, Qian Yongsheng, Xu Dejie, Jia Zhilong, and Huang Zhidan
Volume 2014, Article ID 218465, 6 pages

Multiagent-Based Simulation of Temporal-Spatial Characteristics of Activity-Travel Patterns Using Interactive Reinforcement Learning, Min Yang, Yingxiang Yang, Wei Wang, Haoyang Ding, and Jian Chen
Volume 2014, Article ID 951367, 11 pages

Topological Effects and Performance Optimization in Transportation Continuous Network Design, Jianjun Wu, Xin Guo, Huijun Sun, and Bo Wang
Volume 2014, Article ID 490483, 7 pages

The Model of Severity Prediction of Traffic Crash on the Curve, Jian-feng Xi, Hai-zhu Liu, Wei Cheng, Zhong-hao Zhao, and Tong-qiang Ding
Volume 2014, Article ID 832723, 5 pages

Modeling of Car-Following Required Safe Distance Based on Molecular Dynamics, Dayi Qu, Xiufeng Chen, Wansan Yang, and Xiaohua Bian
Volume 2014, Article ID 604023, 7 pages

Urban Traffic Signal System Control Structural Optimization Based on Network Analysis, Li Wang, Xiao-ming Liu, Zhi-jian Wang, and Zheng-xi Li
Volume 2013, Article ID 706919, 9 pages

Multi-Objective Optimization of Traffic Signal Timing for Oversaturated Intersection, Yan Li, Lijie Yu, Siran Tao, and Kuanmin Chen
Volume 2013, Article ID 182643, 9 pages

Sustainable Transport Data Collection and Application: China Urban Transport Database, Tian Jiang, Zhongyi Wu, Yu Song, Xianglong Liu, Haode Liu, and Haozhi Zhang
Volume 2013, Article ID 879752, 10 pages

Operational Impacts of Using Restricted Passenger Flow Assignment in High-Speed Train Stop Scheduling Problem, Huiling Fu, Benjamin R. Sperry, and Lei Nie
Volume 2013, Article ID 902562, 8 pages

Estimation Trajectory of the Low-Frequency Floating Car Considering the Traffic Control, Zhijian Wang, Min Li, Li Wang, and Xiaoming Liu
Volume 2013, Article ID 762924, 11 pages

A Study on the Model of Traffic Flow and Vehicle Exhaust Emission, Han Xue, Shan Jiang, and Bin Liang
Volume 2013, Article ID 736285, 6 pages

An Approach to Optimize the Departure Times of Transit Vehicles with Strict Capacity Constraints, Huimin Niu and Xiaopeng Tian
Volume 2013, Article ID 471928, 9 pages

Analysis on Residents' Travel Activity Pattern in Historic Urban Areas: A Case Study of Historic Urban Area of Yangzhou, China, Mao Ye, Miao Yu, Xiucheng Guo, Yingshun Liu, and Zhibin Li
Volume 2013, Article ID 620973, 9 pages

Using Principal Component Analysis to Solve a Class Imbalance Problem in Traffic Incident Detection, Changjiang Zheng, Shuyan Chen, Wei Wang, and Jian Lu
Volume 2013, Article ID 524861, 8 pages

State of Charge Estimation Based on Microscopic Driving Parameters for Electric Vehicle's Battery, Enjian Yao, Meiying Wang, Yuanyuan Song, and Yang Yang
Volume 2013, Article ID 946747, 6 pages

Optimizing Schedules of Rail Train Circulations by Tabu Search Algorithm, Mingming Chen and Huimin Niu
Volume 2013, Article ID 102346, 7 pages

Empirical Study of Travel Time Estimation and Reliability, Ruimin Li, Huajun Chai, and Jin Tang
Volume 2013, Article ID 504579, 9 pages

Motor Vehicle Emission Modeling and Software Simulation Computing for Roundabout in Urban City, Haiwei Wang, Huiying Wen, Feng You, Jianmin Xu, and Hailin Kui
Volume 2013, Article ID 312396, 12 pages

Deformation of Overlong Isolated Buildings Caused by Thermal and Concrete Shrinkage, Yu Dang and Ying-ke Liu
Volume 2013, Article ID 139159, 7 pages

Research on Large-Scale Road Network Partition and Route Search Method Combined with Traveler Preferences, De-Xin Yu, Zhao-Sheng Yang, Yao Yu, and Xiu-Rong Jiang
Volume 2013, Article ID 950876, 8 pages

LQR-Based Power Train Control Method Design for Fuel Cell Hybrid Vehicle, Yun Haitao, Zhao Yulan, Liu Zunnian, and Hao Kui
Volume 2013, Article ID 968203, 7 pages

Particle Filter for Estimating Freeway Traffic State in Beijing, Jun Bi, Can Chang, and Yang Fan
Volume 2013, Article ID 382042, 6 pages

Effects of Transverse Baffle Design on Reducing Liquid Sloshing in Partially Filled Tank Vehicles, Xue-lian Zheng, Xian-sheng Li, Yuan-yuan Ren, Yu-ning Wang, and Jie Ma
Volume 2013, Article ID 130570, 13 pages

An Adaptive Model for Calculating the Correlation Degree of Multiple Adjacent Signalized Intersections, Linhong Wang and Yiming Bie
Volume 2013, Article ID 247184, 13 pages

An Efficient Adaptive Denoising Algorithm for Remote Sensing Images, Xiujie Qu, Fu Zhang, and Huan Jia
Volume 2013, Article ID 207461, 5 pages

Railroad Track Deterioration Characteristics Based Track Measurement Data Mining, Peng Xu, Reng-Kui Liu, Feng Wang, Fu-Tian Wang, and Quan-Xin Sun
Volume 2013, Article ID 970573, 7 pages

Research on Face Recognition Based on Embedded System, Hong Zhao, Xi-Jun Liang, and Peng Yang
Volume 2013, Article ID 519074, 6 pages

Empirical Analysis of Traffic Bottleneck at Beijing Expressways, Sheng Jin, Dianhai Wang, and Dongfang Ma
Volume 2013, Article ID 254841, 7 pages

The Departure Characteristics of Traffic Flow at the Signalized Intersection, Zhaowei Qu, Yan Xing, Hongyu Hu, Yuzhou Duan, Xianmin Song, TingTing Chai, and Hu Zhang
Volume 2013, Article ID 671428, 11 pages

Study of Railway Track Irregularity Standard Deviation Time Series Based on Data Mining and Linear Model, Jia Chaolong, Xu Weixiang, Wei Lili, and Wang Hanning
Volume 2013, Article ID 486738, 12 pages

Rational Formations of a Metro Train Improve Its Efficiencies of Both Traction Energy Utilization and Passenger Transport, Xuesong Feng, Haidong Liu, Hanxiao Zhang, Baoshan Wang, and Qipeng Sun
Volume 2013, Article ID 643274, 7 pages

Crossing Reliability of Electric Bike Riders at Urban Intersections, Huan Mei, Yang Xiaobao, and Jia Bin
Volume 2013, Article ID 108636, 8 pages

Research on Quantitative Models of Electric Vehicle Charging Stations Based on Principle of Energy Equivalence, Zhenpo Wang, Peng Liu, Jia Cui, Yue Xi, and Lei Zhang
Volume 2013, Article ID 959065, 10 pages

Editorial

Green Transportation System and Safety

Wuhong Wang,¹ Geert Wets,² and Heiner Bubb³

¹ Department of Transportation Engineering, Beijing Institute of Technology, Beijing 100081, China

² Transportation Research Institute (IMOB), Hasselt University, 3590 Diepenbeek, Belgium

³ Lehrstuhl für Ergonomie, Technische Universität München, 85747 Munich, Germany

Correspondence should be addressed to Wuhong Wang; wangwuhong@bit.edu.cn

Received 1 June 2014; Accepted 1 June 2014; Published 16 June 2014

Copyright © 2014 Wuhong Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Effectiveness and efficiency have been accepted as the measures for the transportation system as well as its components performance. However, highlighted by the concept of sustainability, environmental impacts of the transportation system must be standardized for traffic participants, roads, and infrastructure from planning to operation. With consideration of the road transport accounting for 75% of the world's total carbon dioxide emissions from fossil fuel combustion and the transportation systems being responsible for more than 20% of world energy consumption, statistical results raise the following problems: what efforts should we make to solve such environmental problems and what countermeasures should we conduct to balance the energy saving and the demand of mobility? Within the sustainable approaches, green traffic and green transportation system are now widely discussed.

Green vehicles, smart road, C-2-X communications for green intelligent transportation systems control, and green urban traffic are proved to have considerable effectiveness. Green vehicles, with new alternative energy as their key feature, are intended to be more environmentally friendly than conventional internal combustion engine vehicles running on gasoline or diesel. Electricity, hybrid electricity, compressed air/natural gas, biofuel, solar power, and so forth can be used as the embodied energy of the vehicles which represent the different types of green vehicles. The development of green vehicles triggers the challenges not only in the manufacture, but also in the use of vehicles and the service provided by traffic system.

Smart road; C-2-X communications including communications between/among vehicles, roadways, roadside, back-end infrastructure, and so forth; and the assistance systems

based on driver-vehicle unit lay the foundation of the traffic information network and enable a smooth, efficient, and accident-free traffic. The ongoing information system in ITS and driver assistance systems will provide guidance to the road users and, moreover, contribute to the emission control and safety.

The overbuilt cities experience the consequence of overloaded traffic demand and the worsening environment. Walking and cycling traffic are encouraged by enhancing the role of public transport in order to control CO₂, air pollution, and noise. Focusing on the improvement of traffic conditions, productivity of cities will benefit from green urban traffic by mitigating the traffic congestions and delay. Furthermore, user-oriented urban traffic will contribute to the safety of vulnerable road users.

It is believed that the green transportation system (GTSS) will be a prime target for obtaining eco-friendly transportation systems and safety.

Wuhong Wang
Geert Wets
Heiner Bubb

Research Article

A Simulation-Based Framework for the Cooperation of VMS Travel Guidance and Traffic Signal Control

Meng Li,¹ Han Jiang,² Zhen Zhang,¹ Wei Ni,¹ Pinchao Zhang,¹ and Jingyan Song²

¹ Department of Civil Engineering, Tsinghua University, 315 Shanheng He Building, Beijing 100084, China

² Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Han Jiang; jianghan0521@sina.cn

Received 14 January 2014; Revised 19 May 2014; Accepted 19 May 2014; Published 16 June 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Meng Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, both travel guidance systems and traffic signal control systems are quite common for urban traffic management. In order to achieve collaborative effect, different models had been proposed in the last two decades. In recent years, with the development of variable message sign (VMS) technology, more and more VMS panels are installed on major arterials to provide highly visible and concise graphs or text messages to drivers, especially in developing countries. To discover drivers' responses to VMS, we establish a drivers' en route diversion model according to a stated-preference survey. Basically, we proposed a cooperative mechanism and systematic framework of VMS travel guidance and major arterials signal operations. And then a two-stage nested optimization problem is formulated. To solve this optimization problem, a simulation-based optimization method is adopted to optimize the cooperative strategies with TRANSIMS. The proposed method is applied to the real network of Tianjin City comprising of 30 nodes and 46 links. Simulations show that this new method could well improve the network condition by 26.3%. And analysis reveals that GA with nested dynamic programming is an effective technique to solve the optimization problem.

1. Introduction

In recent years, China has entered the stage of quickened urbanization process. However, traffic problems such as traffic congestions and the deterioration of air quality are getting increasingly serious. Since the obvious unbalance between the rapidly growing traffic demand and the limitation of traffic resources, traffic management is becoming an increasingly serious concern. During the last two decades, how to alleviate traffic congestion by better urban traffic management strategies has captivated consistent attention from both traffic operators and engineers.

To reach this goal, various strategies have been developed for urban traffic management [1, 2]. One important branch of such approaches is traffic flow guidance. Many researches have been conducted on traffic flow guidance by providing drivers with real-time information (e.g., traffic information broadcasting, on-board navigation systems, and

variable message sign). Among them, the variable message sign installed on the roadside is designed to improve road network performance by providing such highly visible and concise information to drivers, especially congested conditions [3]. Instead of being installed on highways in America and European countries, VMS are developed to provide en route information for major arterials in developing countries. The application of VMS technology in congested urban road network in developing country can potentially ease traffic congestion. It was revealed that the key of effective utilization of VMS is to capture the relationship between variable messages and drivers' behavior [4]. Therefore, sound investigations of drivers' responses to VMS should be carried out at first.

Another common urban traffic management method is traffic signal operational strategy, which provides a real-time traffic flow control approach [5]. Obviously, there exist interactions and interdependence between travel guidance

and traffic signal control. Consequently, analytical models or simulation evaluation seems inspiring by using traffic signal control system or travel guidance system separately. On the contrary, adopting traffic signal control and travel guidance simultaneously does not always succeed. Nowadays, more and more serious urban congestion problem in developing countries requires researchers' additional insights into the optimization of cooperative strategies.

In previous attempts, the cooperative mechanism related to traffic signal control and travel guidance is one of hotspots for researchers. In the early stage, most research integrates traffic signal control with user equilibrium [6–8] or dynamic user equilibrium (DUE) [9–11]. The key postulated in these studies is that all users have the same choice criterion and complete information. Taking this hypothesis, we find that this integration different from drivers' en route route-choice process neglects specific drivers' behaviors. More importantly, it is impossible to provide complete dynamic information for all drivers on urban road network. The researches on joint optimization of traffic signal control and UE or DUE cannot be put into practice so far.

To solve the aforementioned problems, some researches have been conducted on cooperation of en route travel guidance and traffic signal control. Among en route travel guidance approaches, advanced on-board navigation systems or traffic information broadcasting can provide real-time information and guidance advice for users. However, there are still some problems needed to be solved before using in developing countries: users of navigation systems only cover a tiny proportion of total travelers in developing countries; and it is isolated from signal control system. In general, these systems designed for travellers have to be separated from signal control systems for security reasons. Different with above approaches, the VMS strategically located on the roadside can directly affect drivers' en route choice. Moreover, VMS technique allows the information communication with road detectors and traffic signal controllers.

In recent years, several studies on VMS travel guidance and traffic signal control have been carried out for urban traffic management in developing countries [12, 13]. Moreover, Lin further proposed a collaborative model integrating traffic control with VMS in a sudden disaster [14]. Nevertheless, these previous studies place emphasis on logistic steps and framework. In addition, only text messages (e.g., diversion advice and congestion information) are considered in previous attempts. In order to avoid the fluctuation of traffic flow, diversion advice should be published only if detour path owns obvious advantage. In other words, the diversion advice owns lag characteristic. To improve the effectiveness of travel guidance, we propose an active travel guidance method by a dynamic graphical display of the traffic status with colors in this paper. The application of VMS will definitely induce distributions of traffic flow in different paths. As a result, not considering the response of travellers may actually increase network-wide congestion. Therefore, we establish a valid driver diversion model on a behavioral survey before implementing the cooperative strategy.

In this paper, we first establish a logit model of drivers' responses to graph and text information on VMS panels

based on a stated-preference survey. And then a cooperative strategy and systematical framework of VMS and traffic signal control are proposed for major arterials. Within this framework, the optimization of the cooperative strategy can be represented as a two-stage nested optimization problem. We name this optimization problem the VMS guidance and signal coordination (VGSC) problem. The desired road status display, diversion advice, and arterial signal parameters (i.e., cycle times and green times) are updated in the first-stage optimization problem by using genetic algorithms. Subsequently, arterial signal offsets are optimized by dynamic programming method in the second-stage. Considering the complexity of interactions between vehicles and signalized or unsignalized intersections on a large network, we further develop a simulation-based optimization (SBO) software package with TRANSIMS to solve the two-stage nested optimization problem.

To explain the mechanism and application of the cooperative strategy, the rest of this paper is organized as follows. Section 2 builds the driver diversion model on the basis of a behavioral survey and propose the active travel guidance approach in detail. Section 3 proposes the cooperative strategy and establishes the two-stage nested optimization model. Section 4 presents the simulation-based optimization method to solve the optimization problem. Computational results on a particular network are presented in Section 5. Finally, Section 6 concludes our findings.

2. Drivers' Diversion Model

2.1. Stated-Preference Behavioral Survey. Many studies have investigated the impact of VMS on drivers' behavior [15, 16] and verified that the text message on VMS panels is a significant factor for route choice. The stated-preference (SP) survey is an effective tool to test the behavior of individuals. However, these SP surveys are not suitable for VMS installed in urban areas. Due to different scenarios including different message content, applied places, and traffic status, drivers' route choices would be quite different. Therefore, it is essential to carry out a sound behavioral survey for modeling drivers' behaviors.

In developing countries, VMS are widely used in urban road network. These VMS panels allow releasing information on "congestion scale," that is, road network with different colors. The "congestion scale" is a classification of average speed calculated by travel time (i.e., red means congested road segment; yellow means heavy traffic volume segment; and green means low-traffic segment). In detail, for major arterials, red represents the real-time speed lower than 20 km/h; yellow represents 20 km/h–40 km/h; and green represents higher speed than 40 km/h. (see Figure 1) Previous research focuses on drivers' response to typical text messages while neglecting the graphical information on VMS panels.

To solve existing problems of drivers' responses to graphical information, we carried out a stated-preference (SP) survey, which is coded into pad computers. Considering road segments with different colors and detour distances that

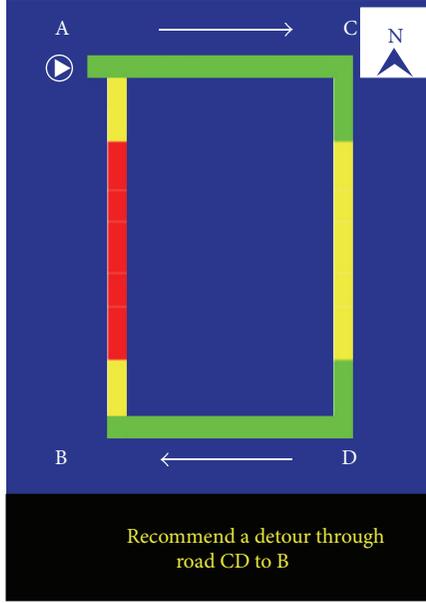


FIGURE 1: An illustration of VMS panel.

drivers might face in the urban area, various variables including distances of pretrip path (AB) and detour route (ACDB), diversion advice, and colors were changed randomly.

Multiple factors vis-à-vis VMS have been considered in the SP survey and the collected information can be classified in the following categories:

- (1) personal information: age, gender, and salary; and the personal attributes which define drivers' characteristics have been identified in Table 1;
- (2) preference information:
 - (a) a drivers' acceptability of VMS;
 - (b) a drivers' response to information on VMS;
- (3) dynamic information on VMS panels:
 - (a) distances: original route and alternative routes;
 - (b) colors: yellow and red segment ratios;
 - (c) diversion advice.

2.2. A Binary Logit Model of Drivers' Diversion Choice. On the basis of the SP survey, drivers' response to VMS is typically modeled by a binary logit model. There are two alternatives, which are denoted by 0 or 1 (i.e., 0 means keeping original route and 1 means choosing to divert). In order to recover all related factors, we take multiple variables into consideration, as follows:

- (1) age: the drivers' ages;
- (2) gender: a dummy variable, which is equal to zero if the driver is male and one otherwise;

TABLE 1: Summary of selected personal attributes.

Personal attributes		Percentage (%)
Gender	Male	44.94
	Female	55.06
Age	18–29	39.33
	30–49	58.43
	50–69	2.25
	<50,000	15.73
	50,001–100,000	21.35
	100,001–150,000	20.22
Personal salary	150,001–200,000	11.24
	200,001–250,000	16.85
	250,001–300,000	1.12
	300,001–350,000	4.49
	350,001–400,000	1.12
	>400,000	7.87

- (3) route length: the length of routes AB and ACB, denoted as l_{AB} and l_{ACB} ;
- (4) advice: a dummy variable which is equal to zero if the VMS shows the advice and one otherwise;
- (5) $YLR_{AB/ACB}$: yellow link ratio in routes AB and ACB;
- (6) $RLR_{AB/ACB}$: red ratio in routes AB and ACB. And, we have

$$RLR = \frac{L_r}{L_y + L_g + L_r}, \quad YLR = \frac{L_y}{L_y + L_g + L_r}, \quad (1)$$

where L_g , L_y , and L_r denote the length of green (smooth), yellow (heavy volume), and red (congested) links.

By analysis of collected behavioral data, we managed to find a set of significant factors. A binary logit model is presented in Table 2. The estimated coefficients of the variables are all significant at 95% confidence level. The likelihood of diversion has been positively affected by certain factors, for example, age, length of routine route l_{AB} , diversion advice, red segment ratio on the routine route RLR_0 and passively affected by length of diversion route l_{ACD} , and red segment ratio on the diversion route RLR_1 .

In addition, the utility function is formulated as

$$\begin{aligned} V_n(C) &= V_n(1) - V_n(0) \\ &= \exp(\beta_0 + \beta_1 \cdot \text{Age} + \beta_2 \cdot l_{AB} + \beta_3 \cdot l_{ACD} \\ &\quad + \beta_4 \cdot RLR_0 + \beta_5 \cdot RLR_1 \\ &\quad + \beta_6 \cdot \text{Advice} + \text{const}). \end{aligned} \quad (2)$$

TABLE 2: Binary logit model for drivers' responses to VMS.

Variables	Coefficient	Estimates	Standard.Error	Significance
Age	β_1	0.044	0.017	0.010
l_{AB}	β_2	0.965	0.096	0.001
l_{ACB}	β_3	-0.439	0.076	0.001
RLR_0	β_4	3.510	0.606	0.000
RLR_1	β_5	-6.240	1.371	0.000
Advice	β_6	0.431	0.211	0.041
Constant	β_0	-3.032	1.314	0.021

The probability that individual n chooses alternative route can be further obtained in the following equation:

$$P_n(C) = \frac{V_n(C)}{1 + V_n(C)}. \quad (3)$$

3. The Cooperative Mechanism of VMS and TSC

In this section, we propose a cooperative mechanism of VMS and traffic signal control system in urban areas. In order to improve effectiveness of cooperation, a two-stage nested optimization problem is formulated.

3.1. Cooperative Mechanism. Based on the proposed logit model in Section 2, we found that the red segments ratios of an arterial displayed on VMS have a profound impact on the behavior of en route diversion decisions. Consequently, it will be a reasonable and practical choice to take the dynamic graph information as an active travel guidance tool (i.e., to display arterial with colors predicted by a reasonable and desirable method), instead of displaying current traffic status passively. With the collaboration of traffic signal control, the new VMS strategy method can not only distribute traffic flow in desirable proportions but also provide traffic status more closely to travellers' driving feeling.

This important improvement of the active VMS strategy is achieved by two preconditions: first, we allow the cooperation of VMS and TSC systems instead of isolation ones; second and more importantly, we can find a desirable solution by joint optimization in a short term and the solution has to be consistent with drivers' driving feelings for a sustainable effectiveness in a long-term use. This approach strikes a very nice balance between accessibility and effectiveness.

By considering signal cycle length and the location of VMS panels, the control period T can be determined at first. As mentioned above, during the j th control interval (t_{j-1}, t_j) , the cooperative strategy engenders a feasible and desired solution for the next control interval (t_j, t_{j+1}) . And the solution for $(j+1)$ th control interval, denoted as x_{j+1} , is a group of control variables:

$$x_{j+1} = [RLR_0, YLR_0, RLR_k, YLR_k, d_k, C_k, g_k, \Phi_k]_{j+1}, \quad (4)$$

where RLR_0 and YLR_0 denote the red segments ratios and yellow segments ratios on the pretrip route and RLR_k and

YLR_k denote ratios on the k th route; d_k is equal to 0 or 1 and $d_k = 1$ represents giving out the diversion suggestion for k th route; C_k , g_k , and Φ_k denote the cycle times, green times, and offsets of signal controllers on the k th route. We assume that there are k alternative routes which can be chosen for drivers. Though the behavioral survey verifies the relation between drivers' behaviors and RLR, it is necessary to compute both RLR and YLR in the joint optimization problem because the two parameters determine the condition of an arterial road together.

When it comes to the aforesaid first precondition, both the VMS and TSC systems are installed and controlled by transportation management agencies. The implementation of the joint optimization for the two systems in the control center is easy to complete.

On the other hand, when it comes to the second precondition, some if-then rules should be satisfied for the optimization process. We assume that the control result is y_{j+1} under the condition of x_{j+1} (see (4)):

$$x'_{j+1} = [RLR'_0, YLR'_0, RLR'_k, YLR'_k], \quad (5)$$

where the RLR'_0 , YLR'_0 , RLR'_k , and YLR'_k represent the estimation of the control results on pretrip route and other routes:

$$\begin{aligned} &\text{if } d_k = 1 \\ &\text{then } RLR_k < RLR_0, \\ &\quad YLR_k < YLR_0 \end{aligned} \quad (6)$$

$$\begin{aligned} &RLR'_k < RLR_k, \\ &YLR'_k < YLR_k, \end{aligned}$$

$$\begin{aligned} &\text{if } d(k) = 0, \quad RLR_k < RLR_0, \\ &\quad YLR_k < YLR_0 \end{aligned} \quad (7)$$

$$\begin{aligned} &\text{then } RLR'_k \leq RLR_k, \\ &\quad YLR'_k \leq YLR_k. \end{aligned}$$

Equation (6) ensures that the advice suggestions will be consistent with the graph on VMS and traffic status on actual road network. Equation (7) ensures that when there is no suggestion on panels the graph also should be consistent with traffic status on actual road network. Thus, the foremost aim of the cooperative strategy is to generate a desired and feasible traffic status on VMS panels in coordination with traffic signal control system.

3.2. The Two-Stage Nested Optimization Problem. As the mechanism mentioned above, how to find a desirable solution of VMS travel guidance and traffic signal control can be represented as an optimization problem. We name this optimization problem the VMS guidance and signal coordination (VGSC) problem. To solve this problem, traffic signal parameters (i.e., cycle times, green times, and offsets) of

arterials shown on the VMS panels should be optimized simultaneously with these arterials' colors and diversion advice. In other words, all variables in (4) should be optimized during a control period.

Among those variables, the offsets need a special attention to their optimization. For major arterials having a number of signalized intersections, optimal offsets under fixed cycle times and green splits need to be selected from a $(n - 1)$ -dimensional solution space C^{n-1} (where C denotes the cycle time). As a result, it is time-consuming to optimize offsets conjunction with other parameters. For the shortage of optimizing offsets with other parameters simultaneously, a two-stage nested optimization model is formulated.

In the first stage, VMS parameters and arterial signal parameters are optimized in order to minimize the total travel time of travellers on a road network. The first-stage optimization problem can be expressed in

$$\min F_t = \sum_{\forall r,s \in M} \int_0^T f_{rs}(t) \eta_{rs}(t) dt, \quad (8)$$

OD constraints,

$$\text{s.t. } d_{rs}(t) = \sum_{p \in P_{rs}} f_{rsp}(t) \quad \forall r, s, t, \quad (9)$$

$$f_{rsp}(t) \geq 0 \quad \forall r, s, t, p \in P_{rs},$$

VMS constraints in (5)-(6);

Signal control constraints,

$$C_k \in (C_{\min}, C_{\max}), \quad g_k \in (g_{\min}, C_k), \quad \forall k, \quad (10)$$

where $f_{rs}(t)$ is the flow on routes between r and s ; $\eta_{rs}(t)$ denotes the travel time between r and s ; M denotes all O-D pairs on a road network; d_{rs} is the total travel demand between r and s ; p represents path and P_{rs} denotes all paths between r and s . The VMS constraints have been explained above. In addition, the arterial signal optimization owns its specific constraints (see (10)).

In the secondstage, the optimization of Φ_k can be regarded as a nested optimization problem. The Φ_k^{opt} can be searched iteratively under the fixed y_{j+1} . The second-stage optimization problem can be expressed in

$$\min G_t = \sum_{\forall l \in k} \eta_l(t) dt, \quad (11)$$

$$\text{s.t. } \Phi_k \in (0, C_k), \quad (12)$$

where $\eta_l(t)$ denotes the average travel time of link l on the k th path.

4. Simulation-Based Method for the Two-stage Nested Optimization Problem

In this section, we provide a simulation-based-optimization (SBO) method for the two-stage nested optimization problem. A SBO systematic framework is put forward in

Section 4.1. Accordingly, the optimization algorithms used in stage I and stage II are explained in detail (see Algorithms 1 and 2).

4.1. Systematic Framework. To better illustrate the cooperative mechanism, we propose a systematic framework in Figure 2. In view of the complexity of network traffic flow and interactions between different strategies, SBO methods have been introduced to solve a few engineering optimization problems, such as a decision support tool for mitigating traffic congestion [17] and regional signal timing strategies optimization [18].

With multiple decision variables in the two-stage optimization problem, the interdependence between VMS and traffic signal control is hard to evaluate by a traditional optimizer. In addition, a microscopic traffic simulator can give a relatively accurate estimation of flow distribution, average speeds, and interactions of vehicles.

Therefore, we adopt the SBO method to solve the aforementioned optimization problem. In Figure 2, the strategy module concerns operational strategies: VMS travel guidance and traffic signal control. Based on the inputs (e.g., network, demand, and strategies), the simulation module in TRNSIMS can produce detailed outputs including accurate traffic conditions and traffic flow assignment with the driver diversion module. Within the framework, the optimization module is designed to solve the first-stage optimization problem by using genetic algorithms (GAs). And the second-stage optimization problem is solved by dynamic programming method.

The SBO framework aims to obtain the feasible optimal strategies that minimize total travel time of all travellers. And the simulation-based algorithms are explained in detail as follows.

4.2. Stage I: Genetic Algorithm for Joint Optimization. As previously mentioned, the optimization of VMS and TSC is difficult to find optimal solution mathematically. As suggested in [19], biological evolution, the essence of GA, is an appealing source of inspiration for addressing computational problems. In this paper, genetic algorithm (GA) is adopted to reduce computation time. The evaluation function, control variables, constraints, and termination condition are explained as follows.

4.2.1. Evaluation Function. As indicated in (8), the objective function of the first-stage problem is the total travel time of all travellers, which can be directly obtained from the output of microsimulator module in TRANSIMS.

After simulation of each individual, the fitness function is applied to each solution indicating how close it meets the overall specification. Based on cellular automata approach, microsimulator can produce specific information for every traveller. Therefore, the temporal summary of total travel time and average speed over a segment of link can be aggregated in given time increments. It allows us to compute the objective function and constraints.

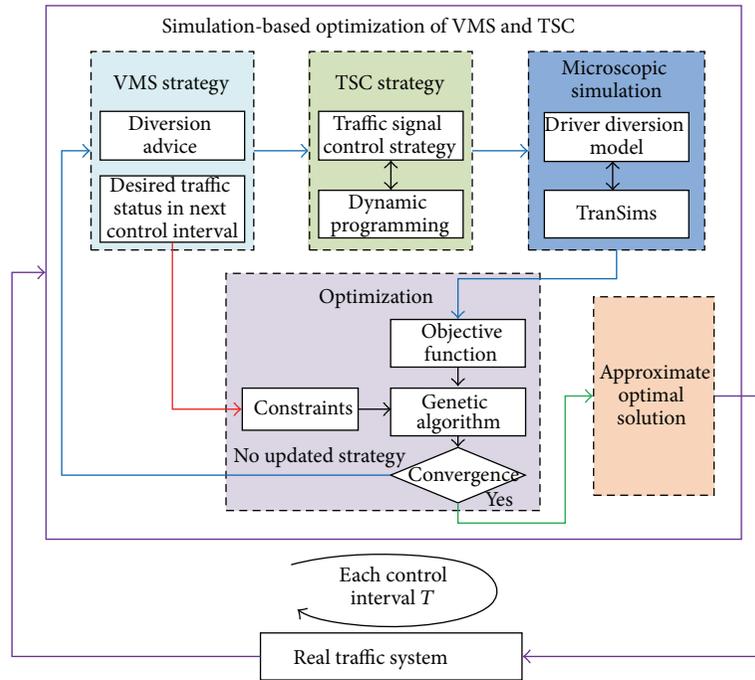


FIGURE 2: A SBO framework for the VGSC Problem.

Initialize $y_{j+1} = [RLR_0, YLR_0, RLR_k, YLR_k, d_k, C_k, g_k]_{j+1}$ as well as $\Phi_k = \{0\}$ (the first time) or Φ_k^{opt} (optimization result in stage II).
DO the following procedure iteratively:
 Find the optimal y_{j+1} by solving the optimization problem;
 Compute of objective function by (8) with respect to the optimal offsets Φ_k^{opt} that is obtained in the second stage;
 Filter solutions by constraints in (8)–(10)
 Update y_{j+1}
Until satisfy the convergence criteria

ALGORITHM 1: The optimization procedures of stage I.

Initialize y_{j+1} as well as Φ_k
DO the following procedure iteratively:
 Find the optimal Φ_k under the condition of y_{j+1} ;
 Compute of objective function by (11)
 Filter solutions by constraints in (12)
 Update Φ_k
Until satisfy the termination criteria
 Return Φ_k^{opt}

ALGORITHM 2: The optimization procedures of stage II.

4.2.2. Control Variables. In a genetic algorithm, a population of candidate solutions is evolved toward better solutions. Each candidate solution has a set of variables (i.e., VMS parameters: RLR_0 , YLR_0 , RLR_k , and YLR_k , d_k ; signal parameters: C_k and g_k).

4.2.3. Constraints. Because of the nature of genetic algorithms, wrapping or truncating individuals in a generation has great influence on optimization performance. In order to deal with the constraints mentioned above without noises, we take advantage of a penalty scheme to the evaluation function. It provides a penalty to the fitness, which is proportional to the constraint violation.

4.2.4. Termination Condition. The iterative process is repeated until a termination condition has been reached. The terminating conditions are as follows.

- (i) Max generation: reach the max generation.
- (ii) Convergence criteria: reach a plateau that successive iterations no longer produce better results.

4.3. Stage II: Dynamic Programming for Offsets Optimization. The second-stage optimization problem was the arterial signal offsets optimization under the condition of fixed green time and cycle time. A recent research by Gartner and Rahul developed a dynamic programming (DP) model which is suitable for signal offsets optimization [20]. On this basis, we further integrate this DP model with simulations in TRANSIMS. In detail, the basic procedure of DP model in a five-intersection arterial is described as follows.

The arterial signalized intersection is denoted as node = 0, 1, 2, 3, 4. Correspondingly, the arterial links are defined as link = 1, 2, 3, 4. For simplicity, we do not show the branches in Figure 3 while taking into account traffic flows of branches when simulating in TRANSIMS. The set of offsets for different nodes is defined as $\Phi = [\Phi_1, \Phi_2, \Phi_3, \Phi_4]$.

Initialization. The DP model needs original input parameters including cycle time (C) and green time (G) and other basic inputs for simulations.

Dynamic Programming. A process of DP is illustrated in Figure 4.

- (1) By setting offset interval δ_{offset} , there are N_j offsets for node j and N_j can be determined by $N_j = C_j/\delta_{\text{offset}}$.
- (2) Every connection in Figure 4 means the average travel time on links. For link j ($j = 2, 3, \text{ and } 4$), offset Φ_j between nodes j and $j - 1$ will be associated with former offset sequence:

$$\Phi_{j-1} = [\Phi_1, \dots, \Phi_{j-1}]. \quad (13)$$

- (3) The average travel time on link j can be further computed by TRANSIMS, which can be denoted as TT_j .

- (4) By comparing this average travel time, we could obtain the offset Φ_j and get a temporary optimized offset sequence from Φ_1 to Φ_j .
- (5) By repeating the above 3 steps for each link, the optimal offset sequence $\Phi_{\text{opt}}^{C,g}$ can be finally determined under the given cycle time and green time.
- (6) A sum of average travel times for the set of links in the arterial is a significant parameter for evaluating the arterial signal control effectiveness. Corresponding to the optimal offset sequence $\Phi_{\text{opt}}^{C,g}$, the minimum travel time can be obtained by microsimulator in TRANSIMS:

$$TT_{\text{min}}^{C,g} = f(\Phi_1, \dots, \Phi_i). \quad (14)$$

5. Application of Methodology

5.1. Network Topology. The proposed methodology is applied in the actual road network of Tianjin Binhai Hi-Tech Industrial Development Park (T.H.I.P, China). This site is located 3 km west to Tianjin urban area, with a two-square-kilometer core area. As a connection to the Tianjin municipal area, road network in this core area is almost filled with heavy traffic flow during rush hours of working days.

In this area, more than 80 percent traffic is due to daily commuters, who work in T.H.I.P. Moreover, traffic management rules and traffic signals are properly maintained, while daily traffic congestions still trouble travelers. Therefore, the road network of T.H.I.P is selected for the proposed study (as shown in Figure 5). It has 30 nodes and 46 unidirectional links. Among them, there are two arterials, Fukang Road (Path 1, 1.97 km, in Figure 6) and Yingshui Road (Path 2, 3.9 km, in Figure 6), which carry the most traffic volume from the Tianjin urban area to T.H.I.P during morning peak hours which remain from 8:00 to 10:00. Besides the links and nodes inside this area, we also set some external zones (blue trapezoids in Figure 6) on the boundary of the road network which are connected with external links to the major traffic flow.

Subsequently, the study site network was coded into TRANSIMS manually by using GIS networks. The network creation required static information including zone, node, link, pocket lane, and vehicle composition. Among them, the vehicle composition contains vehicle type, size, capacity, maximum speed, and acceleration. The traffic signal timing plan for arterial signal control consists of cycle time, green time, yellow time, red time, offset, and phase sequence. And the timing plan can be revised by signal control module dynamically.

If the initial simulation result indicates that default parameters of microsimulator module in TRANSIMS are not acceptable, the parameter calibrations are necessary. In this paper, calibrations for the AM and PM peak time periods were conducted. And the urban arterial network mentioned above was calibrated against field measured traffic count data and travel times by video cameras at signalized intersections.

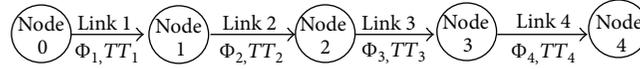


FIGURE 3: Offsets and travel time functions on a five-intersection arterial.

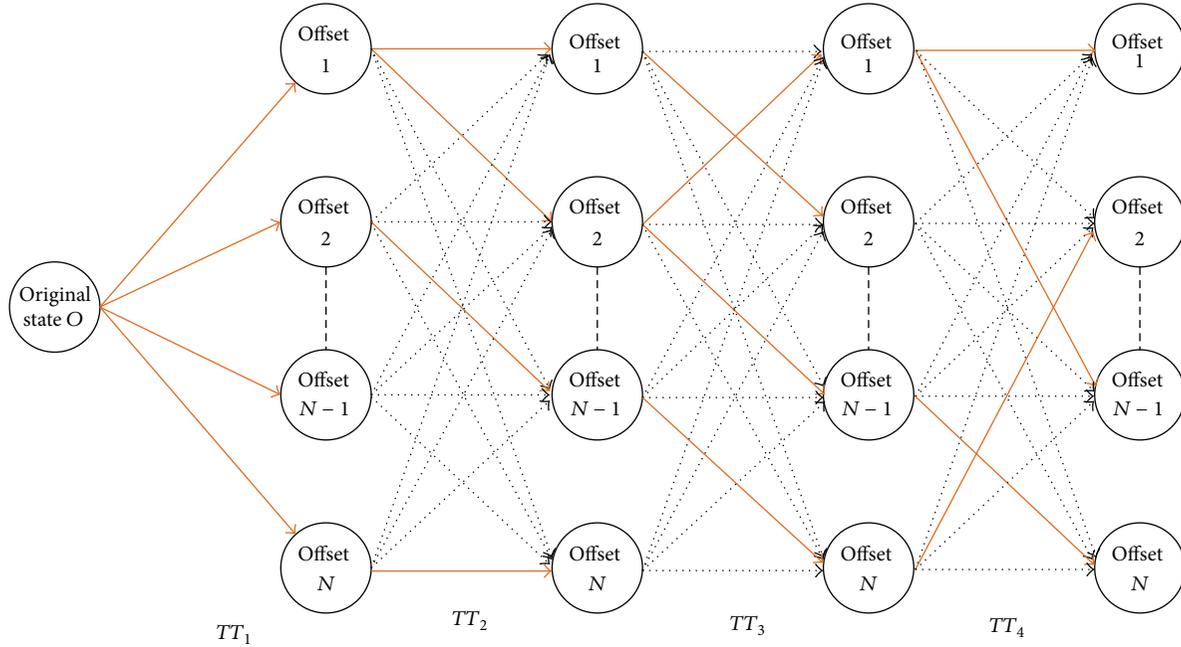


FIGURE 4: An illustration of dynamic programming of offsets.

Based on the basic steps of parameter calibration in TRANSIMS proposed by Park and Kwak [21], we introduce the calibration procedures in detail as follows.

5.1.1. Selection of Calibration Parameters. The first step is the selection of calibration parameters including lane-changing and car-following parameters. In brief, we take the following 5 key calibration parameters as examples.

Maximum Speed. The maximum speed is defined as the 10 km/h above the limited speed.

Slow-Down Probability. The slow-down percentage defines the likelihood that a vehicle will slow down for no apparent reason. The default value is zero (i.e., no random slow-down).

Slow-Down Percentage. The slow-down percentage defines the amount a vehicle will randomly slow down. The default value is zero.

Maximum Waiting Time. The maximum waiting time defines when a vehicle is removed from the simulation. If the vehicle has not moved for this amount of time, the vehicle is removed

from the link and moved to the destination parking lot. The default value is 60 minutes.

Maximum Swapping Speed. To avoid deadlock situations, a cooperative lane swapping concept permits the vehicles to continue their trips. This parameter defines the maximum speed at which lane swapping will be allowed. The default value is 37.5 m/s.

5.1.2. Experimental Design and Multiple Runs. Subsequently, ranges of calibration parameters are determined and multiple sets of calibration parameters are generated. In detail, slow-down probability ranges from 0 to 50%; slow-down percentage ranges from 0 to 50%; maximum waiting time ranges from 60 to 200 minutes; and maximum swapping speed ranges from 0 to 30 m/s. And then 1000 simulation runs are performed for each parameter set.

5.1.3. Feasibility Test. The distribution of simulation results is compared to travel times and volumes generated by field video cameras in order to determine whether current parameter ranges are feasible. As a result, the parameters are calibrated and the results are that the slow-down probability

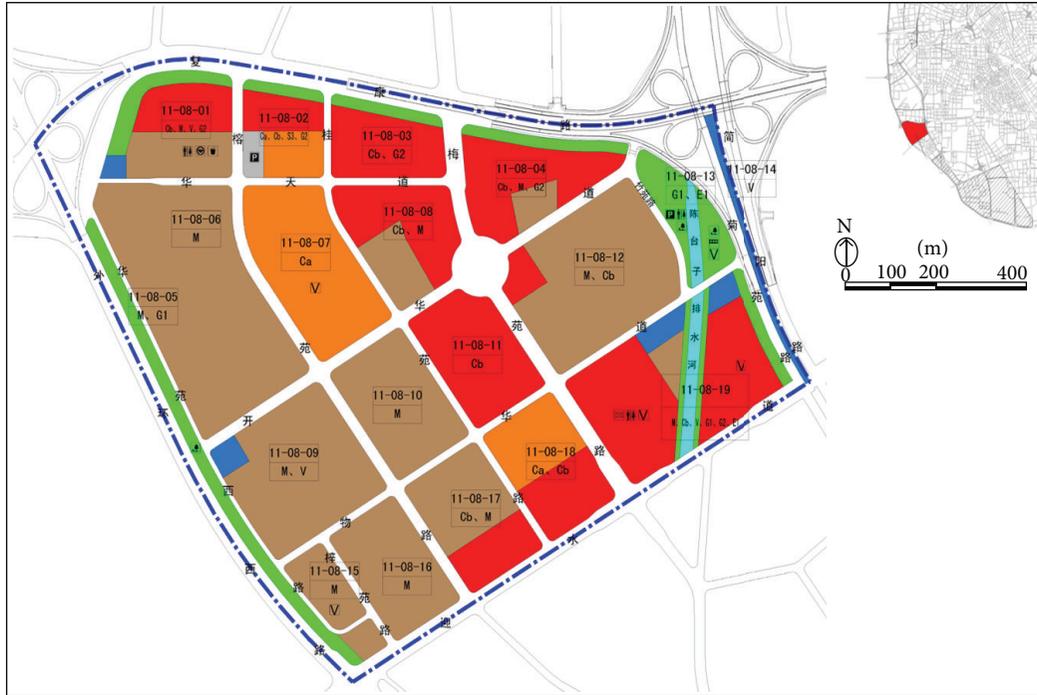


FIGURE 5: Map of T.H.I.P with land use.

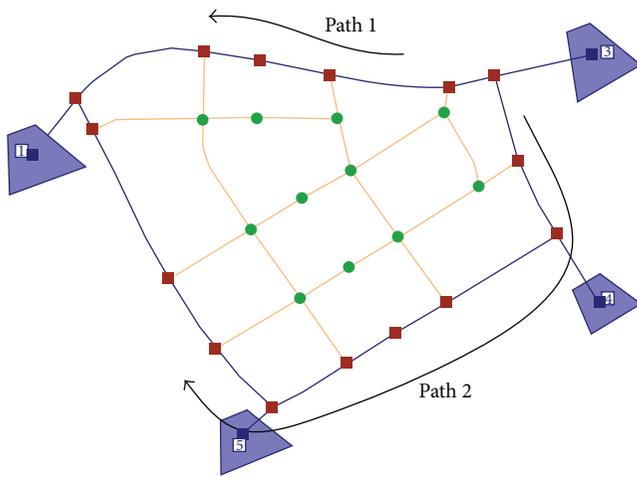


FIGURE 6: Road network topology of T.H.I.P.

is 25.3%, slow-down percentage is 34.5%, maximum waiting time is 83 minutes, and maximum swapping speed is 11.5 m/s.

5.1.4. *Evaluation of Calibrated Parameters.* In the end, it is necessary to conduct 1000 simulation runs to consider the variability of the parameter set.

5.2. *Simulation-Based Optimization Process on TRANSIMS.* TRANSIMS is a travel demand modeling software package that was initially developed by the Los Alamos National Laboratory (LANL) [22]. TRANSIMS separates the simulation process into two stages.

Traffic Assignment. The router module performs traffic assignment using a time-dependent minimum impedance path algorithm based on travel time of links.

Microscopic Simulation. Load travel plans, from the router module, and further computes the time-space information of vehicles second by second.

The open-source software package allows the user to develop new function and customize features of the simulation model. In order to optimize the joint strategy, we develop the driver diversion module, traffic signal control module,

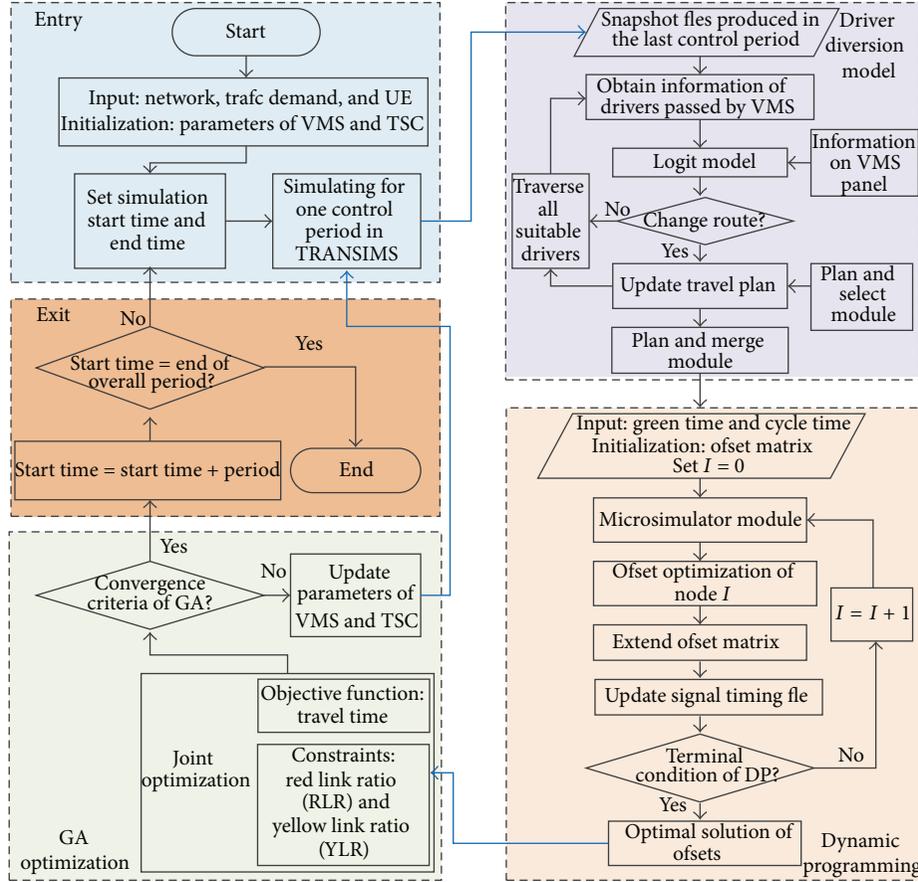


FIGURE 7: Flowchart of simulation program in TRANSIMS.

and optimization module in Python 2.7. Through integrating with postprocessor of TRANSIMS written in C++, the developed three modules can perform various functions. A flowchart in Figure 7 shows the developed procedure in TRANSIMS to solve the joint optimization problem.

At first, the driver diversion module extracts driver space-time information from snapshots files to find drivers who can see the VMS panels. Subsequently, the driver diversion model determines if drivers will make a detour under the influence of VMS information, personal features, and other factors identified previously.

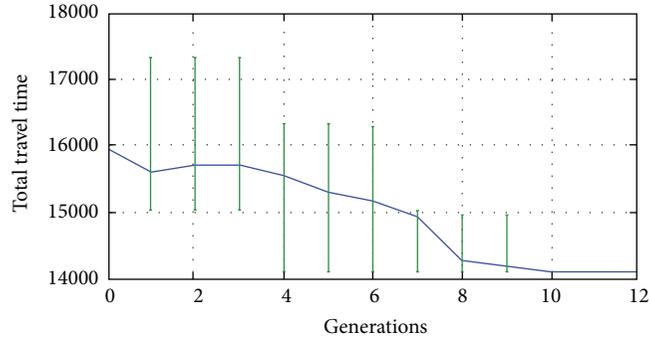
Subsequently, the traffic signal control module can revise signal timing table under the direction of dynamic programming. And total average travel time on arterial links provides a feedback to the signal controllers to evaluate its effectiveness.

The optimization module is developed to evaluate the joint optimization by using genetic algorithm. And the statistical analysis of the simulation results was done to validate the feasibility of solutions and compute the objective value of solutions. The algorithm was implemented in Python using the Pyevolve module [23] and Networkx module [24]. In addition, parallel implementation of the GA is carried out in the optimization process in order to improve the computation

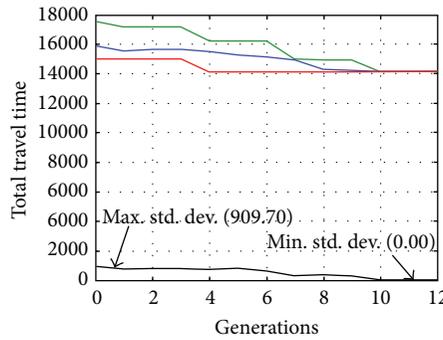
effectiveness. Because of the independence of individuals in a generation, the optimization module can submit population to job manager and evaluate the simulation results after all individuals have completed simulations.

In this network, all 30 nodes are signalized nodes. Among them, 15 nodes on the two arterials (i.e., path 1 and path 2) can be controlled under the direction of optimization module, and other intersections used the fixed timing table. By using the arterial signal control method, we assume that the cycle time and green time for intersections on an arterial are the same. To make faster computation, we further restrict the range of variables' values, as follows.

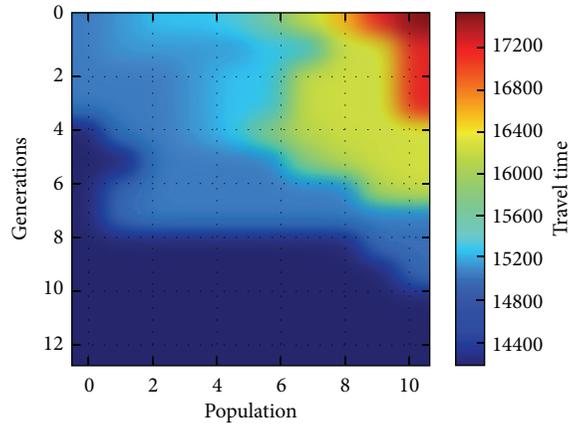
- (1) Cycle times: minimum cycle time C_{\min} is assumed to be 80 s and maximum cycle time C_{\max} is assumed to be 130 s.
- (2) Phase sequence: phase sequence of a given node is kept the same.
- (3) Green time: minimum green time g_{\min} is assumed to be 20 s and the maximum green time is assumed to be $C - g_{\min}$.
- (4) Offsets: minimum green time Φ_{\min} is assumed to be 0 s and the maximum green time is assumed to be C .



(a) The evolution process



(b) The standard deviation of population in generations



(c) Total travel time of population along generations

FIGURE 8: Convergence process of the genetic algorithm.

5.3. *Simulation Results and Analysis.* According to the size of study network, we set the control interval as 5 minutes. And the 12 optimum solutions can be obtained during 8:00-9:00 in the morning peak time. During each interval, 7 variables are optimized in the first-stage GA optimization, including the desired yellow ratios on the two arterials, the cycle times and green times for each arterial, and diversion advice. Moreover the offsets of 10 signalized intersections on the two arterials are optimized by the second-stage dynamic programing.

By the multiruns of the three control methods under traffic demand during the morning peak time, we present the mean values of 20 groups. The comparison of improvement of traffic condition by different strategies is shown in Table 3. The improvement by adopting the active VMS strategy with fixed signal control is 10.2%, while the cooperation of both strategies reaches 26.3%.

The convergence process of each control interval is further explained in Figure 8. Obviously, there exists an obvious

TABLE 3: Comparison of minimum values of objective function.

Control method	Index	
	Objective function values	Improvement
Pretimed signal control without VMS	84756.77	/
Pretimed signal control with VMS strategy	76118.13	10.2%
The cooperative strategy of VMS and arterial signal control	62391.83	26.3%

trend of convergence with the generation increase as shown in Figure 8(a). The maximum number of generations is set to 20; the mutation rate is 0.02; and the crossover rate is 0.9. When the optimization comes to the 10th generation, the standard deviation is close to 0, which satisfies the convergence criteria in GA in Figure 8(b). It can be explicitly displayed that the objective function values of each individual from the 10th generation are kept the same as shown in Figure 8(c).

6. Conclusion and Future Work

In this paper, we study variable message sign (VMS) widely deployed on major arterials in developing countries. In order to appropriately describe drivers' response to VMS, we propose an en route diversion model based on a stated-preference behavioral survey. The vital significance of this new model lies in well considering drivers' responses to graphical road with colors.

On this basis, we propose a cooperative mechanism of VMS travel guidance and arterial signal control for urban traffic management. In this mechanism, traffic control parameters and VMS parameters are optimized together. Particularly for arterial signal control, a two-stage nested optimization problem is formulated. To find the optimal solution, we apply a simulation-based optimization framework for solving this two-stage nested optimization problem. In this cooperative strategy, we allow the communication between VMS and TSC systems to find a desirable and feasible solution during each control interval.

As shown in Table 3, the cooperation strategy managed to reduce total travel time of all travellers when compared to pretimed signal control without VMS strategy and pretimed signal parameters with VMS strategy. Moreover, we examine the convergence process of each interval in Figure 8. It shows that this simulation-based method strikes a very nice balance between accessibility and effectiveness. And the convergence performance is quick and effective.

Besides, it should be pointed out that this simulation-based-optimization method requires relatively high computation source by utilizing microscopic traffic simulation. For future works, we will focus on testing the proposed method for more complex networks of intersecting arterials to investigate effectiveness of the cooperative strategy. In addition,

analytical methods will be researched in this cooperative mechanism to reduce computation costs.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by Hi-Tech Research and Development Program of China under Grant 2014BAG03B04, National Natural Science Foundation of China Grant 51138003, and National Science and Technology Major Project 2012ZX03005016002.

References

- [1] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2065, 2003.
- [2] N. B. Hounsell, B. P. Shrestha, J. Piao, and M. McDonald, "Review of urban traffic management and the impacts of new vehicle technologies," *IET Intelligent Transport Systems*, vol. 3, no. 4, pp. 419–428, 2009.
- [3] A. C. Sutandi, "Evaluation of the impacts of VMS on traffic performance measures in an urban area in Indonesia," *Civil Engineering Dimension*, vol. 10, no. 1, pp. 28–34, 2008.
- [4] S. Peeta, J. L. Ramos, and R. Pasupathy, "Content of variable message signs and on-line driver behavior," *Transportation Research Record*, vol. 1725, no. 1, pp. 102–108, 2000.
- [5] P. Mirchandani and L. Head, "A real-time traffic signal control system: architecture, algorithms, and analysis," *Transportation Research C: Emerging Technologies*, vol. 9, no. 6, pp. 415–432, 2001.
- [6] R. E. Allsop, "Some possibilities for using traffic control to influence trip distribution and route choice," *Transportation and Traffic Theory: Proceedings*, vol. 6, pp. 345–373, 1974.
- [7] M. J. Smith, "The existence, uniqueness and stability of traffic equilibria," *Transportation Research B*, vol. 13, no. 4, pp. 295–304, 1979.
- [8] M. J. Smith, "The existence and calculation of traffic equilibria," *Transportation Research B*, vol. 17, no. 4, pp. 291–303, 1983.
- [9] M. O. Ghali and M. J. Smith, "A model for the dynamic system optimum traffic assignment problem," *Transportation Research B*, vol. 29, no. 3, pp. 155–170, 1995.
- [10] N. H. Gartner and C. Stamatidis, "Integration of dynamic traffic assignment with real-time traffic adaptive control system," *Transportation Research Record*, vol. 1644, no. 1, pp. 150–156, 1998.
- [11] A. S. Abdelfatah and H. S. Mahmassani, "System optimal time-dependent path assignment and signal timing in traffic network," *Transportation Research Record*, vol. 1645, no. 1, pp. 185–193, 1998.
- [12] S. Lu, X. Liu, and Z. Yang, "Study on consistent and integrated traffic management model," *Journal of Wuhan University of Technology, Transportation Science and Engineering*, vol. 32, no. 2, pp. 251–254, 2008 (Chinese).
- [13] H. Cui, *The Coordination Strategy Between Signal Control and VMS and Optimization for information Publishing*, Beijing Jiaotong University, Beijing, China, 2011 (Chinese).

- [14] C. Lin, B. Gong, D. Zhao, and X. Liu, "Traffic control and VMS collaborative technique in sudden disaster," *Journal of Traffic and Transportation Engineering*, vol. 12, no. 6, pp. 104–110, 2012 (Chinese).
- [15] J. J. Cheng and P. E. Firmin, "Perception of VMS effectiveness: a British and Canadian perspective," in *Proceedings of the 12th International Conference on Road Transport Information and Control (RTIC '04)*, pp. 175–185, London, UK, April 2004.
- [16] M. Kraan, N. van der Zijpp, and B. Tuter, "Evaluating network wide effects of VMSs in the Netherlands," in *Proceedings of the 78th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 1999.
- [17] S. H. Melouk, B. B. Keskin, C. Armbrester, and M. Anderson, "A simulation optimization-based decision support tool for mitigating traffic congestion," *Journal of the Operational Research Society*, vol. 62, no. 11, pp. 1971–1982, 2011.
- [18] C. Osorio, *Mitigating network congestion: analytical models, optimization methods and their applications [Ph.D. dissertation]*, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2010.
- [19] M. Melanie, *An Introduction to Genetic Algorithms*, 3, MIT Press, Cambridge, Mass, USA, 5th edition, 1999.
- [20] N. H. Gartner and D. Rahul, "Dynamic programming approach for arterial signal optimization," in *Proceedings of the 92nd Annual Meeting of the Transportation Research Board*, No. 13-0177, 2013.
- [21] B. B. Park and J. Kwak, "Calibration and validation of TRANSIMS microsimulator for an urban arterial network," *KSCE Journal of Civil Engineering*, vol. 15, no. 6, pp. 1091–1100, 2011.
- [22] L. Smith, R. Beckman, and K. Baggerly, "TRANSIMS: transportation analysis and simulation system," Tech. Rep. LA-UR-95-1641, Los Alamos National Laboratory (LANL), Los Alamos, NM, USA, 1995.
- [23] C. S. Perone, "Pyevolve: a Python open-source framework for genetic algorithms," *ACM SIGEVolution*, vol. 4, no. 1, pp. 12–20, 2009.
- [24] A. Hagberg, P. Swart, and D. S. Chult, "Exploring network structure, dynamics, and function using NetworkX," Tech. Rep. LA-UR-08-05495, LA-UR-08-5495, Los Alamos National Laboratory (LANL), Los Alamos, NM, USA, 2008.

Research Article

Research on the Safety Audit Methods for Two-Lane Highway Based on HRV

Zhan-dong Zhu,¹ Ying-chih Lu,¹ Cheng-hong Fu,¹ and Ting Xu²

¹ Department of Transportation, Fujian University of Technology, Fuzhou 350108, China

² School of Automobile, Chang'an University, Xi'an 710064, China

Correspondence should be addressed to Zhan-dong Zhu; zhuzhandong@emails.bjut.edu.cn

Received 15 November 2013; Revised 13 February 2014; Accepted 4 March 2014; Published 31 March 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Zhan-dong Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to explore stable route safety evaluation indicators, this study discusses the relationship between the heart rate variability (HRV) and other factors, such as design speed, operating speed, acceleration rate, and velocity gradient through a lot of experiments. The results show that operating speed coordination and velocity gradient (G_v) outperform others as road safety indicators. Speed coordination evaluation criteria: $|\Delta V_{85}| \leq 10$ km/h, good coordination; 10 km/h $\leq |\Delta V_{85}| \leq 20$ km/h, general coordination, route indicators of adjacent sections should be adjusted, so that the speed difference is not more than 10 km/h; $|\Delta V_{85}| \geq 20$ km/h, poor speed coordination, it needs to readjust the design of the adjacent sections. Speed gradient evaluation criteria: $G_v \leq 0.10$, good road safety; $0.10 \leq G_v < 0.15$, common road safety; $G_v > 0.15$, poor road safety and the adjacent sections need to readjust. The conclusions provide the theoretical reference for highway safety evaluation.

1. Introduction

In order to improve road safety, research abroad explored the relationship between speed statistics and accident rate. Kloeden et al. studied the country road (or national highway) and concluded that the decline of the average speed rather than the decline of velocity gradient contributes to traffic safety [1]. Garber and Gadiraju concluded that the accident rate is proportional to the variation of velocity which increases as the accident rate increases [2]. However, Lassarre in France concluded that the influence of velocity variance on the safety is not significant [3]. Lave found that the influence of velocity variance on accident rate is more significant than that of average velocity [4]. Baruya and Finch studied the relationship between average velocity, velocity gradient, and personal injury accidents and found that average velocity and velocity gradient greatly affect the accident rate [5]. Although the findings discussed above are not consistent and sometimes even contradictory, they indicate that the relationship among average velocity, velocity gradient, and velocity variance to accident rate remains uncertain.

In China, many research results have the same characteristic, which are trying to establish the relationship between

speed and accident number. Unfortunately, the traffic accident data in China was kept by the police departments, so it is not easy to obtain original accident data. Even if accident data can be gained, it cannot meet needs of research. The data from the traffic police department on the cause of the accident, morphology, and classification are not the same with the researchers, that is, not consistent with standard statistical accident. In addition, in recent years, China is in the large-scale construction period, and the roads expansions are very frequent, resulting in shorter service life of road and the lack of stable long-term accident statistics. So the convincing results are scarce, and it is difficult to evaluate the safety of a new road.

Considering the above issues, this study investigates the relationship between road safety and speed statistics, with physiological aspects on drivers through a series of experiments.

2. Materials and Methods

A transportation network consists of human beings, vehicles, roads, and traffic control devices. Any change of these factors

will affect drivers' reactions. While the external reaction may be observed by abnormal driving behavior (such as acceleration), the inner one appears in driver's psychological changes (such as HRV). Therefore, the studies on road safety associated with driving psychological changes become popular in recent years.

Brookhuis and de Waard studied the relation between the driver's psychological pressures and ECG and EEG changes, respectively, in the crowded city ring Expressway and the free flow of Expressway and concluded that psychological pressure can be obtained in the ECG and EEG responses [6, 7]. Pan studied the relationship among drivers' heart rate, blood pressure, and radius of horizontal curve, curve length, curve angle, and so forth in the mountain road test road in a certain speed [8, 9]. Qiao studied the key parameters of the two-lane highway in mountainous area based on driver factor analysis (parameters of what) and put forward the design parameters to enable driving comfortably on two-lane highway in mountainous area [10]. Related literatures selected the changes of the driver real-time heart rate variability, namely, HRV to indicate workload of the drivers and divided its threshold [11–15]. HRV is a balanced and effective detection index between the characterization of sympathetic and vague nerves, measures the continuous sinus heartbeat interval (instantaneous heart rate) of the tiny fluctuations of ECG changes, and is more sensitive to describe physiological index of the driver [16, 17]. Wang et al. proposed a composite measure consisting of three physiological measures, facial skin temperature, eye blinks, and pupil dilation [18]. J. Lee and S. Lee develop a model for evaluation of highway alignment design with a new evaluation index as vehicle spacing [19]. The above literatures studied the traffic safety problems based on EEG and ECG reactions of the drivers, and conclusions show that there was some relationship between the driver's physiological reactions and traffic safety and route parameter.

Because the experiments were all drawn in the fields, taking into account that the EEG anti-interference ability was poor, the heart rate variability (HRV) was chosen as the index. The experiments were drawn to study the stable and reliable relationship between the driver of HRV and different safety evaluation index of the road. Combination of all researches and road safety evaluation indicators include speed consistency, speed coordination, speed reduction coefficient, and speed gradient.

2.1. Data Collection. The devices for the experiments include a test vehicle, high precision GPS, KF2 dynamic multiple physiological parameters detection apparatus (KF2), and a camera. The test vehicle is Santana 2000, and the high precision GPS instrument selected Novatel dynamic GPS to acquire speed change and coordinate continuous velocity and acceleration and other related indicators, corresponding to the road alignment. KF2 dynamic multiple physiological parameters detection can obtain the drivers' HRV indicator at any time, and road traffic environment is recorded by camera, and manual recording special conditions and data processing should be removed without the free-flow speed data. Test sections selected two-lane highways in Jiangxi

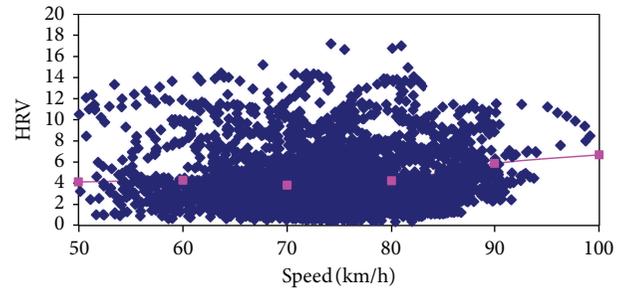


FIGURE 1: Speed versus heart rate variability.

Yichun Anfu, Xinjiang Kuitun-Karamay, Xinjiang alar-hotan, and Xinjiang Akesu-Kuche. Yichun Anfu section is located in the mountainous area, and the remaining roads are in plain area. The total mileage of test sections is about 800 km. The sample size of each type is greater than 50.

At the beginning of experiments, the basic information of drivers, the installation of test equipment, including KF2, GPS, and camera, and the time of opening KF2 should be recorded. After installing the device, drivers were required to obtain a calm state. Then the driver drove on the test section and turned on the GPS and camera to record driving conditions and road traffic environment until the end of the test. The records should be marked for the special moments point and the traffic environment change time point (section) for further data processing. In order to eliminate noises, data for 3 minutes before and after tests are excluded.

2.2. Relationship between Road Safety

Evaluation Indicators and HRV

2.2.1. Design Speed Consistency and HRV. At present in the routine design, design speed is still selected as the control index in China. For the same type highways of terrain, the design speed is a constant value. For example, on two-lane highways in plain areas, it needs to use the design speed of 80 km/h, and 60 km/h in the mountainous areas. The benefits of design speed concept are easy to use. According to the "Guidelines for Safety Audit of Highway," design speed consistency is defined as the absolute value of the difference between the operating speed and design speed. Figure 1 shows a scatter plot between the speed and the driver of HRV test vehicle in two-lane highway.

In Figure 1, design speed increasing, namely, the design speed consistency increasing, the drivers' HRV trend is not obvious, and the distribution range has no significant diffusion difference. There is not a similar trend between design speed consistency and HRV did not show an obvious consistent trend. It is concluded from psychological results that it is not inappropriate to select design speed consistency as a road safety evaluation index. Drivers tends to prefer a higher operating speed because of a better roadway alignment, while poor road alignment forces drive at lower speed, and only in poor route linear design speed will play a role. The minimum value, in the route design specifications only limits the design standard for the most unfavorable

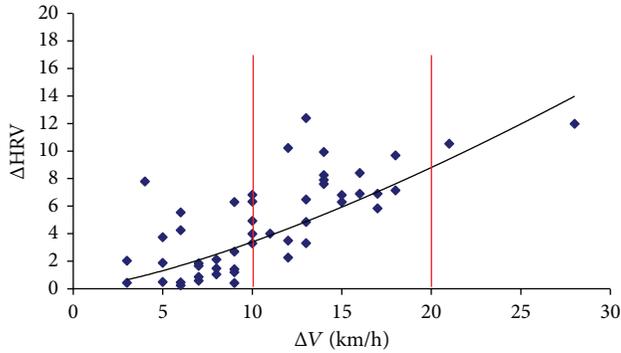


FIGURE 2: Operating speed coordination versus Δ HRV.

situation. For two-lane highways of the western plains region, average speed of good road sections is much higher than the design speed. The increase of speed consistency does not bring a corresponding increase in the accident rates. It is also the same in some foreign speed limit sections. From the experimental results it is illustrated that it is not appropriate to choose the design speed consistency to measure road safety.

2.2.2. Operating Speed Coordination and HRV. The operating speed coordination is the absolute value of the operating speed difference between adjacent sections. The sections are divided and grouped according to the different design elements within the section. Firstly nonfree flow data is removed. Then the whole road is divided into small sections as the straight section of the circular curve sections, longitudinal slope sections, and curved section in accordance with the Guideline. Sections with various features are found out and the corresponding speed value at the right time is selected from GPS. Thirdly, the operating speed coordination is calculated by speed difference between the adjacent sections. HRV data is obtained by KF2 from physiological indicators of the corresponding time. Δ HRV and the ΔV data are then obtained. ΔV is the operating speed difference between the adjacent sections and Δ HRV is the HRV difference between the adjacent sections. The relationship between operating speed coordination and Δ HRV is shown as Figure 2.

In Figure 2, Δ HRV increases significantly when the operating speed coordination is increasing. When operating speed coordination is less than 10 km/h, the average value of Δ HRV is about three. When operating speed coordination is enlarged to 10–20 km/h, the average Δ HRV grows up to about 7. And when operating speed coordination expanded to more than 20 km/h, HRV values are all greater than 10. Results show that the speed of coordination is directly proportional to Δ HRV.

During the driving process, changes of adjacent operating speed will lead to the change of HRV, and especially the dramatic change of adjacent running speed can cause strong reaction of drivers' psychology. Therefore it is appropriate to select the operating speed coordination to measure the safety of the highway and it is also secure to select 20 km/h as an evaluation threshold in the Guideline.

2.2.3. Speed Reduction Coefficient and HRV. Drivers will change acceleration or deceleration transformation according to the environment and road conditions during the running. According to statistics, the speed acceleration of the vehicle in the process of smooth driving generally remains between -0.15 and 0.15 m/s^2 . If the speed acceleration is due to good traffic environment and roadway conditions, it is usually considered to be a safe behavior. In contrast, deceleration process, in particular severe deceleration, is often considered to be not safe. It means that a sudden change of road conditions occurs, and if the deceleration behavior cannot meet the road traffic needs, traffic accidents will occur. The statistics results show that traffic accidents account for a large proportion of accidents. Removing the nonfree-flow speed data, figuring out the speed and acceleration in the process of moving test vehicle, relationship between the acceleration and deceleration acceleration and HRV is analyzed, respectively. And the results are shown in Figure 3.

Derived from Figure 3, in the process of acceleration (Figure 3(a)), there is little correlation between the driver HRV and the acceleration of the vehicle. As the acceleration increases, HRV does not change significantly, reflecting the smaller driver psychological changes. It follows that the acceleration process is safe. This is consistent with the conclusion of the above analysis. In the process of deceleration (Figure 3(b)), as the acceleration increases, HRV also increases rapidly. When the acceleration is greater than 1.5 m/s^2 changes in HRV will be more intense, reflecting that the sudden deceleration behavior can result in strong psychological reaction. And the conclusion is consistent; namely, the deceleration process especially the sudden deceleration process is insecure.

During acceleration, the acceleration value of the driver is generally less than 1.0 m/s^2 , and the maximum acceleration is up to 1.5 m/s^2 . But during deceleration process, the acceleration value of the driver is generally below 1.5 m/s^2 , and the maximum acceleration is about 2.5 m/s^2 . It shows that the response of drivers' deceleration is bigger than that of acceleration. In the course of the research on deceleration, according to Moscow State University conclusions, the acceleration process is divided into the three intervals $0-0.5 \text{ m/s}^2$, $0.5-1.5 \text{ m/s}^2$, and $1.5-2.5 \text{ m/s}^2$. The relationship between speed reduction coefficient and the accident rate was studied in the three cases, and we came to the conclusion that safety would be greatly reduced when the acceleration was more than 1.5 m/s^2 . Deceleration greater than 1.5 m/s^2 is defined as the sudden deceleration behavior, and is often caused due to sudden changes in the road. It is very unsafe.

Choosing adjacent sections in which deceleration is greater than 1.5 m/s^2 , and figuring out the speed reduction coefficient (the speed difference between first section and the latter sections and its own ratio), get the relationship between speed reduction coefficient and HRV, as shown in Figure 4. From the diagram, HRV increases with speed reduction coefficient increasing. When the speed reduction coefficient is more than 0.25, which is the speed of a sudden decrease of more than 25%, HRV changes rapidly.

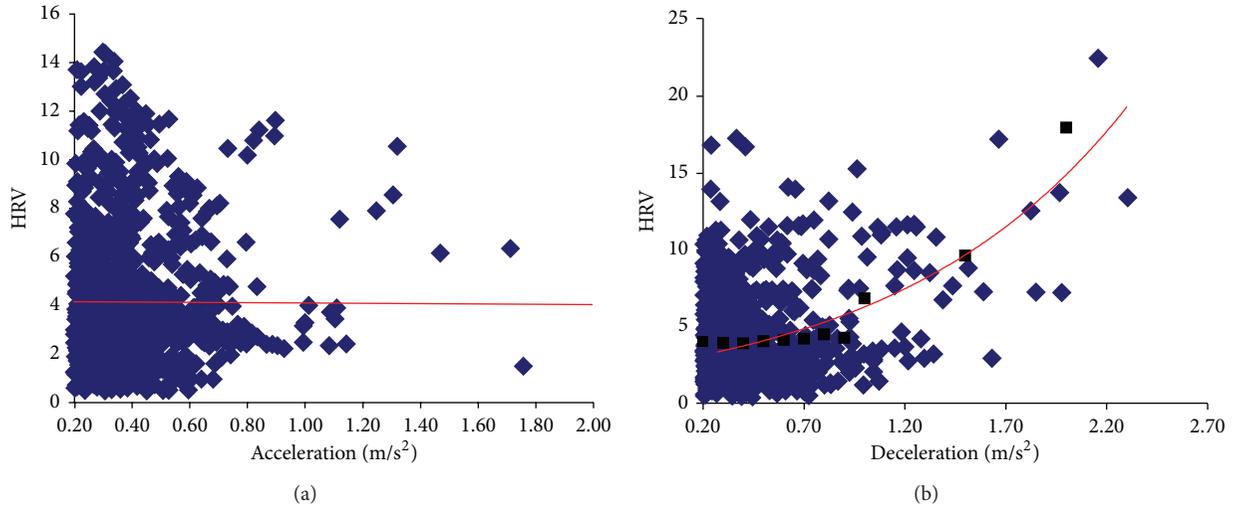


FIGURE 3: Relation between acceleration and HRV.

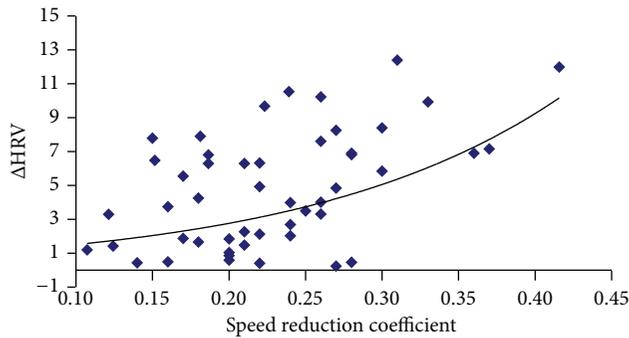
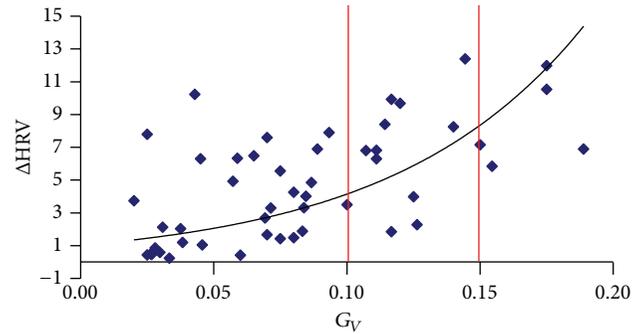
FIGURE 4: Relation between speed reduction coefficient and Δ HRV.

FIGURE 5: Relation between speed gradient and HRV.

From Figure 4, it is concluded that it is appropriate to measure road safety with speed reduction coefficient. Because trend of HRV is more obvious when speed reduction coefficient of the adjacent sections is more than 0.25, and then 0.25 is selected as the critical value (threshold) of the speed coefficient. Compared to the speed at the feature points of previous sections, if the speed coefficient of the latter section reduced more than 0.25, it is considered that the routine design has potential security risks and needs to be improved.

2.2.4. Speed Gradient and HRV. The above analysis shows that operating speed coordination and speed reduction coefficient are closely related with the change of driver's HRV. Both indicators can be measured by the amount of change in operating speed of adjacent sections; the two conclusions are certainly consistent. However, there exists a common defect that they cannot measure how much operating speed changes in alignment curve in unit length sections.

For example, a straight section is followed by a circular curve segment and roads operating speed difference is 15 km/h. the psychological reactions of the driver are different on different lengths of the curve. For drivers, the 100-meter radius circular curve may lead to a decline at a

dramatic rate and a 300-meter radius circular curve may lead to a slower decrease. Although two sections from the coordination of all belong to the general safety section, the previous is clearly larger than the latter on the driver's sudden reaction reflection. Therefore, an index named speed gradient is constructed to describe the above differences as follows:

$$G_v = 2 \times \frac{\Delta V_{85}}{L}, \quad (1)$$

where G_v is the speed gradient; L is curve length, assuming that the deceleration of the vehicle completes in the first half of the curve.

After removing non free-flow data, deceleration speed, which is used to study the relationship between speed gradient and HRV, is calculated by speed of adjacent sections. Statistical analysis is shown as in Figure 5.

It can be obtained from Figure 5 that as the velocity gradient increases, HRV is also gradually increasing, and both approximate logarithmic. When the speed gradient is less than 0.1, HRV growth trend is more gentle; when the speed gradient is over 0.10, HRV increasing trend has accelerated very fast; and when the velocity gradient exceeds 0.15, HRV increasing trend is very fast, reflecting strong

psychological reactions of drivers. It shows that there is a strong correlation between the velocity gradient changes and changes of HRV.

3. Results and Discussion

It can be concluded from the above analysis that the changes in speed gradient and HRV are closely related and HRV increases as speed gradient increases. Speed gradient can be divided into three intervals: less than 0.10, 0.10~0.15, and more than 0.15, three security descending order. And greater than 0.15 speed gradients in the design should be avoided.

In summary, correlation is not strong between design speed consistency and HRV. From the point of view of drivers' physiological responses, it is not suitable to be regarded as a safety evaluation indicator; speed coordination is closely correlated with HRV. 20 km/h as the critical threshold of safety evaluation is also appropriate, and guideline about the speed coordination regulation is reasonable. During the deceleration of the vehicle, the speed reduction coefficient and the deceleration of the adjacent sections and the security have close relationship. Speed reduction coefficient is more than 0.25 and the deceleration greater than 1.5 m/s^2 is considered a potential safety risk; velocity gradient and HRV changes are closely related. According to the analysis the speed gradient should be divided into three groups as less than 0.10, 0.10~0.15, and more than 0.15, and the speed gradient of greater than 0.15 in the design should be avoided.

Because the speed reduction coefficient and operating speed coordination are similar, in order to maintain consistency with the guideline, operating speed coordination is chosen as one audit index. The speed gradient measures road safety from the aspects of unit length of road section, avoiding the defect that operating speed coordination is the unique index. In general, the speed coordination measures the speed variation amplitude of adjacent sections, and the speed gradient measures how fast of the change of the adjacent section, which can be used in combination to evaluate the safety of route.

Operating speed coordination $|\Delta V_{85}|$ is the speed difference of adjacent sections. Speed coordination $|\Delta V_{85}|$ is evaluation standard: when $|\Delta V_{85}| \leq 10 \text{ km/h}$, good route design; $10 \text{ km/h} < |\Delta V_{85}| < 20 \text{ km/h}$, common route design, when conditions permit, adjacent sections of technical indicators should be appropriately adjusted, in the condition that the operating speed difference is less than or equal to 10 km/h; $|\Delta V_{85}| \geq 20 \text{ km/h}$ means bad route design and the design of adjacent sections need to be adjusted.

The speed gradient evaluation standard: when $G_v \leq 0.10$, road safety is good; $0.10 \leq G_v < 0.15$, road safety is common, when conditions permit adjacent sections of technical indicators should be appropriately adjusted; $G_v > 0.15$, road safety is poor and route indicator of adjacent sections need to be adjusted.

4. Conclusions

The paper studied the relationship among drivers heart rate variability of the vehicle in the running process and

design speed consistency, operating speed coordination, speed reduction coefficient, and speed gradient. Then their relationship with safety is obtained from speed statistics. From analysis, operating speed coordination and the speed gradient can be chosen as indicators to measure traffic security. Conclusion provides a strong theoretical support for road safety evaluation.

Conflict of Interests

The authors declared that they have no conflicts of interests to this work.

Acknowledgments

The project is supported by National Natural Science Foundation of China (no. 51308058) and China Postdoctoral Science Foundation (no. 2013M532007).

References

- [1] C. N. Kloeden, G. Ponte, and A. J. McLean, "Traveling speed and the risk of crash involvement on rural roads," Road Accident Research Unit CR 204, Adelaide University, 2003.
- [2] N. J. Garber and R. Gadiraju, "Factors affecting speed variance and its influence on accidents," *Transportation Research Record*, vol. 1213, pp. 64–71, 1989.
- [3] S. Lassarre, "The introduction of the variables traffic volume, speed, and belt-wearing into a predictive model of the severity of accidents," *Accident Analysis and Prevention*, vol. 18, no. 2, pp. 129–134, 1996.
- [4] C. Lave, "Speeding, coordination, and the 55-mph limit," *The American Economic Review*, vol. 75, no. 5, pp. 1160–1164, 1985.
- [5] A. Baruya and D. J. Finch, "Investigation of traffic speeds and accidents on urban roads," in *PTRC European Transport Forum*, pp. 219–230, University of Warwick, 1994.
- [6] K. A. Brookhuis and D. de Waard, "The use of psychophysiology to assess driver status," *Ergonomics*, vol. 36, no. 9, pp. 1099–1108, 1993.
- [7] R. Inoue, "Safety measures against construction accidents in kanto regional development bureau," *Public Works Management Journal*, pp. 9–12, 2002.
- [8] X.-D. Pan, *Application of Human Information Technology in Road Safety Evaluation of Geometric Structure*, Transportation College of Tongji University, 2002.
- [9] S. Cafiso, G. La Cava, and A. Montella, "Safety index for evaluation of two-lane rural highways," *Transportation Research Record*, vol. 2019, pp. 136–145, 2007.
- [10] J.-G. Qiao, *Study on the Key Parameter of Two-Lane Highway in Mountain Area Based on Driver's Factors*, The College of Architecture and Civil Engineering, Beijing University of Technology, 2006.
- [11] R. Bird and I. Hashim, "Exploring the relationship between safety and the consistency of geometry and speed on British Roads," in *Proceedings of the 85th Annual Meeting Compendium of Papers*, Transportation Research Board, National Research Council, 2006.
- [12] D. H. Harwood, F. M. Council, E. Hauer, W. E. Hughes, and A. Vogt, "Prediction of the expected safety performances of

- the rural two-lane highways,” Federal Highway Administration, U.S. Department of Transportation RD-99-207, 2000.
- [13] X. Cao, *Study on the Highway Longitudinal Alignment Safety Evaluation Based on Drivers Workload*, The College of Architecture and Civil Engineering, Beijing University of Technology, 2009.
- [14] K. Fitzpatrick, P. Carlson, M. Brewer, and M. D. Wooldridge, “Design speed, operating speed, and posted speed limit practice,” in *Proceedings of the 82nd Annual Meeting Transportation Research Board*, 2003.
- [15] J. Hu and J. Wang, “Safety evaluation of freeway horizontal curve based on driver workload,” *Journal of Transportation Systems Engineering and Information*, vol. 12, no. 1, pp. 31–37, 2012.
- [16] M. A. Brewer, G. Pesti, and W. Schneider, “Improving compliance with work zone speed limits: effectiveness of selected devices,” *Transportation Research Record: Journal of the Transportation Research Board*, no. 1948, pp. 67–76, 2006.
- [17] G. Kanellaidis, “Human factors in highway geometric design,” *Journal of Transportation Engineering*, vol. 122, no. 1, pp. 59–66, 1996.
- [18] L. -M. Wang, V. G. Duffy, and Y. Du, “A composite measure for the evaluation of mental workload,” in *Digital Human Modeling*, pp. 460–466, 2007.
- [19] J. Lee and S. Lee, “Development of model for highway design consistency evaluation with artificial neural network,” in *Proceedings of the 13th International conference on Multimedia Modeling (MMM '07)*, vol. 4352, Part II, pp. 620–626, Lecture Notes in Computer Science, 2007.

Research Article

The Theory of Dynamic Public Transit Priority with Dynamic Stochastic Park and Ride

Chengming Zhu,^{1,2} Yanyan Chen,¹ and Changxi Ma³

¹ Transportation Research Center, Beijing University of Technology, Beijing 100124, China

² College of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo, Henan 454000, China

³ School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China

Correspondence should be addressed to Chengming Zhu; zhuchengming@hpu.edu.cn

Received 2 December 2013; Revised 24 January 2014; Accepted 10 February 2014; Published 12 March 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Chengming Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Public transit priority is very important for relieving traffic congestion. The connotation of dynamic public transit priority and dynamic stochastic park and ride is presented. Based on the point that the travel cost of public transit is not higher than the travel cost of car, how to determine the level of dynamic public transit priority is discussed. The traffic organization method of dynamic public transit priority is introduced. For dynamic stochastic park and ride, layout principle, scale, and charging standard are discussed. Traveler acceptability is high through the analysis of questionnaire survey. Dynamic public transit priority with dynamic stochastic park and ride has application feasibility.

1. Introduction

There is a fact that public transit cannot compete with car and travelers are not willing to initiatively reduce car use. Public transit travelers make effort for relieving traffic congestion, but they do not benefit from their behaviors. Car travelers increase traffic congestion degree and they benefit from their behaviors. Because of the above reason, travelers are reluctant to actively cooperate with government and traffic administration. For travel behavior intervention, there is a large gap between actual effect and expected effect. Some tough intervention measures on car use are adopted for relieving congestion, but tough intervention measures are easily to cause traveler's antipathy and cannot guide traveler to initiatively reduce car use.

The effective travel intervention measure should be able to make public transit to compete with car. Also, it should be able to make government, traffic administration, and public transit traveler (including travels who initiatively reduce car use) to compete with traveler who do not reduce car use. Therefore, it is needed that public transit traveler are willing to actively cooperate with government and traffic administration. The premise of this cooperation is that public transit

travelers benefit from choosing public transit. But at present time, public transit traveler cannot benefit from choosing public transit because of very low comfort level and low carrying speed. Therefore, other travel behavior intervention model is needed to be studied based on enhancing public transit travel service level. It should increase public transit competitiveness with car and it should ensure that travelers are willing to accept it. It can promote traveler actively to cooperate with government and traffic administration. The start point of this intervention model is based on the fact that public transit travel cost is not higher than car travel cost.

2. The Existing Problem

2.1. The Existing Problem of Current Public Transit Priority. There are many researches on public transit priority from a different point. Tirachini et al. [1] have studied restating modal investment priority with an improved model for public transport analysis. Eichler and Daganzo [2] have researched Bus lanes with intermittent priority. Koehler and Kraus Jr. [3] have studied simultaneous control of traffic lights and bus departure for priority operation. Sun et al. [4] have studied bus detection based on sparse representation for transit signal

priority. Wahlstedt [5] analyzed the impacts of bus priority in coordinated traffic signals. Liu et al. [6] proposed an Elastic analysis on urban public transport priority in Beijing. Zyryanov and Mironchuk [7] studied intermittent bus lane and bus signal priority strategy by simulation. Vedagiri and Arasan [8] discussed on estimating modal shift of car travelers to bus on introduction of bus priority system. The above researches play an important role on public transit priority, but some problems exist. The above researches mainly focus on public transit priority measures, but the research on public transit priority level for specific city is absent. The researches on public transit priority are not discussed from the point of public transit travel cost comparing with car travel cost, which results in low acceptance level of car user on public transit. Then the expected effect of public transit priority cannot be achieved.

Public transit priority policy has been implemented in China for more than ten years, it enhances public transit travel proportion in a certain extent, but traffic congestion is not relieved and car travel proportion during peak period is still increasing. Instead, public transit priority attracts slow travel group to public transit. The nature reason of this result lies in that public transit travel service level is far lower than cars during period. There is no reference point for public transit priority and what priority extent should be achieved for specific city is not discussed, which directly results in no goal of public transit priority and little effect of easing traffic congestion. The key reason is that how to attract car traveler to choose public transit is not considered from traveler point. Simply low fare is proved to be insufficient. Therefore, public transit travel service level should be enhanced through public transit priority. Since passenger crowd level is difficult to be eased in a short period, enhancing carrying speed of public transit is an important goal of public transit priority. Also, carrying speed enhancement for public transit is relative to car carrying speed, rather than simply comparing with carrying speed of public transit itself. It is obvious that public transit carrying speed should dynamic change according to car carrying speed and actual traffic volume. Therefore, dynamic public transit priority theory is needed.

2.2. The Present Situation and Existing Problem of Intervention on Car Use. The intervention on car use includes policy intervention and driving behavior intervention. For driving behavior intervention, Wang et al. [9, 10] discussed model-based digital driving dependability and safety-based behavioural approaching model. This study focuses on policy intervention. The current policy intervention measures include restriction pass, congestion charging, and increasing parking fee. These measures play a certain role to ease traffic congestion in a short period, but they are not suitable for long period because they are difficult to make traveler to initiatively voluntarily cooperate with traffic administration. Borjesson et al. [11] analyzed the Stockholm congestion charges, 5 years on effects, acceptability, and lessons learnt. But current research neglects the study of traveler acceptance on intervention measures. Traveler has low acceptability on intervention measures and is easy to form psychological resistance. These measures are not analyzed from convenient

travel perspective. Also, because public transit service level is too low, there is no travel mode which car traveler can accept when reducing car use is required or car use is restricted in certain time interval. Therefore, other intervention measures on car use with higher traveler acceptability should be studied and dynamic public transit priority combined with dynamic stochastic park and ride is an intervention measure with high studying value.

2.3. The Present Situation and Existing Problem of Park and Ride. There are many researches on park and ride from a different point. Holguin-Veras et al. [12] studied user rationality and optimal park-and-ride location under potential demand maximization. Aros-Vera et al. [13] discussed p-Hub approach for the optimal park-and-ride facility location problem. Farhan and Murray [14] used a multiobjective spatial optimization model for siting park-and-ride facilities. Meek et al. [15, 16] analyzed UK local authority attitudes to park and ride and evaluating alternative concepts of bus-based park and ride. Kepaptsoglou et al. [17] analyzed optimizing pricing policies in park-and-ride facilities: a model and decision support system with application. Qin et al. [18] analyzed park-and-ride decision behavior based on Decision Field Theory. Hounsell et al. [19] studied enhancing park and ride with access control.

The above researches focus on the layout around rail transit station of urban peripheral or some public transit hub. This layout model plays a certain role for reducing car volume of driving into city center, but it is not conducive for traveler to stochastic selecting park and ride according to their travel demand. Therefore, it can influence the proportion of actual park and ride to possible park and ride. The layout and scale of park and ride should be hierarchical and classified in order to be adapted to dynamic stochastic park and ride demand for car traveler.

3. Dynamic Public Transit Priority Theory

3.1. The Connotation of Dynamic Public Transit Priority. Dynamic public transit priority dynamically adjusts space-time resources for public transit according to road traffic volume and saturation, public transit vehicle volume, and intersection saturation. It includes public transit lane allocation on road, intersection entrance lane, and pass time allocation for public transit at intersection. Its purpose is to ensure that public transit carrying speed is so high that public transit can compete with car. When travelers select public transit, their profits are not lower than car traveler, or public transit travel cost is not higher than car travel. The connotation of dynamic public transit priority can be concretely expressed as follows.

- (1) For unit travel distance, the proportion of public transit travel profit to cost is not lower than car travel.
- (2) For same travel distance, public transit travel cost is not higher than car travel cost.

Dynamic public transit priority theory is conducive to promote car traveler to reduce car use and shift to public transit.

3.2. The Level of Dynamic Public Transit Priority. The level of dynamic public transit priority refers to what real-time dynamic priority level should be provided for public transit according to dynamic traffic demand under certain road traffic facilities conditions. That is how to dynamically allocate time and space for public transit priority pass. At the same time, dynamic public transit priority is relative to car travel service level (an average carrying speed of car). The purpose of dynamic public transit priority is to realize that public transit travel cost is not higher than car travel cost. Therefore, the key of dynamic public transit priority is to determine the proportion of public transit carrying speed to car carrying speed. Because car carrying speed is influenced by the saturation of road and intersection, the carrying speed of public transit dynamic changes too. This is the essential connotation of dynamic public transit priority.

3.2.1. The Equilibrium Point of Public Transit Priority Carrying Speed (The Minimum Carrying Speed). Judging current public transit priority, key factor influencing public transit priority effect is carrying speed. Enhancing public transit carrying speed is a powerful measure for enhancing public transit competitiveness. Therefore, dynamic public transit should ensure that public transit carrying speed is higher than car carrying speed. Because car speed is influenced by traffic saturation, the carrying speed of dynamic transit priority also should be depended on specific traffic saturation.

Assume that public transit travel and car travel have same travel OD and approximate same travel path. That is travel distance of two travel modes is approximately equal. From the point of travel cost, what carrying speed should be provided for public transit is analyzed under certain road-traffic facilities and traffic saturation.

The travel cost of choosing public transit is C_P . It includes time cost C_P^t (including walking time of arriving at station, away from station, transfer, and waiting time) and ticket cost C_P^e , the energy consumption cost caused by the crowd in public transit and energy consumption cost per unit time C_P^n , without considering energy consumption cost of walking (this walking can be regarded as a kind of physical training).

The travel cost of choosing car is C_C . It includes time cost C_C^t , fuel cost C_C^f , and parking cost C_C^p .

Note that travel distances are L , walking distance of choosing public transit is L_W , walking speed is V_W , public transit carrying speed is V_P , car carrying speed is V_C , and fuel cost of per km is C_0^f .

For the same traveler, per time cost of choosing public transit and car is equal. Assume that Per time cost is C_0^t , then travel cost of choosing public transit is as shown in (1):

$$\begin{aligned} C_P &= C_P^t + C_P^e + C_P^n \cdot \frac{L - L_W}{V_P} \\ &= C_0^t \cdot \left(\frac{L_W}{V_W} + T_W + \frac{L - L_W}{V_P} \right) + C_P^e + C_P^n \cdot \frac{L - L_W}{V_P}. \end{aligned} \quad (1)$$

Travel cost of choosing car is as shown in (2):

$$C_C = C_C^t + C_C^f + C_C^p = C_0^t \cdot \frac{L}{V_C} + C_0^f \cdot L + C_C^p. \quad (2)$$

Under real-time dynamic traffic flow conditions, the relationship model between V_P and V_C is obtained in (3) according to $C_P = C_C$:

$$V_P = \frac{(C_0^t + C_P^n) \cdot (L - L_W) V_W V_C}{C_0^t L V_W + [(C_0^f L + C_C^p - C_P^e - C_0^t T_W) V_W - C_0^t L_W] V_C}. \quad (3)$$

Car speed is average speed and it is related to road traffic flow density and waiting time at intersection. The model between V_C and road traffic flow density K and waiting time T_D at intersection will be constructed. Assume that speed is a linear relationship with density, which is shown in (4). Limitation Assume that high speed of urban road is V_f , jam density is K_j , and the travel speed of car is V_R , then

$$V_R = V_f \left(1 - \frac{K}{K_j} \right), \quad (4)$$

$$V_C = \frac{L V_f (K_j - K)}{L K_j + T_D V_f (K_j - K)}. \quad (5)$$

Put (5) to (3), then

$$\begin{aligned} V_P &= \left((C_0^t + C_P^n) \cdot (L - L_W) V_W \cdot \frac{L V_f (K_j - K)}{L K_j + T_D V_f (K_j - K)} \right) \\ &\quad \times \left(C_0^t L V_W + [(C_0^f L + C_C^p - C_P^e - C_0^t T_W) V_W - C_0^t L_W] \right. \\ &\quad \left. \cdot \frac{L V_f (K_j - K)}{L K_j + T_D V_f (K_j - K)} \right)^{-1}. \end{aligned} \quad (6)$$

From the above model, it is shown that if parking cost C_C^p , walking distance L_W , walking speed V_W , and fuel cost C_0^f of per km and ticket cost are invariable, when the travel cost of choosing public transit and choosing car is equal for the same travel distance, public dynamic V_P is determined by car flow density, car waiting time at intersection, and energy consumption cost per unit time. Energy consumption cost per unit time is the function of crowding degree J in public transit. Therefore, V_P is the level of dynamic public transit priority.

3.2.2. The Public Transit Priority Demand Based on Minimum Carrying Speed. According to above priority level, road lane numbers of one direction, intersection form, entrance lane, and passing time for public transit priority can be determined. For public transit, the necessary driving speed on road and allowable waiting time at intersection can be

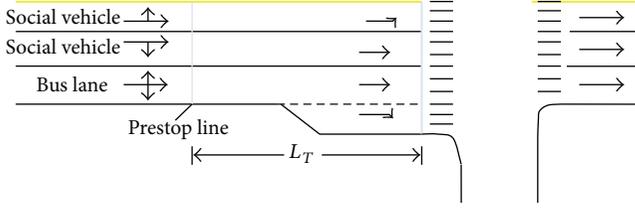


FIGURE 1: Traffic organization design sketch of dynamic public transit priority.

deduced from public route number N , departure interval τ , stop time S_T at station, and average spacing between intersection. Necessary public transit lane number can be determined by the necessary driving speed on road and public transit volume. The pass rule for public transit priority can be determined by allowable waiting time at intersection.

For public transit, the available passing time T_R on road and waiting time T_J at intersection are analyzed firstly:

$$T_R + T_J = \frac{L}{V_P} - \frac{L}{d} \cdot S_T. \quad (7)$$

Actually, the travel speed of public transit on public transit lane can reach V_R , thus,

$$T_J = \frac{L}{V_P} - \frac{L}{d} \cdot S_T - T_R. \quad (8)$$

It is obvious that available waiting time T_J at intersection is related to carrying speed V_P and stop time S_T . The longer T_R and S_T are, the shorter T_J is. Intersection signal timing is mainly determined by T_J .

Generally, one public transit lane is enough on road and two public transit lanes is needed at intersection under general condition. Stop line in advance is designed for public transit and car. Specific traffic organization and design is showed as Figure 1.

3.3. Dynamic Traffic Organization and Design of Dynamic Public Transit Priority. The dynamic traffic organization and design of dynamic public transit priority mainly include variable public transit lane design, pass space, and stop rule design at intersection and signal timing design. Variable public transit lane design is deduced from above model, with setting variable public transit lane sign and releasing information about variable public transit lane in advance in order to enable car traveler to select appropriate path or park and ride in advance. Two stop lines are adopted at intersection entrance lane. Vehicle firstly stops at stop line far away intersection for prewaiting. Vehicle which will pass intersection in next green signal waits before stop line near intersection. Specific organizational design is as shown in Figure 1.

The distance L_T between two stop lines is determined by public transit vehicle volume Q_B , which needs to pass intersection during one green time and entrance lane numbers n (relating to permit through rule influenced by signal timing).

The space way of public transit vehicle queue is l_B . The model used to determine L_T is as shown in (9):

$$L_T = \frac{Q_B l_B}{n}. \quad (9)$$

3.4. Model Analysis. The equation of $C_P = C_C$ is the basic requirement of dynamic public transit priority. The aim of dynamic public transit priority is to make traveler initiative to choose public transit, but the equation of $C_P = C_C$ cannot make sure that traveler will prefer to choose public transit, therefore, the equation of $C_P = C_C$ is adjusted as $C_P = K C_C$. Here, K is a coefficient. In order to determine K , questionnaire survey is carried out in Jiaozuo city. The question is that when the total travel cost including travel economic cost, travel time, and comfort is considered, R is the ratio of public transit travel cost and car travel cost, what value is R ? you will prefer to choose public transit. Two hundred questionnaires are carried out. Questionnaire survey shows that 60% of travelers choose 70 percentage and 22% of travelers choose 80 percentage. Therefore, it is rational to let $K = 0.7$, then $C_P = 0.7 C_C$. That is,

$$\begin{aligned} C_0^t \cdot \left(\frac{L_W}{V_W} + T_W + \frac{L - L_W}{V_P} \right) + C_P^e + C_P^n \cdot \frac{L - L_W}{V_P} \\ = 0.7 \left(C_0^t \cdot \frac{L}{V_C} + C_0^f L + C_C^p \right). \end{aligned} \quad (10)$$

Here, L is average travel distance, T_W is waiting time. For Jiaozuo city, average travel distance is 6 km, L_W is 0.6 km, T_W is 12 min, C_P^e is 1 yuan, C_0^f is 0.7 yuan, V_C is 22.9 km, C_C^p is 5 yuan, C_0^t is 10 yuan, and C_P^n is 8 yuan. Then according to (10), V_P can be calculated. $V_P = 25.116$ km/h. It is the level of dynamic public transit priority when the car carrying speed is 22.9 km.

But according to survey, actual V_P is 12.92 km/h. Now, V_P is difficult to be increased in Jiaozuo city; according to (10), C_C^p should be raised to 10 yuan.

4. Dynamic Stochastic Park and Ride Theory

4.1. The Connotation of Dynamic Stochastic Park and Ride. Current park and ride generally lies in urban peripheral. For urban center with serious traffic congestion, firstly, it is necessary to encourage traveler to reduce car use as far as possible, secondly, park and ride should be provided for traveler who have selected car when they feel too crowd. Also when dynamic public priority is carried out, some traveler of using car will want to park and ride because car carrying speed is lower than public transit carrying speed, then it is needed to provide parking facility for park and ride. This park and ride facility is located around general public transit station, with smaller scale comparing with general park and ride facility. Therefore, it is different from general park and ride. When dynamic public priority is carried out, not only public transit carrying speed is guaranteed, but also other choices are provided for car traveler on road. This intervention measures can be willingly accepted by traveler.

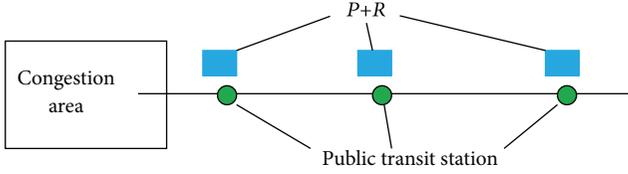


FIGURE 2: The layout sketch of dynamic stochastic park and ride.

When traveler has selected car and drives on road, at the beginning stage of travel, traveler is not sure that park and ride is necessary. If park and ride is needed, which site will be selected is not clear too. These are related to real-time dynamic traffic condition. That is, whether park and ride is needed and which park and ride site is selected is determined by dynamic traffic condition, which means that this park and ride has dynamic. Also it is obvious that this park and ride has uncertainty, which means it is stochastic. Therefore, this park and ride is called dynamic stochastic park and ride. At the same time, the above dynamic public transit priority will promote travel behavior of dynamic stochastic park and ride.

4.2. The Layout and Scale of Dynamic Stochastic Park and Ride. Dynamic stochastic park and ride layout is generally close to public transit station and corresponding road with more traffic volume will be priority selected. Figure 2 is layout sketch.

For dynamic stochastic park and ride, the closer it is to congestion area, the smaller the scale is. The diminishing size method mainly lies in two reasons. First, the closer it is to congestion area, the greater the difficulty of arranging park and ride land is. Second, traveler is encouraged and guided to park and ride in advance as far as it is possible in order to release traffic pressure of road and intersection close to congestion area.

The specific size of dynamic stochastic park and ride is determined by car volume of peak period and corresponding traffic volume with travel destination in congestion area. The corresponding traveler is investigated for park and ride proportion and site. Then, the park and ride volume for each station can be obtained and the park and ride scale for park and ride facility is determined. For car travel destination investigation, license plate photo can be adopted. Each road with park and ride facility should be investigated by license plate photo. The car volume with travel destination at congestion area is analyzed and for each car, whether park and ride will be adopted and which station will be selected are analyzed. If the station number with park and ride facility is N , corresponding road section number is N and corresponding intersection number is N . The layout sketch and charging standard for different site of dynamic stochastic park and ride are shown for car traveler during investigation (delivering card and feedback).

When investigation is completed, the car volume with passing road Section 1 and travel destination at congestion area are obtained, the park and ride volume at each park and ride facility of this volume, respectively, is PQ_{1j} ($j = 1, 2, \dots, N$). Car volume entering investigation road section

from i ($i = 2, \dots, N$) intersection combined with travel destination at congestion area is investigated and its ride volume at each park and ride facility is PQ_{ij} ($j = i, \dots, N$), then the park and ride volume of i park and ride facility is as shown in (11):

$$PR^i = PQ^i = \sum_{j=1}^i PQ_{ji} \quad (i = 1, \dots, N). \quad (11)$$

Therefore, the i park and ride facility scale can be preliminary determined as shown in (11).

After PR^i is determined, the car volume of each road section during peak period is Q^i ($i = 1, \dots, N$); when car traveler selects park and ride, the car volume of i road section reduces to as shown in (12):

$$Q_P^i = Q^i - \sum_{j=1}^i PQ_{ji} \quad (i = 1, \dots, N). \quad (12)$$

This model can be directly used to analyze V_p of dynamic public transit; that is the relationship between dynamic public transit priority and dynamic stochastic park and ride volume is established.

4.3. Charging Standard of Dynamic Stochastic Park and Ride.

The charging principle of dynamic stochastic park and ride is lower than charging standard of central area. For park and ride facility, the closer it is to congestion area, the higher the charging standard is. Charging standards of park and ride can be based on the charge standard of congestion area and it is reduced according to the distance of park and ride apart from congestion area. The charging standard of congestion area is F_J yuan per hour. The charging standard of the park and ride with large scale at urban peripheral is F_E yuan per hour and its distance apart from congestion area is L_{EJ} km; dynamic stochastic park and ride distance apart from congestion area is L_{PJ} km, then corresponding charging standard for park and ride is as shown in (13):

$$F_{PJ} = F_J - \frac{F_J - F_E}{L_{EJ}} \cdot L_{PJ} \quad \text{yuan per hour}. \quad (13)$$

4.4. The Guidance for Dynamic Stochastic Park and Ride.

The guidance for dynamic stochastic park and ride includes two aspects. On the one hand, it is needed to guide travelers to park and ride in advance as far as possible based on reducing car volume on road; on the other hand, the closer it is to congestion area, the smaller the scale of park and ride is, so, it is needed to guide travelers to park and ride in advance as far as possible based on park and ride capacity. When dynamic public transit priority is carried out, the closer it is to congestion area, the greater the car travel resistance is, and the parking cost is very high in area close to congestion area or in congestion area. The above two factors prompt traveler to park and ride in advance. At the same time, marked sign is set before each park and ride and propaganda slogan are adopted with showing the benefits and charge standard of park and ride. In order to provide convenience for park and

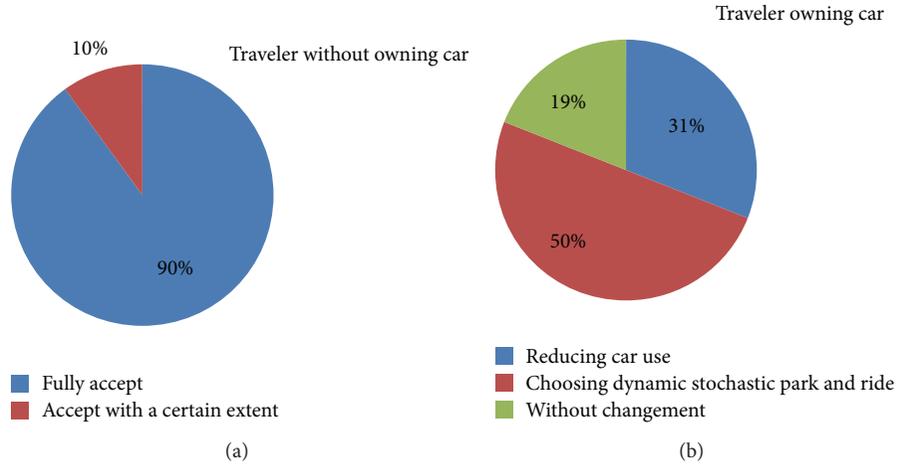


FIGURE 3: Traveler acceptability on dynamic public transit priority with dynamic stochastic park and ride.

ride traveler to transfer to public transit, public transit station is set at dynamic stochastic park and ride with bigger transfer demand.

5. Traveler Acceptability Analysis

The traveler acceptability on dynamic public transit priority with dynamic stochastic park and ride includes public transit traveler acceptability on dynamic public transit priority and car traveler acceptability on dynamic public transit priority with dynamic stochastic park and ride. Acceptability analysis is conducted by questionnaire investigation. The goal of dynamic public transit priority and charging standard of dynamic stochastic park and ride is stated in questionnaire. 300 questionnaires are handed out. One hundred respondents own cars. The result from investigating on traveler without owning car at present shows that ninety percent of them fully accept dynamic public transit priority and ten percent of them, who prepare to buy car recently, accept with a certain extent.

The result from investigating on traveler owning car shows that thirty-one percent of them intend to reduce car use by choosing public transit in peak period, fifty percent of them intend to choose dynamic stochastic park and ride, and nineteen percent of them do not intend to change car use pattern. The result is showed as in Figure 3.

It can be observed from above survey result that traveler has higher acceptability on dynamic public transit priority with dynamic stochastic park and ride. Therefore, dynamic public transit priority with dynamic stochastic park and ride has a good application prospect and it can relieve traffic congestion.

6. Further Discussion on Dynamic Public Transit Priority

The above discussion on dynamic public transit priority is mainly from the point of carrying speed, which is mainly from the point of dynamic space-time priority. But sometimes, the expected level of dynamic public transit priority

cannot be achieved because of many factors, such as road and traffic condition and car user acceptance. Then, other measures should be adopted in order to achieve the expected level of dynamic public transit priority. Other measures can also be obtained from the point that the travel cost of using public transit is not higher than travel cost of using car. If congestion charging C_C^c is considered, the travel cost of using car can be modified as follows:

$$C_C = C_C^t + C_C^f + C_C^p + C_C^c = C_0^t \cdot \frac{L}{V_C} + C_0^f \cdot L + C_C^p + C_C^c \quad (14)$$

According to $C_P = C_C$,

$$\begin{aligned} C_0^t \cdot \left(\frac{L_W}{V_W} + T_W + \frac{L - L_W}{V_P} \right) + C_P^e + C_P^n \cdot \frac{L - L_W}{V_P} \\ = C_0^t \cdot \frac{L}{V_C} + C_0^f \cdot L + C_C^p + C_C^c. \end{aligned} \quad (15)$$

In this equation, L_W , V_P , C_P^e , C_P^n , V_C , C_C^p , and C_C^c can be dynamically adjusted.

At present time, for most cities, the adjustable range of L_W , C_P^e , C_P^n is very small; that is, in order to achieve the equation of $C_P = C_C$, V_P , V_C , C_C^p , C_C^c should be key dynamic adjustment parameter.

Based on traveler acceptance, V_P and V_C should be firstly adjusted. When the equation of $C_P = C_C$ cannot be achieved, C_C^p should be adjusted. The last adjustment parameter is C_C^c . For traffic management, the key of dynamic public transit priority is how to achieve the balance among V_P , V_C , C_C^p , and C_C^c to make traveler to initiative accept it. This is the emphasis of future research.

7. Conclusion

Based on the point that the travel cost of choosing public transit is not higher than the travel cost of choosing car, the real-time dynamic relationship model among public transit

carrying speed, car carrying speed, and traffic density is established. The model of determining the level of dynamic public transit priority is constructed. For dynamic public transit priority, if parking cost and congestion charge do not reach a certain level, public transit carrying speed must be higher than car carrying speed. Dynamic stochastic park and ride is a matching measure for dynamic public transit priority and its layout and scale have feasibility, with promoting car traveler to dynamic stochastic transfer to public transit. Questionnaire survey shows that traveler has higher acceptability on dynamic public transit priority with dynamic stochastic park and ride and it illustrates that the application of dynamic public transit priority with dynamic stochastic park and ride is feasible. How to achieve the balance among public transit carrying speed, car carrying speed, parking cost, and congestion charging is the emphasis of future research.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is sponsored by Science and Technology Planning Project of Beijing Municipal Science and Technology Commission (no. Z1211000003120100) and the Natural Science Foundation of China (no. 61164003).

References

- [1] A. Tirachini, D. A. Hensher, and S. R. Jara-Díaz, "Restating modal investment priority with an improved model for public transport analysis," *Transportation Research E*, vol. 46, no. 6, pp. 1148–1168, 2010.
- [2] M. Eichler and C. F. Daganzo, "Bus lanes with intermittent priority: strategy formulae and an evaluation," *Transportation Research B*, vol. 40, no. 9, pp. 731–744, 2006.
- [3] L. A. Koehler and W. Kraus Jr., "Simultaneous control of traffic lights and bus departure for priority operation," *Transportation Research C*, vol. 18, no. 3, pp. 288–298, 2010.
- [4] X. Sun, H. P. Lu, and J. Wu, "Bus detection based on sparse representation for transit signal priority," *Neurocomputing*, vol. 118, pp. 1–9, 2013.
- [5] J. Wahlstedt, "Impacts of bus priority in coordinated traffic signals," *Procedia*, vol. 16, pp. 578–587, 2011.
- [6] M. J. Liu, J. Yang, L. L. Yang, and Z. Zhangb, "An elastic analysis on urban public transport priority in Beijing," *Procedia*, vol. 42, pp. 79–85, 2012.
- [7] V. Zyryanov and A. Mironchuk, "Simulation study of intermittent bus lane and bus signal priority strategy," *Procedia*, vol. 48, pp. 1464–1471, 2012.
- [8] P. Vedagiri and V. T. Arasan, "Estimating modal shift of car travelers to bus on introduction of bus priority system," *Journal of Transportation Systems Engineering and Information Technology*, vol. 9, no. 6, pp. 120–129, 2009.
- [9] W. H. Wang, H. W. Guo, H. Bubb, and K. Ikeuchi, "Numerical simulation and analysis procedure for model-based digital driving dependability in intelligent transport system," *KSCE Journal of Civil Engineering*, vol. 15, no. 5, pp. 891–898, 2011.
- [10] W. H. Wang, W. Zhang, H. W. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [11] M. Borjesson, J. Eliasson, M. B. Hugosson, and K. Brundell-Freij, "The Stockholm congestion charges—5 years on. Effects, acceptability and lessons learnt," *Transport Policy*, vol. 20, pp. 1–12, 2012.
- [12] J. Holguin-Veras, W. F. Yushimito, F. Aros-Vera et al., "User rationality and optimal park-and-ride location under potential demand maximization," *Transportation Research B*, vol. 46, pp. 949–970, 2012.
- [13] F. Aros-Vera, V. Marianov, and J. E. Mitchell, "p-hub approach for the optimal park-and-ride facility location problem," *European Journal of Operational Research*, vol. 226, no. 2, pp. 277–285, 2013.
- [14] B. Farhan and A. T. Murray, "Siting park-and-ride facilities using a multi-objective spatial optimization model," *Computers & Operations Research*, vol. 35, no. 2, pp. 445–456, 2008.
- [15] S. Meek, S. Ison, and M. Enoch, "UK local authority attitudes to Park and Ride," *Journal of Transport Geography*, vol. 18, no. 3, pp. 372–381, 2010.
- [16] S. Meek, S. Ison, and M. Enoch, "Evaluating alternative concepts of bus-based park and ride," *Transport Policy*, vol. 18, no. 2, pp. 456–467, 2011.
- [17] K. Kepaptsoglou, G. K. Matthew, and Z. Z. Li, "Optimizing pricing policies in Park-and-Ride facilities: a model and decision support system with application," *Journal of Transportation Systems Engineering and Information Technology*, vol. 10, no. 5, pp. 53–65, 2010.
- [18] H. M. Qin, H. Z. Guan, and Y. J. Wu, "Analysis of park-and-ride decision behavior based on Decision Field Theory," *Transportation Research F*, vol. 18, pp. 199–212, 2013.
- [19] N. Hounsell, B. Shrestha, and J. N. Piao, "Enhancing Park and Ride with access control: a case study of Southampton," *Transport Policy*, vol. 18, no. 1, pp. 194–203, 2011.

Research Article

Pedestrian Evacuation Time Model for Urban Metro Hubs Based on Multiple Video Sequences Data

Ji-biao Zhou,¹ Hong Chen,^{1,2} Jing Yang,¹ and Jiao Yan¹

¹ School of Highway, Chang'an University, Xi'an 710064, China

² 902 Transportation Science & Technology Building, Chang'an University, Middle Section of Nanerhuan Road, Xi'an 710064, China

Correspondence should be addressed to Hong Chen; hongchen82@126.com

Received 25 November 2013; Accepted 23 January 2014; Published 10 March 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Ji-biao Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evacuation time is a significant safety coefficient for Urban Metro Hubs (UMHs). Usually, a reasonable model for evacuation time will effectively promote the safety for pedestrian when emergency incidents occur in UMHs. In this paper, we propose a pedestrian evacuation time model for UMHs to improve the accuracy and reliability of its evacuation time. Firstly, we design an experiment survey based on the multiple video sequences to analyze the characteristics of pedestrian flow. Then, we decompose the evacuation process on the basis of the parameters, which involve the evacuation characteristics, the speed-density variation law, the pedestrian drop-off time, the platform evacuation time, and the channel evacuation time. Finally, we take the Bei Da-jie metro hub in Xi'an as an example, and verify the feasibility of the proposed pedestrian evacuation time model. The results show that the relative error for the evacuation time between the experiment result and the actual data is only 1.90%, where the experiment time is 169.87 s and the actual time is 166.64 s. Moreover, the proposed model strictly follows the Code for Design of Metro (GB 50157-003) and hence it can provide a good theoretical guidance for innovating the evacuation efficiency and the design reasonability of UMHs.

1. Introduction

Transport plays an important role in enhancing the quality of our living environment. The Urban Metro Hubs (UMHs), known as a key node of the integrated transport system, are the distribution center of passenger flow and play a crucial role in providing multimodal access to people and services in a manner that is convenient, safe, affordable, sustainable, efficient, and enjoyable. With the steady and rapid growth of passenger flow as well as the intensive transit, the closed environment space, and the complex transfer networks, the UMHs have become places vulnerable to extreme events, such as fire, stampede, and delay. Therefore, a time model for pedestrian evacuation is vital, for one thing, it can improve the efficiency of pedestrian evacuation, for another, it also ensure the passengers' life with a reasonable evacuation strategy.

In recent years, the pedestrian evacuation dynamics theory has become a hot topic in the academic field; they focus their attention on the passenger distribution characteristics and traffic behavior. It has made considerable progress since

the 1980s; the researchers observed the characteristics of pedestrian movement regularity. With the deepening of the research on the regularity in the pedestrian movement, some also put physiological and psychological characteristics of pedestrians as a factor into consideration and set up a large amount of pedestrian simulation models, which is considered to be a powerful tool for evaluating pedestrian flows in facilities. These models can be classified as macroscopic and microscopic models. These scholars who study macroscopic models believe that pedestrian movement behavior is similar to the flow of gas or liquid [1–3]. The early macroscopic models include Game theory [4], Decision theory and Propagation model, Queuing model [5], Transformation Matrix model [6], Stochastic model [7], and Route Choice Behaviors model [8]. However, because the above models lack the consideration of the self-organizing behavior of pedestrians, these models do not have accurate prediction results. Therefore, some scholars turn their attention to microscopic pedestrian simulation models [9]. References [10–14] study normal state population self-organization phenomenon and

proposed the social force model; many phenomena were simulated, which include the formation of the trail, oscillation at the bottleneck, and sheep-flock effect. Hereafter, [15–17] develop and improve the social force model to some extent. Reference [18] puts forward cellular automata model for the first time; it is fancied by most scholars because of its advantage in simulating the microscopic behavior of pedestrians. With the deepening of research some improved models are put forward, such as two process model [19], Prefixed Probabilities model [20], two floors model [21], and lattice-gas model [22, 23]. In addition to the social model and cellular automata model, [24] also puts forward magnetic force model, which has the characteristic of not using empirical relationships crowd density and speed, whereas it adopts Coulomb's theorem describing pedestrians as the object in magnetic field. With the development of computer technology, the computer simulation of pedestrian evacuation process becomes possible. It can simulate pedestrians directly in the moving process of the building and record the different time of pedestrians position change and meanwhile calculate the pedestrian evacuation time. There are about 22 pedestrian evacuation models, among them EVACNET [25, 26], Building EXODUS [27, 28], FIRECAM [29], SIMULEX [30, 31], and HAZARD [32, 33] have received more attention because of their relatively authentic reflection in pedestrians' choice of escape routes and pedestrians' choice evacuation.

In China, the study of pedestrian and evacuation dynamics theory is in early stages, which is mainly focused on pedestrian traffic characteristics in urban traffic environment and the crowd evacuation in large public places or buildings (e.g., a theatre, a stadium, or a shopping mall). In the first part, [34] proposes the centrifugal force model of pedestrian dynamics and simulates for pedestrian movement process; [35] proposes a cellular automata-based alighting and boarding microsimulation model for passengers in Beijing metro stations; [36, 37] propose an evacuation time model for passengers in metro platforms from field data by considering crowd density and physical characteristic of stations; [38–40] investigate pedestrian traffic characteristics of Chinese metropolis and get the basic data of pedestrian traffic characteristics and level of service of crossing facilities by establishing model. In the second part, [41] identifies seven methodological approaches (e.g., cellular automata models, lattice gas models, social force models, fluid-dynamic models, agent-based models, game theoretic models, and approaches based on experiments with animals) for crowd evacuation which has been studied over the last decades; [42] presents a new emergency evacuation system based on Multi-Agent framework and Geographic Information System (GIS), which can simulate human's typical behaviors in multiexit public place and present the public evacuation process and in-time distribution during emergency; [43–45] investigate the route choice of pedestrians during evacuation under conditions of both good and zero visibility in a classroom and build a microscopic pedestrians model with discrete space representation.

Through analyzing the current research on pedestrian, this paper maintains that the study of pedestrians and evacuation dynamics is more focused on a building (e.g., a theatre,

a stadium, or a shopping mall), rather than the UMH. What is more, these models do not consider the queuing delay caused by limited capacity in stairs or escalator, gate and China's pedestrian characteristics are ignored. On the basis of summarizing traditional evacuation time models, this paper proposes to do research, respectively, on the character of the UMHs structure and Chinese pedestrian traffic characteristics and meanwhile use queuing theory and fluid mechanics simulation theory to establish a pedestrian evacuation time model in UMHs, which includes the pedestrian drop-off time, pedestrian evacuation time in platform, and pedestrian evacuation time in the channel.

2. Experiment Survey Based on Multiple Video Sequences

2.1. Experiment Design

(1) *The Selection of Experimental Observation Time and Place.* The morning peak and evening peak are different to some extent. Through the survey in the UMHs, the pedestrian traffic volume in the evening peak is higher than the morning peak; the trip by the metro is a commuter trip more often; by contrast, the peak hour mainly focuses on the on and off duty and business travel at the normal working day. The survey time should choose the time in the tremendous traffic demand. Hence, the normal working days are chosen. In addition, through analysis of the pedestrian flow characteristics in the UMHs, business people travel mainly by elastic travel, office people and living people mainly by rigid travel. At the same time, the analysis period should be selected in the maximum hour to analyze in the UMHs. Eventually, we determine the survey period is evening peak. The time of experiment survey based on multiple video sequences is chosen in 17:00–19:00 on the working day and sums up to 2 hours in the evening peak. Moreover, for the purpose of eliminating the influence of human factors, the measurement of geometric parameters in the UMHs is selected at the time of fewer flow. Through the live measurements on the UMHs and analysis of the pedestrian flow characteristic, coupled with the fact where there is only a transfer hub site in Xi'an, that is, Bei Da-jie station, at the end, we select the Bei Da-jie station as the best observation in the experiment survey, as shown in Figure 1.

(2) *The Specific Design Scheme.* According to the Bei Da-jie site's CAD drawings of facility layout and the investigation, the position of video cameras is determined, as shown in Figure 2 (the circle is the video camera location); the multiple video sequences can be obtained in the video cameras. In the experiment survey, there are 21 video cameras in the Bei Da-jie station, which record the evolution process of multiple infrastructure in the evening peak.

A specific operation is as follows.

- (1) At the exit: we select the video cameras, which represent the purple circle1, the purple circle2, the purple circle3, the purple circle4, and the purple circle5, to observe the pedestrian flow's parameters.

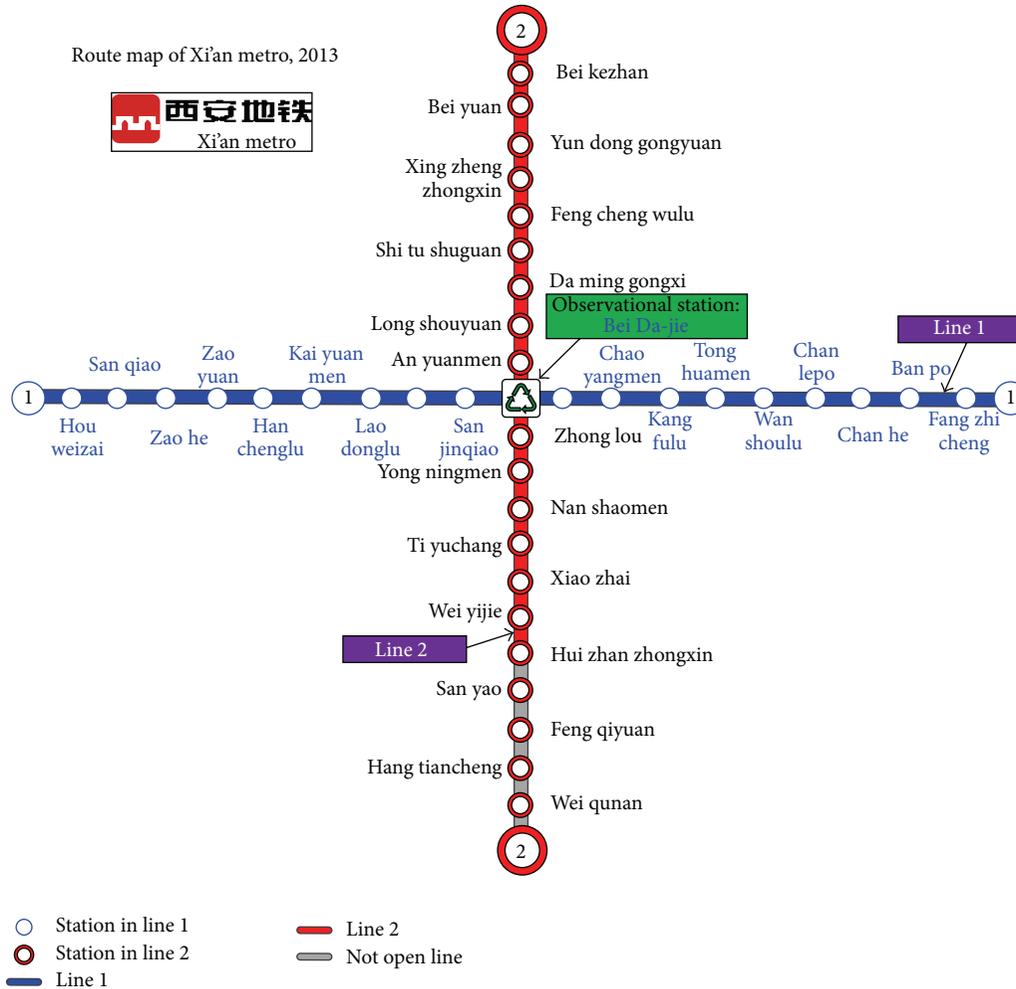


FIGURE 1: Observational station (Bei Da-jie) for pedestrian flow.

- (2) On the stairs and escalators: we select the video cameras, which represent the blue circle1, the blue circle2, the blue circle3, the blue circle4, the blue circle5, and the blue circle6, to observe the pedestrian flow's parameters. Where the blue circle1, the blue circle2, and the blue circle3 are used to observe the parameters of down direction of the stairs and escalators, meanwhile, the blue circle4, the blue circle5, and the blue circle6 are used to observe the parameters of up direction of the stairs and escalators.
- (3) In the TVM: we select the video cameras, which represent the green circle1, the green circle2, the green circle3, and the green circle4 to observe the pedestrian flow's parameters.
- (4) In the auto gate: we select the video cameras, which represent the red circle1, the red circle2, the red circle3, and the red circle4 to observe the pedestrian flow's parameters.

- (5) In the transfer channel: we select 1 video camera independently to observe the pedestrian flow's parameters in the transfer channel.
- (6) On the platform: we select 1 video camera independently to observe the pedestrian flow's parameters in the platform.

2.2. Data Collecting and Data Processing

(1) *Data Collecting.* In order to obtain the pedestrian traffic characteristics exactly in the Bei Da-jie hub, we use the video camera method to obtain the multiple video sequences data, to record video information to collect the pedestrian flow parameter data, which is shown in Figure 3.

Video acquisition method, which is one of the widely methods in the investigation. The advantage of the method is convenient for storage and revisit detail information. At the same time, the advantage of the method is almost collects all of the multiple video sequence data in the pedestrian flow characteristics. The observation pedestrian flow is used to the video acquisition method such as speed, and density in

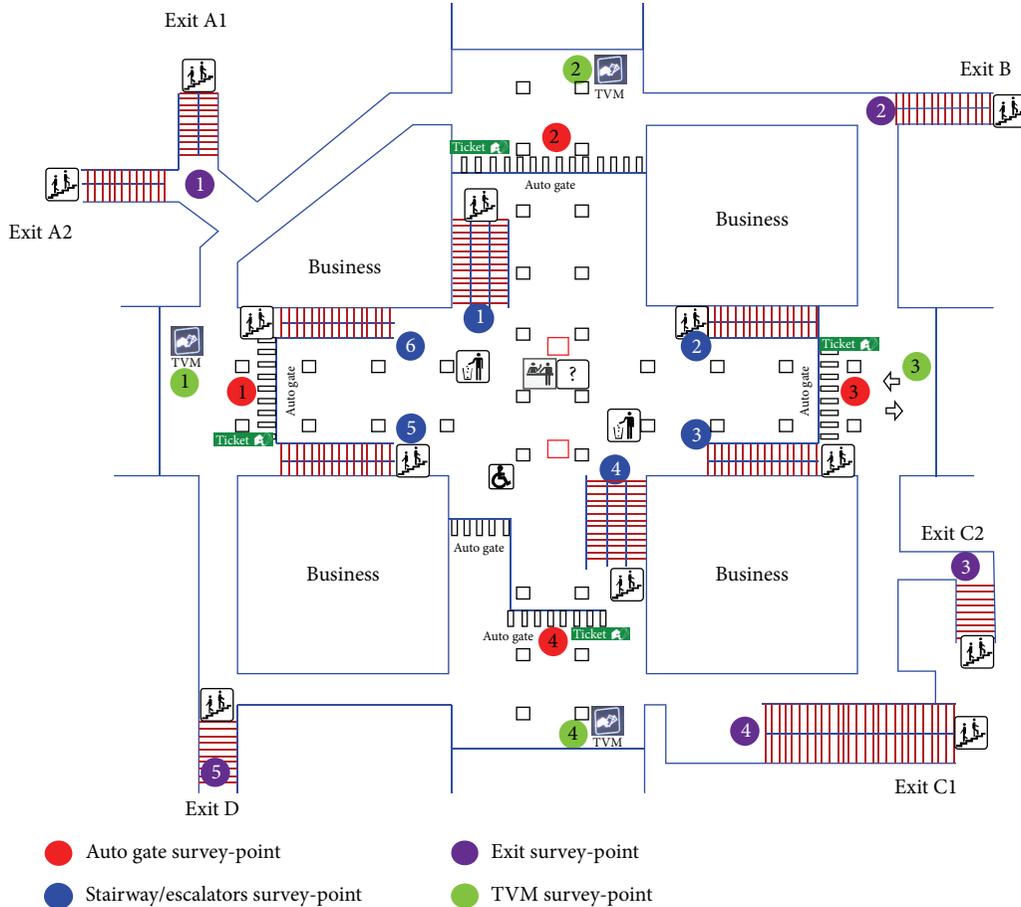


FIGURE 2: The video camera layout drawing in Bei Da-jie hub.



FIGURE 3: Data collection by the video camera.

the platforms, transfer channels, auto gates, stairs and escalators, exits, and so on.

(2) *Data Processing.* According to the multiple video sequence data from the camera, every video is divided into frames in a second and play video frame by frame in processing, then, we record the correct time in the observation area between the pedestrian getting into the observation area and leaving observation area.

The pedestrian speed is calculated by the length and time of the observation area in the camera. In the observation, not only record the speed but also the pedestrian density and the corresponding pedestrian flow of the moment, and then put all the multiple video data into the tables, for the purpose of subsequent data analysis. In the process of data analysis, we use the SPSS to fit the curve relationship of pedestrian flow parameters, and we also test the fitting degree using the value of R -squared.

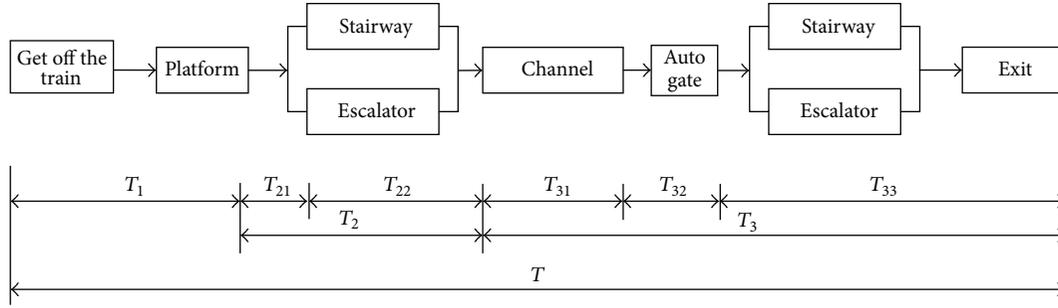


FIGURE 4: Decomposition process of safety evacuation in UMHS.

3. Pedestrian Evacuation Time Model for UMHS

Through the observation and analysis in Xi'an Bei Da-jie hub, the walking speed and spaces characteristic are different from each other; based on these differences between them, we decompose the evacuation process, and the pedestrian evacuation time (T) is divided into three parts; the first part is pedestrian drop-off time (T_1), the second part is safety evacuation time in platform (T_2), and the third part is time of pedestrian pass channel, stairway (or escalators) (T_3). Safety evacuation time is shown in Figure 4, and detailed description for every part of evacuation time is shown as follows.

3.1. Pedestrian Drop-Off Time Model. Pedestrian drop-off time refers to the train arriving at the station stably and the time all passengers dropping off from the train to the platform. Drop-off time is related to drop-off numbers, get-up numbers, and the width of the train gate door. Here we suppose the pedestrian drop-off time is T_1 (s).

Through the analysis, we can conclude that the relationship between the single door drop-off time and the number of passengers is different; it can be power exponent relationship, linear relationship, exponential relationship, and natural logarithm. Among them, the power exponent relationship's correlation coefficient (R^2) is the biggest, with the value 0.9924, as shown in Figure 5.

In Figure 5, the number of passengers and drop-off time satisfy the following formula:

$$T_1 = \alpha_1 x_{\max}^{\beta_1} = 0.6311x^{0.9884}, \quad (1)$$

where x_{\max} : the largest passengers of drop-off passenger of gate (person), α_1, β_1 : Parameters for calibration.

3.2. Pedestrian Evacuation Time in Platform. Safety evacuation time in platform is described in metro design code in China; namely, the width of the exit stairs and evacuation corridor should guarantee the passengers both on the train and the platform and staffs working in the station to escape the fire disaster in 6 minutes. Specific safety evacuation time in platform in metro specification can be seen as follows:

$$T_0 = 1 + \frac{(Q_1 + Q_2)}{\{0.9 \cdot [A_1 \cdot (N - 1) + A_2 \cdot B]\}}. \quad (2)$$

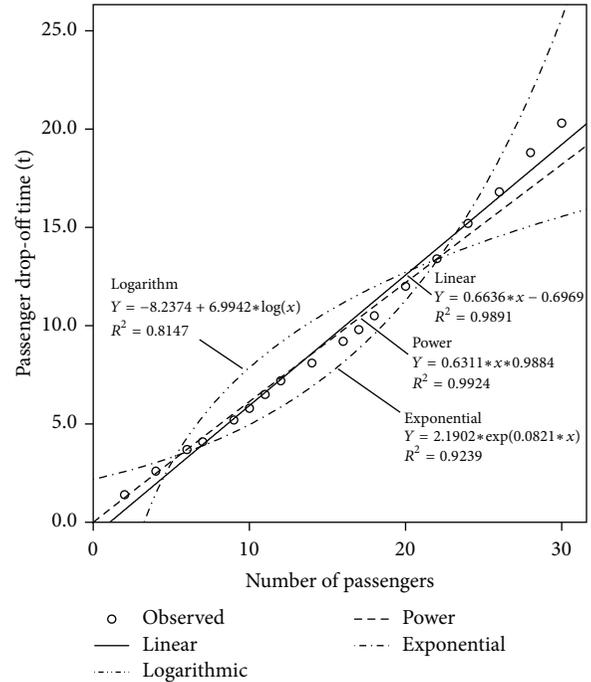


FIGURE 5: The function curve fitting between passenger number and drop-off time.

The evacuation time in platform in metro specification above only considers the biggest evacuation capacity of different facilities and thinks little of the pedestrian density and the impact of environment on pedestrian speed, and the influence of guideline information (directional signs, direction signs, radio evacuation command, radio evacuation message, etc.) is also insightful and thinks little about the factors that may affect traffic organization in the transport hub, making the large deviation time between the theoretical calculation result and the actual evacuation. Hence, in the process of channel (stairway, escalators) evacuation, we consider the speed-density change law of the traffic flow, taking wave theory as an example to simulate the pedestrian flow, as well as considering the accumulation and dissipation characteristic of pedestrian, compared to only considering the evacuation capability on the platform which has improved, further to improve the traditional evacuation time model.

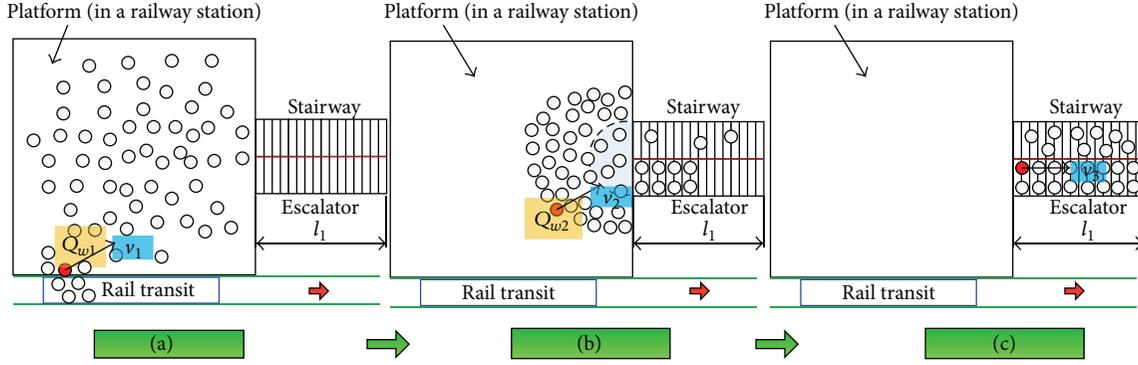


FIGURE 6: Pedestrian evacuation process.

Platform evacuation time can be divided into three phases, as shown in Figure 6. Here we suppose the final pedestrian as the research object, as shown in the red circle in the figure. The final pedestrian (red circle) of the evacuation process can be divided into two stages. The first stage is shown in Figures 6(a) and 6(b), at this period of time, the speed of him is mainly influenced by the evanescent wave which transmits from high density to low density, and the evanescent wave is beginning to transform into the gathering wave. During the forward process, there will be a speed mutation point; the pedestrian (red circle) will suddenly slow down at this point, which is due to the pedestrian waiting in line in front of the evacuation platform, so the flow density is very large and the corresponding speed is very low; the second stage is shown in Figure 6(b) to Figure 6(c), pedestrian (red circle) is mainly affected by the fluctuation of the pedestrian flow, and the gathering wave is beginning to transform into evanescent wave.

Here we suppose the safety evacuation time in platform is $T_2(s)$, pedestrian passing time to the stairway (or escalators) is $T_{21}(s)$, and pedestrian passing time in stairway (or escalators) is $T_{22}(s)$. Thus the safety evacuation time mode in the platform is shown below

$$\begin{aligned} Q_{w1} &= \frac{(v_1 - v_2)}{(1/k_1 - 1/k_2)}, \\ Q_{w2} &= \frac{(v_3 - v_2)}{(1/k_3 - 1/k_2)}, \\ T_{21} &= T_1 \cdot \frac{Q_{w1}}{(Q_{w1} - Q_{w2})}, \end{aligned} \quad (3)$$

where Q_{w1} —wave flow of pedestrians aggregation wave after getting off (p/(m·s)), Q_{w2} —wave flow of pedestrians evanescent wave on stairway (or escalator) (p/(m·s)), v_1 —pedestrian normal walking speed after getting off the train (m/s), v_2 —pedestrian dissipation speed on stairway (or escalators) (m/s), v_3 —escalator running speed or pedestrians speed climbing up the stairway (m/s), k_1 —pedestrian density under v_1 state (p/m²), k_2 —pedestrian density under v_2 state (p/m²), and k_3 —pedestrian density under v_3 state (p/m²):

$$T_{22} = \frac{l_1}{v_3}, \quad (4)$$

where l_1 —stairway (or escalator) length (m).

Thus the evacuation time in platform T_2 is

$$T_2 = T_{21} + T_{22}. \quad (5)$$

3.3. Pedestrian Evacuation Time in Channel (Stairway, Escalators). We suppose the time of pedestrian pass channel, stairway (or escalators), is $T_3(s)$, the time of pedestrian pass channel is $T_{31}(s)$, pedestrian dissipation time on auto gate is $T_{32}(s)$, and pedestrians pass time on stairway (or escalators) is $T_{33}(s)$.

(1) Calculation of the Time of Pedestrian Pass Channel. Pedestrian speed in the corridor or stairway is largely influenced by flow density. The pedestrians occupied smaller place when the passenger flow volume is larger, and the pedestrian speed is slower. While on the other hand, the smaller the flow density, the faster the pedestrian speed. According to the collected data in the Bei Da-jie station, we find the relationship between speed and density in pass channel, which is shown in Figure 7.

According to the fitting functions between the speed and density of pedestrians, we can see that there is a strong correlation between the speed and density of pedestrians in a subway corridor or in the stairway (upward direction and downward direction of the stairs), and they comply with third-degree polynomial curve (R^2 is the biggest, with the value 0.934), which is shown in Table 1.

Thus the time of pedestrian pass channel can be shown as follows:

$$T_{31} = \frac{l_2}{v(k)} = \frac{l_2}{(\alpha_2 k^3 + \beta_2 k^2 + \gamma k + C)}, \quad (6)$$

where l_2 —channel length (m), k —pedestrian density in corridor (p/m²), and $\alpha_2, \beta_2, \gamma, C$ —parameters for calibration.

(2) Calculation of Pedestrian Dissipation Time on Auto Gate. Here we suppose the mean arrival rate of pedestrian flow is λ and the average service rate of auto gate is μ ; thus the traffic intensity of auto gate is ρ . And we assume that the time of the pedestrian pass through the auto gate meets the standard M/M/1 model (Figure 8).

In the standard M/M/C and M/M/1/C model, the regulation of characteristics of those models is the same with

TABLE 1: Model summary and parameter estimates.

Equation	Model summary					Parameter estimates			
	R Square	F	df1	df2	Sig.	Constant	b1	b2	b3
Logarithmic	0.795	236.71	1	61	0	1.228	-0.355		
Cubic	0.934	276.576	3	59	0	1.651	-0.229	-0.113	0.022
Exponential	0.894	516.655	1	61	0	1.836	-0.349		

The independent variable: density.
Dependent variable: speed.

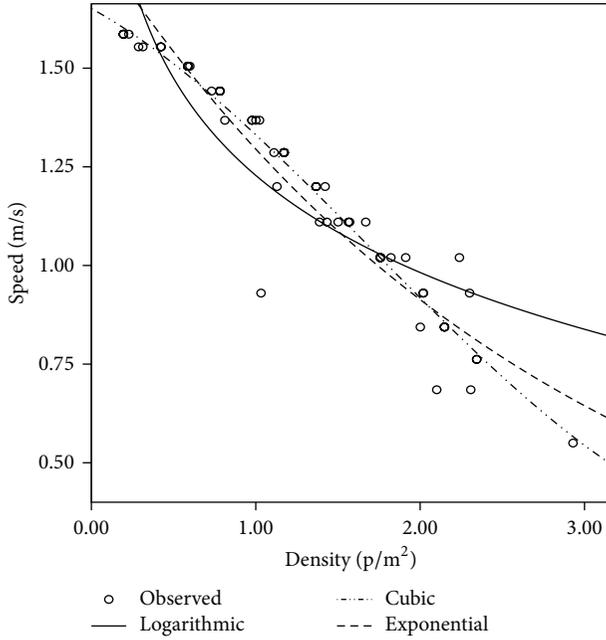


FIGURE 7: Create fit curve between speed and density.

the standard M/M/1. In addition, specified the auto gates are mutual independent and the average service rates are the same; namely, $\mu_1 = \mu_2 = \dots = \mu_c = \mu$. So the average service rate for the whole services agency is $c\mu$ ($n \geq c$) ($n\mu$ ($n < c$)).

Similarly, we find

(a) the entire gates idle probability

$$P_0 = \left[\sum_{k=1}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k + \frac{1}{c!} \cdot \frac{1}{1-\rho} \cdot \left(\frac{\lambda}{\mu} \right)^c \right]^{-1}, \quad (7)$$

(b) average number of customers in system L_s

$$L_s = L_q + \frac{\lambda}{\mu} \quad (8)$$

(c) average numbers of customer in the queue L_q

$$L_q = \sum_{n=c+1}^{\infty} (n-c) P_n = \frac{(c\rho)^c \rho}{c!(1-\rho)^2} P_0 \quad (9)$$

(d) expected value of stay time in the system

$$W_s = \frac{L_s}{\lambda}, \quad (10)$$

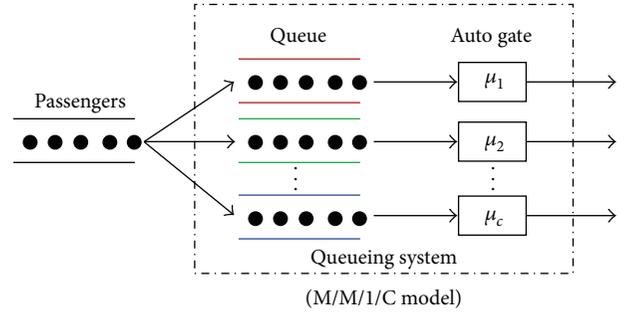


FIGURE 8: Passengers queuing at the auto gate.

Thus, pedestrian dissipation time on auto gate is T_{32}

$$T_{32} = W_s. \quad (11)$$

(3) Calculation of Pedestrians Pass Time on Stairs (or Escalators). For escalator, the speed and width of it are fixed, and the pedestrians are upward and downward with it automatically without walking on it, so the time is basically fixed. As shown in the following:

$$T_{33} = \frac{l_1}{v_3}, \quad (12)$$

where l_1 —stairway (or escalator) length (m).

The whole pedestrian evacuation time in the Urban Metro Hub is as follows:

$$T_3 = T_{31} + T_{32} + T_{33}. \quad (13)$$

3.4. Pedestrian Evacuation Time Model. The safety evacuation time of pedestrian is divided into three parts, and with those three parts added up we can get the whole evacuation time. Taking the pedestrian drop-off time, safety evacuation time in platform, time of pedestrian pass channel, and stairway (or escalators) into consideration, the safety evacuation time model in UMHS is established and is shown in what follows:

$$T = \sum_{i=1}^3 T_i. \quad (14)$$

3.5. Example Verification

3.5.1. General Situation. In this paper, we take the Bei Dajie metro hub in Xi'an as an example, mainly based on two

TABLE 2: Observed values of model parameters in various stages.

Number	Model name	Index name	Observed values	Unit
1	Pedestrian drop-off time	T_1	20	—
		v_1	1.495	(m/s)
		v_2	1.05	(m/s)
		v_3	0.93	(m/s)
2	Pedestrian evacuation time in platform	k_1	1.167	(p/m ²)
		k_2	1.193	(p/m ²)
		k_3	2.181	(p/m ²)
		l_1	28.24	m
		k	1.314	(p/m ²)
		l_2	100	m
3	Pedestrian evacuation time in channel	c	6	—
		μ	0.5	(p/min)
		λ	1.35	(p/min)

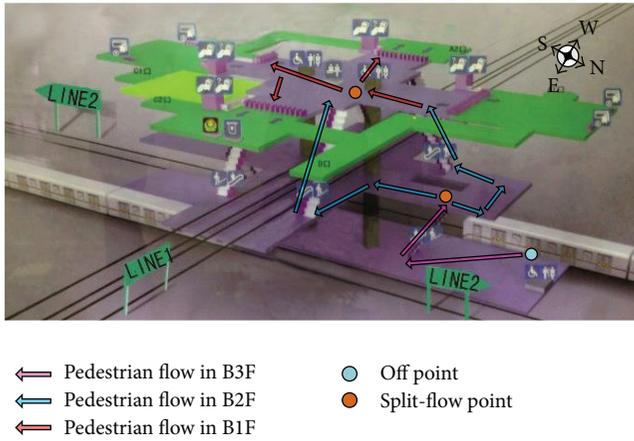


FIGURE 9: Pedestrians' walking flow lines.

reasons. On the one hand, the station is a typical “ten” shaped metro interchange station, line 1 and line 2 meet here, line 1 site is side platform, line 2 site is an island platform, and the pedestrians' walking flow lines are shown in Figure 9. On the other hand, the station is the only staggered floor transfer station based on a ladder in Xi'an, which is characterized by the shortest distance traveled by passengers; especially in peak hours the transfer time is not more than one minute. Therefore, a higher request for passenger transport organization and train time interval coordinate should be put forward.

3.5.2. *The Relevant Data Measured.* Before studying the pedestrians' transfer time model in Xi'an Bei Da-jie station, some parameters are measured, which are shown in Table 2.

3.5.3. Results

(1) *The Observed Results.* Transferring the case to the Bei Da-jie when the peak hours of pedestrian flow is short, and a cycle of pedestrian flow evacuation time is measured; each stage

TABLE 3: The observed pedestrian evacuation time in each stage.

Stage name	Observed time (s)
Pedestrian drop-off time	11.16
Pedestrian evacuation time in platform	44.86
Pedestrian evacuation time in channel	113.85
Total time	169.87

TABLE 4: The calculated pedestrian evacuation time in each stage.

Stage name	Calculated time (s)
Pedestrian drop-off time	12.23
Pedestrian evacuation time in platform	42.76
Pedestrian evacuation time in channel	111.64
Total time	166.64

of the pedestrian evacuation time is shown in Table 3. The total number of pedestrians is 317 that transfer to line 1 from line 2, the last pedestrian used 172.82 seconds totally when he arrived at line 1 platform.

(2) *The Calculated Results.* The established evacuation time model in this paper is used to calculate each stage of the pedestrian evacuation time; the result is shown in Table 4.

(3) *Comparison Analysis.* Comparing the observed time with the calculated time, we can see that the error between them is only 1.90%. What is more, the error in each stage is also very tiny; the result is shown in Table 5. In the end, the estimation results of the established model in the paper and measured values are consistent; the applicability of the model is reasonable.

4. Conclusions

(1) Through calculating the whole evacuation time in Xi'an Bei Da-jie transfer hub, the observed results in the model are 169.87 s, the calculated results in the model are 166.64 s as well, and the relative error between them is 1.90%, where

TABLE 5: The relative error analysis.

Stage name	Observed time (s)	Calculated time (s)	Relative error
Pedestrian drop-off time	11.16	12.23	-9.59%
Pedestrian evacuation time in platform	44.86	42.76	4.68%
Pedestrian evacuation time in channel	113.85	111.64	1.94%
Total time	169.87	166.64	1.90%

we have the following. (a) The pedestrian drop-off time's measured value is 11.16 s, the value of calculation is 12.23 s, and the relative error is -9.59%. (b) The value of measure in platform evacuation time is 44.86 s, the value of calculation is 42.76 s, and the relative error is 4.68%. (c) The value of measure in channel evacuation time is 113.85 s, the value of calculation is 111.64 s, and the relative error is 1.94%. Compared to the Code for Design of Metro (GB 50157-003) promulgated by Ministry of Housing and Urban-Rural Construction of the People's Republic of China (MOHURD), the whole time is 166.64 s, less than the 6 min in Code for Design of Metro (GB 50157-003), and the applicability of the pedestrian evacuation time model was verified.

(2) In terms of the pedestrian drop-off time model, through fitting analysis with the multiple video sequence data, the four functions of the relationship are fitted between passengers getting off them through a single door and a number of alighting passengers; the fitting functions are composed of the exponent relationship, linear relationship, exponential relationship, and the natural logarithm relationship. The correlation coefficient of exponent relationship is the largest and the correlation coefficient (R^2) is 0.9924. The relationship between passengers getting off time and the number of alighting passengers is an exponent relationship. For the platform evacuation time model, the pedestrian evacuation process obeys the wave theory. As an improvement, compared to Code for Design of Metro (GB 50157-003), we consider the pedestrian flow speed-density changing laws and combining with the flow wave theory, the platform evacuation time model has been improved for better reflection of the pedestrian flow distribution characteristics in this paper.

(3) In terms of the channel evacuation time models, the pedestrian evacuation time model has been split into the time that pedestrians pass through the corridor, the auto gate, and stairs or escalators. The research result indicates that there is a strong quartic polynomial relation between pedestrian flow velocity and density of the subway corridor and upward direction and downward direction of the stairs. The time of pedestrians through the auto gate is to meet the standard M/M/C model. The actual evacuation time in the hub corridor can be better reflected while comprehensively considering the time, which consider pedestrians passing through the corridor, the auto gate, and stairs or escalators into the channel evacuation time model.

(4) In addition, as a future work, we only take the Bei Da-jie transfer hub in Xi'an as an example and verify the feasibility of the proposed pedestrian evacuation time model. Next, we should do lots of experiment surveys, such as other Urban Metro Hubs in Xi'an, verify and modified the

model. In addition, could the Urban Metro Hubs (UMHs) become a full-time laboratory? What if you could analyze every transaction, capture insights from every passenger interaction, and did not have to wait for months to get data from the field? What if...? Big data are flooding in at rates never seen before—doubling every 18 months. In the future, how to set up the platform for big data, big discovery, and big decision in the UMHs? The answer is technology. Technology for capturing and analyzing big data is widely available at ever-lower price points. With the recent developments of sensing, networking, and computing technologies, more and more UMHs-related big data and computational resources become available.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgements

This paper has been supported by the National Natural Science Foundation of China (Grant no. 51208051) and the Fundamental Research Funds for the Central Universities (CHD2011ZD014). The authors deeply appreciate the supports. The authors wish to acknowledge the contribution of all the expert groups and other team members, including Associate Professor Yong-gang WANG, Dr. Long-fei WANG, Dr. Xiao-wei LI, and M.E. Yu-ting HE. The authors also wish to express their gratitude to the anonymous reviewers for their insightful comments and suggestions to an earlier version of the paper. Special gratitude is extended to those construction academics and industry practitioners who have responded to and contributed their valuable input to complete the survey questionnaires.

References

- [1] L. F. Henderson, "The statistics of crowd fluids," *Nature*, vol. 229, no. 5284, pp. 381–383, 1971.
- [2] L. F. Henderson, "On the fluid mechanics of human crowd motion," *Transportation Research*, vol. 8, no. 6, pp. 509–515, 1974.
- [3] L. F. Henderson and D. J. Lyons, "Sexual differences in human crowd motion," *Nature*, vol. 240, no. 5380, pp. 353–355, 1972.
- [4] D. Helbing, "A mathematical model for the behavior of individuals in a social field," *Journal of Mathematical Sociology*, vol. 19, pp. 189–219, 1994.

- [5] S. J. Yuhaski Jr. and J. M. Smith, "Modeling circulation systems in buildings using state dependent queueing models," *Queueing Systems*, vol. 4, no. 4, pp. 319–338, 1989.
- [6] D. Garbrecht, *Describing Pedestrian and Car Trips by Transition Matrices*, vol. 27 of *Traffic Quarterly*, 1973.
- [7] A. J. Mayne, "Some further results in the theory of pedestrians and road traffic," *Biometrika*, vol. 41, pp. 375–389, 1954.
- [8] A. Borgers and H. Timmermans, "A model of pedestrian route choice and demand for retail facilities within inner-city shopping areas," *Geographical Analysis*, vol. 18, no. 2, pp. 115–128, 1986.
- [9] M. Asano, T. Iryo, and M. Kuwahara, "Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour," *Transportation Research C*, vol. 18, no. 6, pp. 842–855, 2010.
- [10] D. Helbing, "A fluid dynamic model for the movement of pedestrians," *Complex Systems*, vol. 6, pp. 391–415, 1998.
- [11] D. Helbing, L. Buzna, A. Johansson, and T. Werner, "Self-organized pedestrian crowd dynamics: experiments, simulations, and design solutions," *Transportation Science*, vol. 39, no. 1, pp. 1–24, 2005.
- [12] D. Helbing, I. J. Farkas, P. Molnar, and T. Vicsek, "Simulation of pedestrian crowds in normal and evacuation situations," *Pedestrian and Evacuation Dynamics*, vol. 21, pp. 21–58, 2002.
- [13] D. Helbing, P. Molnár, I. J. Farkas, and K. Bolay, "Self-organizing pedestrian movement," *Environment and Planning B*, vol. 28, no. 3, pp. 361–383, 2001.
- [14] D. Helbing and T. Vicsek, "Optimal self-organization," *New Journal of Physics*, vol. 1, pp. 13.1–13.17, 1999.
- [15] D. R. Parisi and C. O. Dorso, "Microscopic dynamics of pedestrian evacuation," *Physica A*, vol. 354, no. 1–4, pp. 606–618, 2005.
- [16] K. Teknomo, "Application of microscopic pedestrian simulation model," *Transportation Research F*, vol. 9, no. 1, pp. 15–27, 2006.
- [17] A. Seyfried, B. Steffen, and T. Lippert, "Basics of modelling the pedestrian flow," *Physica A*, vol. 368, no. 1, pp. 232–238, 2006.
- [18] T. Toffoli and N. H. Margolus, "Invertible cellular automata: a review," *Physica D*, vol. 45, no. 1–3, pp. 229–253, 1990.
- [19] V. J. Blue and J. L. Adler, "Cellular automata microsimulation for modeling bi-directional pedestrian walkways," *Transportation Research B*, vol. 35, no. 3, pp. 293–312, 2001.
- [20] F. Weifeng, Y. Lizhong, and F. Weicheng, "Simulation of bi-direction pedestrian movement using a cellular automata model," *Physica A*, vol. 321, no. 3–4, pp. 633–640, 2003.
- [21] C. M. Henein and T. White, "Macroscopic effects of microscopic forces between agents in crowd models," *Physica A*, vol. 373, pp. 694–712, 2007.
- [22] X. Guo, J. Chen, S. You, and J. Wei, "Modeling of pedestrian evacuation under fire emergency based on an extended heterogeneous lattice gas model," *Physica A*, vol. 392, pp. 1994–2006, 2013.
- [23] X. Guo, J. Chen, Y. Zheng, and J. Wei, "A heterogeneous lattice gas model for simulating pedestrian evacuation," *Physica A*, vol. 391, no. 3, pp. 582–592, 2012.
- [24] S. Okazaki and S. Matsushita, "A study of simulation model for pedestrian movement with evacuation and queuing," in *Proceedings of the International Conference on Engineering for Crowd Safety*, pp. 271–280, 1993.
- [25] T. M. Kisko, R. Francis, and C. Nobel, *Evacnet4 User's Guide*, University of Florida, 1998.
- [26] A. T. Syed and M. A. Hassanain, "A simulation model for emergency evacuation time of a library facility using EVACNET4," *Structural Survey*, vol. 31, pp. 1–10, 2013.
- [27] E. R. Galea and J. M. Perez Galparsoro, "A computer-based simulation model for the prediction of evacuation from mass-transport vehicles," *Fire Safety Journal*, vol. 22, no. 4, pp. 341–366, 1994.
- [28] L. Yang, P. Rao, K. Zhu, S. Liu, and X. Zhan, "Observation study of pedestrian flow on staircases with different dimensions under normal and emergency conditions," *Safety Science*, vol. 50, no. 5, pp. 1173–1179, 2012.
- [29] D. Yung, G. Hadjisophocleous, and G. Proulx, "A description of the probabilistic and deterministic modelling used in FiRE-CAM," *International Journal on Engineering Performance-Based Fire Codes*, vol. 1, pp. 18–26, 1999.
- [30] R. Fahy, "A practical example of an evacuation model for complex spaces," in *Proceedings of the Human Behaviour in Fire-1st International Symposium*, pp. 743–751, 1999.
- [31] C. Rogsch and W. Klingsch, "Basics of software-tools for pedestrian movement-identification and results," *Fire Technology*, vol. 48, no. 1, pp. 105–125, 2012.
- [32] N. T. N. Anh, Z. J. Daniel, N. H. Du, A. Drogoul, and V. D. An, "A hybrid macro-micro pedestrians evacuation model to speed up simulation in road networks," in *Advanced Agent Technology*, vol. 7068, pp. 371–383, Springer, 2012.
- [33] P. A. Thompson and E. W. Marchant, "A computer model for the evacuation of large building populations," *Fire Safety Journal*, vol. 24, no. 2, pp. 131–148, 1995.
- [34] W. J. Yu, R. Chen, L. Y. Dong, and S. Q. Dai, "Centrifugal force model for pedestrian dynamics," *Physical Review E*, vol. 72, no. 2, Article ID 026112, 2005.
- [35] Q. Zhang, B. Han, and D. Li, "Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations," *Transportation Research C*, vol. 16, no. 5, pp. 635–649, 2008.
- [36] S.-K. Chen, S.-Y. Li, X. Li, J. Hong, J.-X. Lai, and P. Cai, "Modeling evacuation time for passengers from metro platforms," *Journal of Transportation Systems Engineering and Information Technology*, vol. 8, no. 4, pp. 101–107, 2008.
- [37] S.-K. Chen, S. Liu, X. Xiao, J. Hong, and B.-H. Mao, "M/G/c/c-based model of passenger evacuation capacity of stairs and corridors in metro stations," *Journal of the China Railway Society*, vol. 34, no. 1, pp. 7–12, 2012.
- [38] S. Feng, N. Ding, T. Chen, and H. Zhang, "Simulation of pedestrian flow based on cellular automata: a case of pedestrian crossing street at section in China," *Physica A*, vol. 392, pp. 2847–2859, 2013.
- [39] B. X. Song, Z. Z. Wu, and Z. H. Xie, "Study on occupant evacuation model considering influence of direction sign," *China Safety Science Journal*, vol. 12, pp. 27–33, 2011.
- [40] L.-X. Yang, X.-M. Zhao, Z.-Y. Gao, and J.-F. Zheng, "Bi-directional pedestrian flow model with traffic convention," *Acta Physica Sinica*, vol. 60, no. 10, Article ID 100501, 2011.
- [41] X. Zheng, T. Zhong, and M. Liu, "Modeling crowd evacuation of a building based on seven methodological approaches," *Building and Environment*, vol. 44, no. 3, pp. 437–445, 2009.
- [42] Y. Bo, W. Yong-gang, and W. Cheng, "A multi-agent and GIS based simulation for emergency evacuation in park and public square," in *Proceedings of the IEEE International Conference on Computational Intelligence and Security (CIS '09)*, pp. 211–216, 2009.

- [43] R. Y. Guo and H. J. Huang, "Theoretical analysis and simulation of pedestrian evacuation under invisible conditions," *Simulation*, vol. 88, pp. 1138–1148, 2012.
- [44] R.-Y. Guo, H.-J. Huang, and S. C. Wong, "Route choice in pedestrian evacuation under conditions of good and zero visibility: experimental and simulation results," *Transportation Research B*, vol. 46, pp. 669–686, 2012.
- [45] Y. Xu and H.-J. Huang, "Simulation of exit choosing in pedestrian evacuation with consideration of the direction visual field," *Physica A*, vol. 391, no. 4, pp. 991–1000, 2012.

Research Article

Networked Timetable Stability Improvement Based on a Bilevel Optimization Programming Model

Xuelei Meng,¹ Bingmou Cui,¹ Limin Jia,² Yong Qin,² and Jie Xu²

¹ School of Traffic and Transportation, Lanzhou Jiaotong University, P.O. Box 405, Anning West Road, Anning District, Lanzhou, Gansu 730070, China

² State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, No. 3 Shangyuancun, Haidian District, Beijing 100044, China

Correspondence should be addressed to Xuelei Meng; mengxuelei@gmail.com

Received 28 November 2013; Revised 13 January 2014; Accepted 23 January 2014; Published 4 March 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Xuelei Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Train timetable stability is the possibility to recover the status of the trains to serve as arranged according to the original timetable when the trains are disturbed. To improve the train timetable stability from the network perspective, the bilevel programming model is constructed, in which the upper level programming is to optimize the timetable stability on the network level and the lower is to improve the timetable stability on the dispatching railway segments. Timetable stability on the network level is defined with the variances of the utilization coefficients of the section capacity and station capacity. Weights of stations and sections are decided by the capacity index number and the degrees. The lower level programming focuses on the buffer time distribution plan of the trains operating on the sections and stations, taking the operating rules of the trains as constraints. A novel particle swarm algorithm is proposed and designed for the bilevel programming model. The computing case proves the feasibility of the model and the efficiency of the algorithm. The method outlined in this paper can be embedded in the networked train operation dispatching system.

1. Introduction

Train timetable is the fundamental file for organizing the railway traffic, which determines the inbound and outbound time of trains. Railways are typically operated according to a planned (predetermined) timetable, and the quality of the timetable determines the quality of the railway service. So it is most important to map a high quality timetable for all kinds of trains.

But there is a dilemma that we place as much as possible trains on the timetable chart, and simultaneously we should enhance the possibility to adjust the timetable when disruptions occur. The randomly occurring disturbances may cause train delays and even disrupt the entire train operation plan. In the railway network, every station and section are planned to serve the trains according to the schedule, often compactly. So a slightly delayed train may cause a domino effect of secondary delays over the thorough network. Although the buffer times added to the minimum

running time in the sections and minimum dwell at stations in scheduled timetables may absorb some train delays and assure some degree of timetable stability, the large buffer time will reduce the capacity of the railway.

Therefore, to ensure both the capacity and the order of the train operation, a reliable, stable, robust timetable, and the feasible efficient rescheduling of the planned timetable must be worked out. A superior quality timetable cannot only decide the inbound and outbound time at stations, and the more important, can offer the possibility to recover the operation according to the planned timetable when the trains are disturbed by accidents randomly. Timetable stability is the index to measure the *possibility*. Timetable stability is related to the train number assigned to the railway sections and the buffer time distributed to each station and section, the probability that the train is disrupted at the stations and in the sections. In this paper, the buffer time refers to the time added to the minimum running time in a section. It equals the planned period of running minus the minimum running

time in the section. And it also refers to the time added to the dwelling time at a station, which equals the planned dwelling time at the station minus the minimum dwelling time.

So when assigning the trains paths and mapping out the train schedule, the train number assigned to the sections and buffer time distribution should be designed carefully, not only considering the section capacity and station capacity, but also the minimal running time at each section and the minimal dwell on the station.

We define the networked timetable stability quantitatively, considering that the railway network with the goal is to optimize the timetable stability to offer more possibilities to reschedule the trains on the railway network to deal with disturbs in the train operation process.

The outline of the paper is as follows. First, Section 2 proposes the literature review on timetable stability improvement. Section 3 builds the bilevel optimization model for the networked timetable stability. Then Section 4 introduces the hybrid fuzzy particle swarm algorithm, improving the velocity equation. Section 5 applies hybrid fuzzy particle swarm optimization algorithm in solving the bilevel model for improving the networked timetable stability. The computing case is presented in Section 6. Finally, Section 7 gives some conclusions.

2. Literature Review

It is a hot topic now to assure the reliability, safety, and stability of the traffic control system as discussed in [1–3]. The timetable stability optimizing is a relatively new issue in the field of railway operation research. The research work was published in the 1990s.

The research experienced two developing periods. In the earliest period, the focus was the timetable on the dispatching section, according to the operation mode of the railway and the basis to study the train timetable stability is formed. A discrete dynamic system model was built to describe the timetable with the max-plus algebra based on the discussion of the timetable periodicity and analyzed the timetable stability, as proposed by Goverde [4]. Carey and Carville developed a simulation model to test the schedule performance and reliability for train stations in [5]. Hansen pointed out that the effect of the stochastic disturbance on trains relied on the adjustment of the running time and buffer time in the timetable and assessed the advantages and the disadvantages of the capacity and stability of evaluating model [6]. These researchers promoted the train timetable stability theory from the perspective of the running time in railway sections, the dwelling time on railway stations, and the buffer time for running and dwelling. De Kort et al. proposed a method to evaluate the capacity determined by the timetable and took the timetable stability as a part of the capacity; see [7]. The goal was tantamount to place train running lines as much as possible, while taking the timetable into consideration at the same time. Goverde presented a method based on max-plus algebra to analyze the timetable stability. He proved the feasibility of the method with data of the Netherlands national railway timetable; see [8]. We

defined and qualified timetable stability and took it as a goal when rescheduling trains on the dispatching sections in [9]. So it is easy to understand that the time is the key factor when studying the timetable of a dispatching railway section. Focusing on the delay time, the behind schedule ratio, the buffer time, and time deviation, researchers studied the timetable adjustability, equilibrium, stability, using statistics theories, max-plus algebra, and so forth.

Research on timetable stability progressively expanded to the railway network, for the study focusing on the timetable stability of dispatching cannot suit the networked timetable design and optimization. Engelhardt-Funke and Kolonko considered a network of periodically running railway lines. They built a model to analyze stability and investments in railway networks and designed an innovative evolutionary algorithm to solve the problem in [10]. Goverde analyzed the dependence of the timetable on the busy degree of the railway network. He again hired the max-plus algebra, to analyze the timetable stability of the railway network. On this basis, he proposed a novel method to generate the paths for the trains on a large-scale railway network; see [11]. Vromans built a complex linear programming model to optimize the timetable on the railway network level, taking the total delayed time as the optimizing goal. And they designed the stochastic optimization algorithm for the model; see [12]. Delorme et al. presented a station capacity evaluating model and evaluated the stability on the key parts of the railway network: stations; see [13]. We analyzed the complex characteristic analysis of passenger train flow network in the former study work [14] and have done some research work to support the networked train timetable stability optimization, from transportation capacity calculation [15], paths generating [16], and line planning [17], which can be seen as the constraint of timetable stability optimizing.

And we can see that the networked timetable stability is related to not only time but also the utilization coefficient capacity of the railway network, as discussed in [11–13]. That is to say, the networked timetable stability study requires the combination of the railway network capacity utilization and the buffer time distribution of the buffer time in the sections and at the stations. However, most of the publications are about the stability of the timetable for a definite dispatching railway section. And there are limited publications about the networked timetable stability. Furthermore, the research on the timetable is in the stage of evaluating, mostly qualitative, not the quantization of the timetable stability.

3. The Bilevel Optimization Programming Model for Networked Timetable Stability Improvement

Networked timetable stability must be studied from two levels. The upper level is to study the relation between the trains flow and the capacity of the sections and stations and the ability to recover the timetable when an emergency occurs determined by the relation. The lower level is to study the distribution plan of the buffer time for each train in

the sections running process and the stations dwelling, to eliminate the negative effects of the disturbs.

The goal of the upper programming is to decide the number of trains assigned on each railway section and at the stations. The fundamental restriction is that the number of trains assigned to the sections and stations must not exceed the capacity of the sections and the stations. And the number of trains received by the stations must be equal to the total number of the trains running through the sections which are connected to the relative station.

The lower programming is to determine the buffer time distribution plan. The running time through a whole section planning in the timetable is more than that it requires if it runs at its highest speed. So there is a period of time called buffer time that can be distributed for the sections running and stations dwelling to absorb the delay caused by the random disturbances. The restriction is that buffer time allocated to each station and section must be longer than or equal to zero.

3.1. The Timetable Stability Improvement Programming on the Network Level. To define the timetable stability on the network level, the load on the sections and stations is the key factor. So the load index numbers must be defined first.

Definition 1. The load index number of a station on the railway network is

$$Z_{ST} = \text{Var} (e^{-\rho_i} w_{ST,i}^{\rho_i - \rho}), \quad (1)$$

where Var is the function to calculate the variance of a vector, ρ_i is the load of the i th station; the bigger ρ_i is, the smaller the stability value is. $\rho_i = F_i/B_i$, B_i is the receiving and sending capacity of the station. F_i is the number of the receiving and sending trains by the i th station according to the trains distribution plan. ρ is a threshold value of a station load. $w_{ST,i}$ is the weight of the i th station, and K is the number of the stations on the railway network.

The index number of the capacity of a station is

$$IX_{ST,i} = \frac{B_i D_i}{\sum_{i=1}^K B_i D_i}, \quad (2)$$

where D_i is the degree of the i th station.

Then the station weight is

$$w_{ST,i} = \frac{IX_{ST,i}}{\sum_{i=1}^K IX_{ST,i}}. \quad (3)$$

Definition 2. The load index number of a section on the railway network is

$$Z_{SE} = \text{Var} (e^{-\lambda_i} w_{SE,i}^{\lambda_i - \lambda}), \quad (4)$$

where Var is the function to calculate the variance of a vector, λ_i is the load of the i th section; the bigger λ_i is, the smaller the stability value is. $\lambda_i = G_i/C_i$, C_i is the capacity of the section. G_i is the number of the trains running through i th section according to the trains distribution plan. λ is a threshold value

of a section load. w_i is the weight of the i th section, and L is the number of the sections on the railway network.

The index number of the capacity of a section is

$$IX_{SE,i} = \frac{C_i}{\sum_{i=1}^L C_i}. \quad (5)$$

Then the weight of the section is

$$w_{SE,i} = \frac{IX_{SE,i}}{\sum_{i=1}^L IX_{SE,i}}. \quad (6)$$

Then with the load index numbers of the stations and sections, the timetable stability on the network level is defined as

$$S_{NET} = e^{-Z_{ST}} \times e^{-Z_{SE}} = e^{-(Z_{ST} + Z_{SE})}. \quad (7)$$

The goal of the upper programming is to optimize the timetable stability on the network level, so S_{NET} is taken as the optimization goal. That is to say, the goal is to maximize the timetable on the network level: S_{NET} .

Restrictions require that the number of the trains running through a section cannot be greater than the number of the trains that the section capacity allows. Likewise, the total trains number going through a station cannot exceed the station capacity of receiving and sending off trains.

And the total numbers of the trains distributed on the sections connected to the station must be equal to the number of arriving trains at the station:

$$\begin{aligned} F_i &\leq B_i, \\ G_i &\leq C_i, \\ F_i &= \sum_{l=1}^U G_{i,l}, \end{aligned} \quad (8)$$

where U is the number of sections connected to station i .

3.2. The Timetable Stability Improvement Programming on the Dispatching Section Level. Take it for granted that there are M trains going through section p , which is the result of the upper programming. The running times of all the M trains form a vector $T_R^p = \{t_{R,i}^p\}_M$. The minimum running time of all the M trains forms a vector $T_R^{p,\min} = \{t_{R,i}^{p,\min}\}_M$. Then the margin vector of the M trains is $\Delta T_R^p = \{\Delta t_{R,i}^p\}_M = \{t_{R,i}^p - t_{R,i}^{p,\min}\}_M$. Set the $A_R^p = \{\Delta t_{R,i}^p / t_{R,i}^p\}_M$ to be the running adjustability vector. To evaluate the equilibrium of the distribution of the buffer time in the sections, the running adjustability dispersion is defined as

$$\text{Var} (A_R^p) = \text{Var} \left(\frac{\Delta t_{R,i}^p}{t_{R,i}^p} \right). \quad (9)$$

The smaller the value of the $\text{Var}(A_R^p)$ is, the more balanced the buffer time distribution plan is, and the timetable is more stable.

Likewise, take it for granted that there are N trains going through station q , with stop or without stop. The planned

dwelling time according to the timetable of the N trains forms a vector $T_D^q = \{t_{D,i}^q\}_N$. The minimal dwelling time of all the N trains at q also forms a vector $T_D^{p,\min} = \{t_{D,i}^{p,\min}\}_N$. Then the margin vector of the N trains is $\Delta T_D^q = \{\Delta t_{D,i}^q\}_N = \{t_{D,i}^q - t_{D,i}^{q,\min}\}_N$. Set the $A_D^q = \{\Delta t_{D,i}^q / t_{D,i}^q\}_N$ to be the dwelling adjustability vector. To evaluate the equilibrium of the distribution of the buffer time at stations, the adjustability dispersion is defined as

$$\text{Var}(A_D^q) = \text{Var}\left(\frac{\Delta t_{D,i}^q}{t_{D,i}^q}\right). \quad (10)$$

The smaller the value of the $\text{Var}(A_D^q)$ is, the more balanced the buffer time distribution plan is, and the timetable is more stable.

On the basis of considering of the running adjustability dispersion and the dwelling adjustability dispersion, the timetable stability on the dispatching section level is defined as

$$S_{\text{DIS}} = e^{-\text{Var}(A_R^p) \times \text{Var}(A_D^q)} = e^{-\text{Var}(\Delta t_{R,i}^p / t_{R,i}^p) \times \text{Var}(\Delta t_{D,i}^q / t_{D,i}^q)}. \quad (11)$$

Then we take the timetable stability on the network level as the optimizing goal of the upper programming:

$$\max S_{\text{DIS}}. \quad (12)$$

When rescheduling the trains on the sections, the minimum running time and the minimum dwelling time must be considered. The rescheduled running time and dwelling time must be longer than the minimum time, which is described in (13) and (14). And the margins between inbound times of different trains at the same stations must be bigger than the minimum interval time to ensure the safety of train operation. This constraint is defined in (15). Likewise, the margins between outbound times of the trains at a station have the same characteristic, as shown in (16). And the number of the trains dwelling at a same station cannot be bigger than the number of the tracks in a station, described in (17):

$$t_{R,i}^p \geq t_{R,i}^{p,\min}, \quad (13)$$

$$t_{R,i}^q \geq t_{R,i}^{q,\min}, \quad (14)$$

$$\left| a_{i,j}^p - a_{l,j}^p \right| > I_{a-a}, \quad l \neq i, \quad (15)$$

$$\left| d_{i,j}^p - d_{l,j}^p \right| > I_{d-d}, \quad l \neq i, \quad (16)$$

$$LN_j - \sum_{p \in P} \sum_{k=1}^{\text{TN}} n_{p,j}^k \geq 0. \quad (17)$$

3.3. The Networked Timetable Stability Definition. Depending on the analysis in Sections 3.1 and 3.2, networked timetable stability is defined with the following:

$$S = S_{\text{NET}} \times S_{\text{DIS}}. \quad (18)$$

We can see that the networked timetable stability is directly related to timetable stability on the network level and the dispatching section level. The programming in Sections 3.1 and 3.2 can optimize the networked timetable stability, through optimizing the stability on the two levels.

4. The Hybrid Fuzzy Particle Swarm Optimization Algorithm

4.1. Fuzzy Particles Swarm Optimization Algorithm. Considerable attention has been paid to fuzzy particle swarm optimization (FPSO) recently. Abdelbar et al. proposed the FPSO [18]. Abdelbar and Abdelshahid brought forward the instinct-based particle swarm optimization with local search applied to satisfiability in [19]. Abdelshahid analyzed variations of particle swarm optimization and gave evaluation on maximum satisfiability; see [20]. Mendes et al. proposed a fully informed particle swarm in [21]. Bajpai and Singh studied the problem of fuzzy adaptive particle swarm optimization (FAPSO) for bidding strategy in uniform price spot market in [22]. Saber et al. attempted to solve the problem of unit commitment computation by FAPSO in [23]. Esmin studied the problem of generating fuzzy rules and fitting fuzzy membership functions using hybrid particle swarm optimization (HPSO); see [24, 25].

Computation in the PSO paradigm is based on a collection (called a swarm) of fairly primitive processing elements (called particles). The neighborhood of each particle is the set of particles with which it is adjacent. The two most common neighborhood structures are p_g , in which the entire swarm is considered a single neighborhood, and p_i , in which the particles are arranged in a ring, and each particle's neighborhood consists of itself, its immediate ring-neighbor to the right, and its immediate ring-neighbor to the left.

PSO can be used to solve a discrete combinatorial optimization problem whose candidate solutions can be represented as vectors of bits; i is supposed to be a given instance of such a problem. Let N denote the number of elements in the solution vector for i . Each particle i would contain two N -dimensional vectors: a Boolean vector x_i , which represents a candidate solution to i and is called particle i 's state, and a real vector v_i , called the velocity of the particle. In the biological insect-swarm analogy, the velocity vector represents how fast, and in which direction, the particle is flying for each dimension of the problem being solved.

Let $K(i)$ denote the neighbors of particle i , and let p_i denote the best solution ever found by particle i . In each time iteration, each particle i adjusts its velocity based on

$$v_{i+1} = \omega v_i + c_1 r_1 (p_i - x_i) + c_2 r_2 (p_g - x_i), \quad (19)$$

$$x_{i+1} = x_i + v_{i+1},$$

where ω , called inertia, is a parameter within the range $[0, 1]$ and is often decreased over time as discussed in [26]; c_1 and c_2 are two constants, often chosen so that $c_1 + c_2 = 4$, which control the degree to which the particle follows the herd thus stressing exploitation (higher values of c_2), or goes its own way thus stressing exploration (higher values of c_1); r_1 and r_2 are uniformly random number generator function that returns values within the interval $(0, 1)$; and p_g is the particle in i 's neighborhood with the current neighborhood-best candidate solution.

Fuzzy PSO differs from standard PSO in only one respect: in each neighborhood, instead of only the best particle in

the neighborhood being allowed to influence its neighbors, several particles in each neighborhood can be allowed to influence others to a degree that depends on their degree of charisma, where charisma is a fuzzy variable. Before building a model, there are two essential questions that should be answered. The first question is how many particles in each neighborhood have nonzero charisma. The second is what membership function (MF) will be utilized to determine level of charisma for each of the k selected particles.

The answer to the first question is that the k best particles in each neighborhood are selected to be charismatic, where k is a user-set parameter, k can be adjusted according to the required precision of the solution.

The answer to the other question is that there are numerous possible functions for charisma MF. Popular MF choices include triangle, trapezoidal, Gaussian, Bell, and Sigmoid MFs; see [27]. The Bell, Gaussian, sigmoid, Trapezoidal, and Triangular MFs

$$\text{bell}(x; c, \sigma) = \frac{1}{1 + ((x - c) / \sigma)^2},$$

$$\text{gaussian}(x; c, \sigma) = \exp\left(-\frac{1}{2}\left(\frac{x - c}{\sigma}\right)^2\right),$$

$$\text{sig}(x; c, \sigma) = \frac{1}{1 + \exp(-\sigma(x - c))},$$

$$\text{triangle}(x; a, b, c) = \max\left(\min\left(\frac{x - a}{b - a}, \frac{c - x}{c - b}\right), 0\right),$$

$$a < b < c,$$

$$\text{trapezoid}(x; a, b, c, d) = \max\left(\min\left(\frac{x - a}{b - a}, 1, \frac{d - x}{d - c}\right), 0\right),$$

$$a < b \leq c < d. \quad (20)$$

Let h be one of the k -best particles in a given neighborhood, and let $f(p_g)$ refer to the fitness of the very-best particle for the neighborhood under consideration. The charisma $\varphi(h)$ is defined as

$$\varphi(h) = \frac{1}{1 + ((f(p_h) - f(p_g)) / \beta)^2} \quad (21)$$

if the MF is based on Bell function.

The charisma $\varphi(h)$ is defined as

$$\varphi(h) = \exp\left(-\frac{1}{2}\left(\frac{f(p_h) - f(p_g)}{\beta}\right)^2\right) \quad (22)$$

if the MF is based on Gaussian function.

The charisma $\varphi(h)$ is defined as

$$\varphi(h) = \frac{1}{1 + \exp(-\beta(f(p_h) - f(p_g)))} \quad (23)$$

if the MF is based on Sigmoid function.

The charisma $\varphi(h)$ is defined as

$$\begin{aligned} \varphi(h) &= \max\left(\min\left(\frac{f(p_h) - f(p_g^0)}{f(p_g) - f(p_g^0)}, \frac{f(p_g^1) - f(p_h)}{f(p_g^1) - f(p_g)}\right), 0\right) \end{aligned} \quad (24)$$

if the MF is based on Triangle function.

The charisma $\varphi(h)$ is defined as

$$\begin{aligned} \varphi(h) &= \max\left(\min\left(\frac{f(p_h) - f(p_g^0)}{f(p_g) - f(p_g^0)}, 1, \frac{f(p_g^1) - f(p_h)}{f(p_g^1) - f(p_g)}\right), 0\right) \end{aligned} \quad (25)$$

if the MF is based on Trapezoid function.

Because $(p_h) \leq f(p_g)$, $\varphi(h)$ is a decreasing function that is 1 when $f(p_h) = f(p_g)$, and asymptotically approaches zero as $f(p_h)$ moves away from $f(p_g)$. To avoid dependence on the scale of the fitness function, β is defined as $\beta = f(p_g) / \ell$ where ℓ is a user-specified parameter. For a fixed $f(p_h)$, the larger the value of ℓ , the smaller the charisma $\varphi(h)$ will be. $f(p_g^0)$ and $f(p_g^1)$ are two functional values that decide the verge of the triangle and trapezoid function.

In Fuzzy PSO, velocity equation is

$$v_{i+1} = \omega v_i + c_1 r_1 (p_i - x_i) + \sum_{h \in B(i,k)} \varphi(h) c_2 r_2 (p_h - x_i), \quad (26)$$

where $B(i, k)$ denotes the k -best particles in the neighborhood of particle i . Each particle i is influenced by its own best solution p_i and the best solutions obtained by the k charismatic particles in its neighborhood, with the effect of each weighted by its charisma $\varphi(h)$. It can be seen that if k is 1, this model reduces to the standard PSO model.

4.2. Hybrid Fuzzy PSO. Hybrid rule requires selecting two particles from the alternative particles at a certain rate. Then the intersecting operation work needs to be done to generate the descendant particles. The positions and velocities of the descendant particles are as follows: according to the intersecting rule, inheriting from the FPSO, see [28]

$$v_{i+1}^1 = \omega v_i^1 + c_1 r_1 (p_i - x_i^1) + \sum_{h \in B(i,k)} \varphi(h) c_2 r_2 (p_h^1 - x_i^1),$$

$$v_{i+1}^2 = \omega v_i^2 + c_1 r_1 (p_i - x_i^2) + \sum_{h \in B(i,k)} \varphi(h) c_2 r_2 (p_h^2 - x_i^2),$$

$$v_{i+1}^1 = \frac{v_{i+1}^1 + v_{i+1}^2}{|v_{i+1}^1 + v_{i+1}^2|} |v_{i+1}^1|,$$

$$v_{i+1}^2 = \frac{v_{i+1}^1 + v_{i+1}^2}{|v_{i+1}^1 + v_{i+1}^2|} |v_{i+1}^2|,$$

$$x_{i+1}^1 = x_i^1 + v_{i+1}^1,$$

$$x_{i+1}^2 = x_i^2 + v_{i+1}^2,$$

$$x_{i+1}^1 = p \times x_{i+1}^1 + (1 - p) \times x_{i+1}^2,$$

$$x_{i+1}^2 = p \times x_{i+1}^2 + (1 - p) \times x_{i+1}^1. \quad (27)$$

Thus, the HFPSO is built. In the model, x is a position vector with D dimensions. x_i^j stands for the particle j of the i th generation particles. p is a random variable vector with D dimensions which obeys the equal distribution. Each dimension of p is in $[0, 1]$.

4.3. Adaptability and Applicability HFPSO. It can be seen that the network timetable stability optimizing model is a nonlinear one and it is an NP-hard problem. Generally, it is very difficult to solve the problem with mathematical approaches. Evolutionary algorithms are often hired to solve the problem for their characteristics. First, its rule of the algorithm is easy to apply. Second, the particles have the memory ability which results in convergent speed and there are various methods to avoid the local optimum. Thirdly, the parameters which need to select are fewer, and there is considerable research work on the parameters selecting.

In addition, the HFPSO hires the fuzzy theory and the hybrid handling method when designing the algorithm. Thus it has the ability to improve the computing precision when solving the optimization problem. And it utilizes intersecting tactics to generate the new generation of particles to avoid the precipitate of the solution. It is adaptive to solve the timetable stability optimization. And its easy computing rule determines the applicability in the solving of timetable stability optimization problem.

5. Hybrid Fuzzy Particle Swarm Algorithm for the Model

5.1. The Particle Swarm Design for the Upper Level Programming. The size of the particle swarm is set to be 30 to give consideration to both the calculating degree of accuracy and computational efficiency. For F_i and G_i are the decision variables, L is the number of the sections on the railway network. K is the number of the stations on the railway network. So a particle can be designed as

$$pa_{UP} = \{F_1, F_2, \dots, F_K, G_1, G_2, \dots, G_L\}. \quad (28)$$

5.2. The Particle Swarm Design for the Lower Level Programming. The size of the particle swarm is also set to be 30. $\Delta t_{R,i}^p$ and $\Delta t_{D,i}^q$ are the decision variables. So a particle can be designed as

$$pa_{LO} = \left\{ \begin{aligned} &\Delta t_{R,1}^1, \Delta t_{R,2}^1, \dots, \Delta t_{R,M_1}^1, \\ &\Delta t_{R,1}^2, \Delta t_{R,2}^2, \dots, \Delta t_{R,M_2}^2, \dots \\ &\Delta t_{R,1}^i, \Delta t_{R,2}^i, \dots, \Delta t_{R,M_i}^i, \dots \\ &\Delta t_{R,1}^L, \Delta t_{R,2}^L, \dots, \Delta t_{R,M_L}^L, \\ &\Delta t_{D,1}^1, \Delta t_{D,2}^1, \dots, \Delta t_{D,M_1}^1, \\ &\Delta t_{D,1}^2, \Delta t_{D,2}^2, \dots, \Delta t_{D,M_2}^2, \dots \\ &\Delta t_{D,1}^j, \Delta t_{D,2}^j, \dots, \Delta t_{D,M_j}^j, \dots \\ &\Delta t_{D,1}^K, \Delta t_{D,2}^K, \dots, \Delta t_{D,M_K}^K \end{aligned} \right\}, \quad (29)$$

where M_i is the number of trains going through section i and M_j is the number of the trains going through station j .

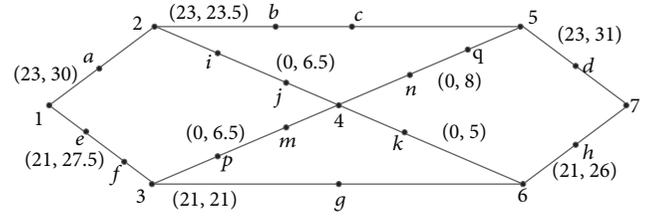


FIGURE 1: The railway network in the computing case and the original distribution plan of the trains.

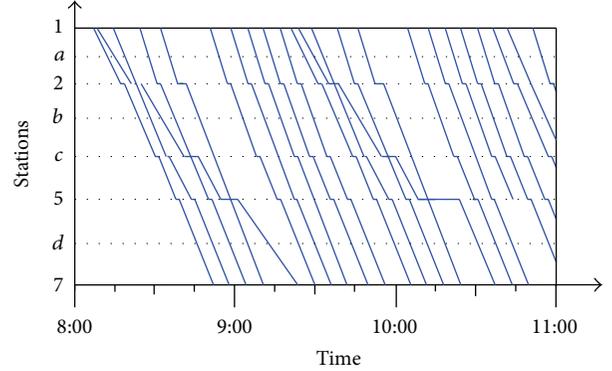


FIGURE 2: Planned operation diagram on path 1-2-5-7.

6. Experimental Results and Discussion

6.1. Computing Case Assumptions. There are 22 stations and 10 sections in the network, as indicated in Figure 1. We assume that the 44 trains operate on the network from station 1 to station 7. The numbers in parentheses beside the edges are the numbers of the trains distributed on the sections according to the planned timetable and the section capacity.

And the planned operation diagram on path 1-2-5-7 is shown in Figure 2, with 23 trains. The planned operation diagram on path 1-3-6-7 is illustrated in Figure 3.

According to the networked timetable stability definition in Section 3, the timetable stability value of the planned timetable can be calculated out. Detailed computing data are presented in Table 1.

According to Table 1(a), the Z_{ST} can be calculated with $Z_{ST} = \text{Var}(e^{-\rho_i} w_{ST,i}^{\rho_i - p}) = 1.047978$, while Z_{SC} can be calculated out with $Z_{SE} = \text{Var}(e^{-\lambda_i} w_i^{\lambda_i - \lambda}) = 7.1940742$. Then the goal of the upper programming $S_{NET} = e^{-Z_{ST}} \times e^{-Z_{SE}} = e^{-(Z_{ST} + Z_{SE})} = 0.000263$.

6.2. The Upper Level Computing Results and Discussion. We reallocate all the 44 trains on the modest railway network, as shown in Figure 4, according to the computing results of the upper level programming. The numbers in parentheses beside the edges are the numbers of the trains allocated on the sections according to the rescheduled timetable and the section capacity. Based on the capacity of every section and station, the computing result of the upper level programming is presented in Table 2.

TABLE 1: Computing results of the planned timetable stability.

(a) Related computing results of Z_{ST}

Key nodes	Trains number through station	Nodes capacity	Load	Capacity index	Nodes degree	Degree index	Stations weight
1	44	66.0	0.6667	0.3099	2.0000	0.1111	0.2245
2	23	34.5	0.6667	0.1620	3.0000	0.1667	0.1760
3	21	31.5	0.6667	0.1479	3.0000	0.1667	0.1607
4	0	15.0	0.0000	0.0704	4.0000	0.2222	0.1020
5	23	36.0	0.6389	0.1690	3.0000	0.1667	0.1837
6	21	30.0	0.7000	0.1408	3.0000	0.1667	0.1531

(b) Related computing results of Z_{SC}

Section	Trains number through section	Sections capacity	Load	Capacity index	Sections weight
1-2	23	30.0	0.7667	0.1622	0.1622
1-3	21	27.5	0.7636	0.1486	0.1486
2-4	0	6.5	0.0000	0.0351	0.0351
2-5	23	23.5	0.9787	0.1270	0.1270
3-4	0	6.5	0.0000	0.0351	0.0351
3-6	21	21.0	1.0000	0.1135	0.1135
4-5	0	8.0	0.0000	0.0432	0.0432
4-6	0	5.0	0.0000	0.0270	0.0270
5-7	23	31.0	0.7419	0.1676	0.1676
6-7	21	26.0	0.8077	0.1405	0.1405

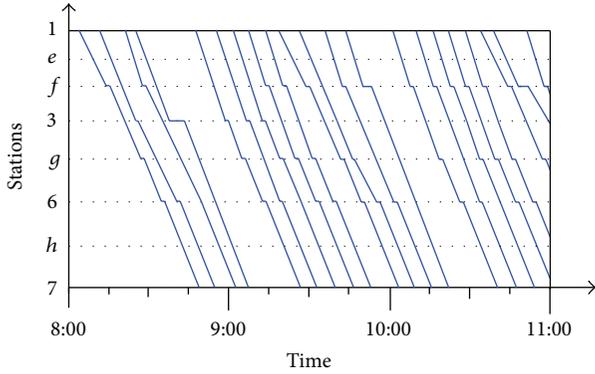


FIGURE 3: Planned operation diagram on path 1-3-6-7.

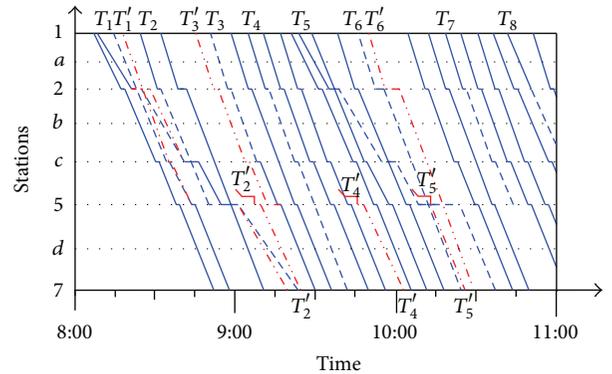


FIGURE 5: Rescheduled timetable of path 1-2-5-7 according to the computing results of the lower level programming.

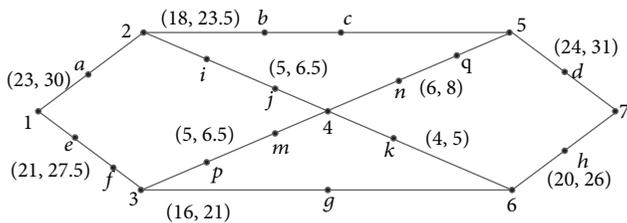


FIGURE 4: Distribution plan of the trains according to the computing results.

According to Table 2(a), the Z_{ST} can be calculated with $Z_{ST} = \text{Var}(e^{-\rho_i} w_{ST,i}^{\rho_i - \rho}) = 0.000223814$, while Z_{SC} can be calculated out with $Z_{SE} = \text{Var}(e^{-\lambda_i} w_i^{\lambda_i - \lambda}) = 0.002432614$. Then the goal of the upper programming $S_{NET} = e^{-Z_{ST}} \times e^{-Z_{SE}} = e^{-(Z_{ST} + Z_{SE})} = 0.997$.

6.3. *The Lower Level Computing Results and Discussion.* According to the computing results of the lower level programming, we adjust the timetable, moving some of the running lines of the trains. The two-dot chain line is the newly planned running trajectory and the dotted line is the previously planned trajectory. The timetable on path 1-2-5-7 is shown in Figure 5. Figure 6 is the timetable of path 2-4-5. Figures 7 and 8 are the timetables on paths 1-3-6-7 and 3-4-6, respectively.

From Figures 5 and 6 we can see that eight trains are rescheduled on path 1-2-5-7, which are numbered $T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8$. T_1, T_3 , and T_6 run on the original path as planned, but the inbound and outbound times at the stations are changed. T_2, T_4, T_5, T_7 , and T_8 modify the path when they arrive at station 2. They run through path 2-4-5, with arriving time 8:32, 9:12, 9:41, 10:25, and 10:48, respectively at station 2.

TABLE 2: Computing results of rescheduled timetable stability.

(a) Related computing results of Z_{ST}							
Key nodes	Trains number through station	Nodes capacity	Load	Capacity index	Nodes degree	Degree index	Stations weight
1	44	66.0	0.6667	0.3099	2	0.1111	0.2245
2	23	34.5	0.6667	0.1620	3	0.1667	0.1760
3	21	31.5	0.6667	0.1479	3	0.1667	0.1607
4	10	15.0	0.6667	0.0704	4	0.2222	0.1020
5	24	36.0	0.6667	0.1690	3	0.1667	0.1837
6	20	30.0	0.6667	0.1408	3	0.1667	0.1531

(b) Related computing results of Z_{SC}					
Section	Trains number through section	Sections capacity	Load	Capacity index	Sections weight
1-2	23	30.0	0.7667	0.1622	0.1622
1-3	21	27.5	0.7636	0.1486	0.1486
2-4	5	6.5	0.7692	0.0351	0.0351
2-5	18	23.5	0.7660	0.1270	0.1270
3-4	5	6.5	0.7692	0.0351	0.0351
3-6	16	21.0	0.7619	0.1135	0.1135
4-5	6	8.0	0.7500	0.0432	0.0432
4-6	4	5.0	0.8000	0.0270	0.0270
5-7	24	31.0	0.7742	0.1676	0.1676
6-7	20	26.0	0.7692	0.1405	0.1405

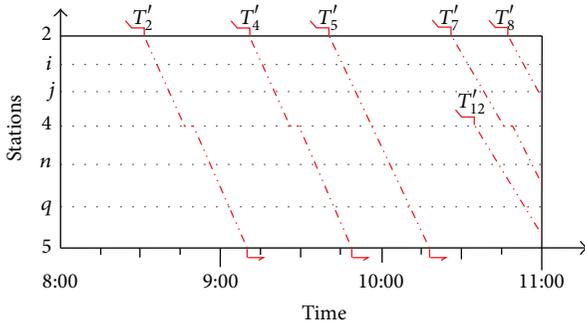


FIGURE 6: Rescheduled timetable of path 2-4-5 according to the computing results of the lower level programming.

T_2 , T_4 , and T_5 leave arrive station 5 at 9:10, 9:48, and 10:18, respectively, then finish the rest travel along Sections 5-7.

From Figures 7 and 8 we can see that eight trains are rescheduled on path 1-3-6-7, which are numbered T_9 , T_{10} , T_{11} , T_{12} , T_{13} , T_{14} , T_{15} , T_{16} . T_9 , T_{10} , T_{11} , and T_{13} run on the original path as planned, but the inbound and outbound time at the stations are changed. T_9 , T_{10} , T_{11} , T_{12} , and T_{13} change the path when they arrive at station 3. T_9 , T_{10} , T_{11} , and T_{13} run along path 3-4-6 with the arriving time 8:22, 9:16, 9:44, and 10:43 at station 3, respectively. T_9 , T_{10} , and T_{11} arrive at station 6 at 8:53, 9:47, and 10:15, respectively. T_{12} runs along the path 3-4-5. It arrives at station 3 at 10:30 and at station 4 at 10:42. From Figure 4 we can see that it runs along path 4-5 to finish the rest travel.

The timetable stability on the dispatching section level is $S_{DIS} = e^{-\text{Var}(A_R^p) \times \text{Var}(A_D^q)} = e^{-\text{Var}(\Delta t_{R,i}^p / t_{R,i}^p) \times \text{Var}(\Delta t_{D,i}^q / t_{D,i}^q)} = 0.879$.

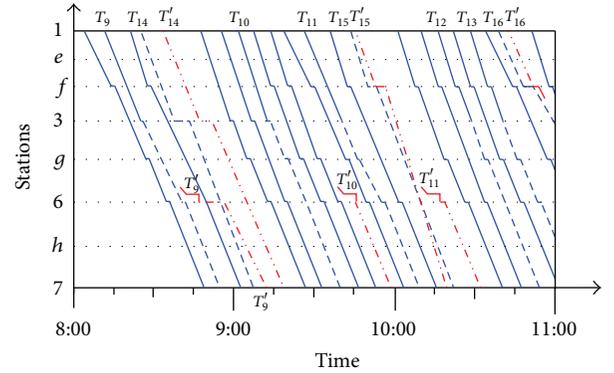


FIGURE 7: Rescheduled timetable of path 1-3-6-7 according to the computing results of the lower level programming.

Then, the networked timetable is $S = S_{NET} \times S_{DIS} = 0.997 \times 0.879 = 0.876$.

7. Conclusion

The bilevel programming model is appropriate for the networked timetable stability optimizing. It comprises the timetable stability of the network level and the dispatching section level. Better solution can be attained via hybrid fuzzy particle swarm algorithm in networked timetable optimizing. The timetable is more stable, which means that it is more feasible for rescheduling in the case of disruption, when it is optimized by hybrid fuzzy particle swarm algorithm. The timetable rearranged based on the timetable stability with bilevel networked programming model can make the

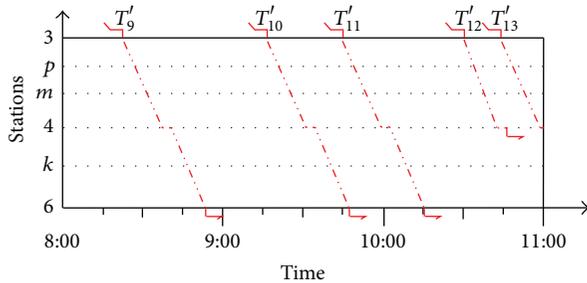


FIGURE 8: Rescheduled timetable of path 3-4-6 according to the computing results of the lower level programming.

real train movements very close to, if not the same with, the planned schedule, which is very practical in the daily dispatching work.

The results also show that hybrid fuzzy particle algorithm has significant global searching ability and high speed and it is very effective to solve the problems of timetable stability optimizing. The novel method described in this paper can be embedded in the decision support tool for timetable designers and train dispatchers.

We can do some microcosmic research work on the timetable optimizing based on the railway network in the future based on the method set out in the present paper, enlarging the research field, adding the inbound time, and outbound time of the trains at stations.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is financially supported by National Natural Science Foundation of China (Grant 61263027), Fundamental Research Funds of Gansu Province (Grant 620030), and New Teacher Project of Research Fund for the Doctoral Program of Higher Education of China (20126204120002). The authors wish to thank anonymous referees and the editor for their comments and suggestions.

References

- [1] H. Guo, W. Wang, W. Guo, X. Jiang, and H. Bubb, "Reliability analysis of pedestrian safety crossing in urban traffic environment," *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.
- [2] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research Part C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [3] W. Wang, Y. Mao, J. Jin et al., "Driver's various information process and multi-ruled decision-making mechanism: a fundamental of intelligent driving shaping model," *International Journal of Computational Intelligence Systems*, vol. 4, no. 3, pp. 297–305, 2011.
- [4] R. M. P. Goverde, "Max-plus algebra approach to railway timetable design," in *Proceedings of the 6th International Conference on Computer Aided Design, Manufacture and Operation in the Railway and Other Advanced Mass Transit Systems*, pp. 339–350, September 1998.
- [5] M. Carey and S. Carville, "Testing schedule performance and reliability for train stations," *Journal of the Operational Research Society*, vol. 51, no. 6, pp. 666–682, 2000.
- [6] I. A. Hansen, "Station capacity and stability of train operations," in *Proceeding of the 7th International Conference on Computers in Railways VII*, pp. 809–816, Bologna, Italy, September 2000.
- [7] A. F. De Kort, B. Heidergott, and H. Ayhan, "A probabilistic (max, +) approach for determining railway infrastructure capacity," *European Journal of Operational Research*, vol. 148, no. 3, pp. 644–661, 2003.
- [8] R. M. P. Goverde, "Railway timetable stability analysis using max-plus system theory," *Transportation Research Part B*, vol. 41, no. 2, pp. 179–201, 2007.
- [9] X. Meng, L. Jia, and Y. Qin, "Train timetable optimizing and rescheduling based on improved particle swarm algorithm," *Transportation Research Record*, no. 2197, pp. 71–79, 2010.
- [10] O. Engelhardt-Funke and M. Kolonko, "Analysing stability and investments in railway networks using advanced evolutionary algorithm," *International Transactions in Operational Research*, vol. 11, no. 4, pp. 381–394, 2004.
- [11] R. M. P. Goverde, *Punctuality of railway operations and timetable stability analysis [Ph.D. thesis]*, TRAIL Research School, Delft University of Technology, Delft, The Netherlands, 2005.
- [12] M. J. Vromans, *Reliability of railway systems [Ph.D. thesis]*, TRAIL Research School, Erasmus University Rotterdam, Rotterdam, The Netherlands, 2005.
- [13] X. Delorme, X. Gandibleux, and J. Rodriguez, "Stability evaluation of a railway timetable at station level," *European Journal of Operational Research*, vol. 195, no. 3, pp. 780–790, 2009.
- [14] X. Meng, L. Jia, J. Xie, Y. Qin, and J. Xu, "Complex characteristic analysis of passenger train flow network," in *Proceedings of the Chinese Control and Decision Conference (CCDC '10)*, pp. 2533–2536, Xuzhou, China, May 2010.
- [15] X.-L. Meng, L.-M. Jia, Y. Qin, J. Xu, and L. Wang, "Calculation of railway transport capacity in an emergency based on Markov process," *Journal of Beijing Institute of Technology*, vol. 21, no. 1, pp. 77–80, 2012.
- [16] X.-L. Meng, L.-M. Jia, C.-X. Chen, J. Xu, L. Wang, and J.-X. Xie, "Paths generating in emergency on China new railway network," *Journal of Beijing Institute of Technology*, vol. 19, no. 2, pp. 84–88, 2010.
- [17] L. Wang, L.-M. Jia, Y. Qin, J. Xu, and W.-T. Mo, "A two-layer optimization model for high-speed railway line planning," *Journal of Zhejiang University: Science A*, vol. 12, no. 12, pp. 902–912, 2011.
- [18] A. M. Abdelbar, S. Abdelshahid, and D. C. Wunsch II, "Fuzzy PSO: a generalization of particle swarm optimization," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '05)*, pp. 1086–1091, Montreal, Canada, August 2005.
- [19] A. M. Abdelbar and S. Abdelshahid, "Instinct-based PSO with local search applied to satisfiability," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, pp. 2291–2295, July 2004.
- [20] S. Abdelshahid, *Variations of particle swarm optimization and their experimental evaluation on maximum satisfiability [M.S. thesis]*, Department of Computer Science, American University in Cairo, May 2004.

- [21] R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: simpler, maybe better," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 204–210, 2004.
- [22] P. Bajpai and S. N. Singh, "Fuzzy adaptive particle swarm optimization for bidding strategy in uniform price spot market," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2152–2160, 2007.
- [23] A. Y. Saber, T. Senjyu, A. Yona, and T. Funabashi, "Unit commitment computation by fuzzy adaptive particle swarm optimisation," *IET Generation, Transmission and Distribution*, vol. 1, no. 3, pp. 456–465, 2007.
- [24] A. A. A. Esmín, "Generating Fuzzy rules from examples using the particle swarm optimization algorithm," in *Proceedings of the 7th International Conference on Hybrid Intelligent Systems (HIS '07)*, pp. 340–343, September 2007.
- [25] A. A. A. Esmín and G. Lambert-Torres, "Fitting fuzzy membership functions using hybrid particle swarm optimization," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 2112–2119, Vancouver, Canada, July 2006.
- [26] Y. Shi and R. Eberhart, "Modified particle swarm optimizer," in *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC '98)*, pp. 69–73, May 1998.
- [27] J. S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice-Hall, 1996.
- [28] P. J. Angeline, "Using selection to improve particle swarm optimization," in *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC '98)*, pp. 84–89, May 1998.

Research Article

A Novel Fast and Robust Binary Affine Invariant Descriptor for Image Matching

Xiujie Qu, Fei Zhao, Mengzhe Zhou, and Haili Huo

School of Information and Electronics, Beijing Institute Technology, Beijing 100081, China

Correspondence should be addressed to Xiujie Qu; quxiujie@bit.edu.cn

Received 20 November 2013; Revised 3 January 2014; Accepted 23 January 2014; Published 4 March 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Xiujie Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the current binary descriptors have disadvantages of high computational complexity, no affine invariance, and the high false matching rate with viewpoint changes, a new binary affine invariant descriptor, called BAND, is proposed. Different from other descriptors, BAND has an irregular pattern, which is based on local affine invariant region surrounding a feature point, and it has five orientations, which are obtained by LBP effectively. Ultimately, a 256 bits binary string is computed by simple random sampling pattern. Experimental results demonstrate that BAND has a good matching result in the conditions of rotating, image zooming, noising, lighting, and small-scale perspective transformation. It has better matching performance compared with current mainstream descriptors, while it costs less time.

1. Introduction

The local feature descriptor is the core of many computer vision technologies, such as object recognition, image retrieval, and 3D reconstruction. How to design a local feature descriptor that has excellent performance and low complexity is an important and difficult research. Many scholars have proposed a variety of descriptors in this area, such as SIFT (Scale Invariant Feature Transform) [1–3], which has good scale invariance and robustness to illumination and viewpoint changes. Because of constructing scale space and many other steps to improve accuracy, SIFT has high computational complexity. All of these make it prohibitively slow.

For these problems, many improved algorithms were proposed. SURF (Speeded Up Robust Features) [4] has similar matching rate with much faster performance by describing key points with the responses of few Haar-like filters compared with SIFT. PCA-SIFT [5] reduced dimensionality of the descriptor from 128 to 36 by using the principal component analysis, but the time of descriptor formation was increased. The GLOH descriptor [6] used the circular pattern in log-polar coordinates in order to enhance the robustness and uniqueness of the descriptors, but it is also

more expensive to compute than SIFT. ASIFT descriptor [7] achieved matching feature points well when there were large changes in viewpoint.

To some extent, above-mentioned descriptors improved the performances, but they still have high computational complexity because of the local histogram statistics. Besides, each dimension of the descriptor is a decimal number, which makes them need a lot of memory. All of these make them difficult to be achieved on a low-power, low-memory application. In recent years, many binary descriptors appear, such as BRIEF (Binary Robust Independent Elementary Features) descriptor [8], ORB (Oriented FAST and Rotated BRIEF) descriptor [9], BRISK (Binary Robust Invariant Scalable Keypoints) descriptor [10], RIFF (Rotation-Invariant Fast Features) descriptor [11], and FREAK (Fast Retina Keypoint) descriptor [12]. These binary descriptors need lower memory than SIFT-like descriptors, and the similarity of descriptors can be evaluated by computing the Hamming distance, which is very efficient to be computed. But one of their shortcomings is that the shape of pattern cannot be changed. In other words, there are only circular pattern and rectangle pattern, which makes them have no affine invariance. In this case, the high false matching rate between different viewpoints images will appear. Another shortcoming is that

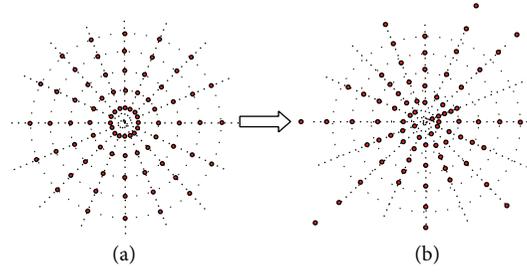


FIGURE 1: (a) is a circular pattern and (b) is an irregular shape one. The red balls are the sampling points.

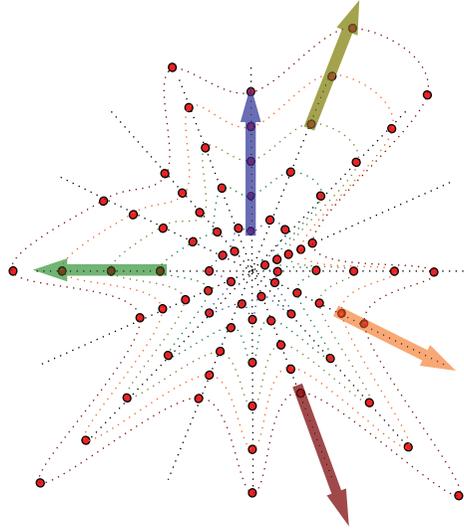


FIGURE 2: Multidirectional rotation invariant builder. The arrows represent the five local orientations. We can see that the pattern is irregular.

one pattern only has one local orientation, and they are not fast enough because of the histogram statistics. In this paper we propose a new binary descriptor with low computational complexity. The orientations of descriptor are found by the five concentric circles, and the affine invariant regions are found by 16 rays which are emanating from the feature point. Experimental results show that the binary affine invariant descriptor has good matching result in the conditions of rotating, image zooming, noising, lighting, and small-scale perspective transformation. In this paper we call this binary descriptor BAND (Binary Affine Invariant Descriptor).

2. Method

We describe 3 key steps of BAND in this section, namely, affine invariant radius builder, multidirectional rotation invariant builder, and simple random sampling pattern builder.

2.1. Affine Invariant Radius Builder. There are many algorithms that extract the local affine invariant regions, such as MSER (Maximally Stable Extremal Regions) [13], EBR (Edge-Based Regions), IBR (Intensity Extrema-Based Regions) [14], and Salient Region [15]. Here we refer to the idea of IBR. With the aim of extracting local affine invariant regions, the 16 rays

which are emanating from the feature point are extracted. The gray value of each sampling point located in the rays is obtained through quadratic interpolation. Here we define the affine invariant radius R_m as

$$R_m = T \cdot \max(R_1, R_2),$$

$$R_1 = L \min \left(\text{abs} \int_c^i (g_i - g_c) dr \right), \quad (1)$$

$$R_2 = L \max \left(\text{abs} \left(\int_c^i (g_i - g_c) dr \right) \right),$$

where R_m is the affine invariant radius in the direction of p , i is the Euclidean arc length along the ray R_m , g_i is the gray value of the radius at the position i , $\text{abs}(x)$ is the absolute value of the x , $L \min(y)$ is the minimum value of the distance between local minimum of the y and the feature point, $L \max(y)$ is the maximum value of the distance between local maximum of the y and the feature point, and T is a constant, and we use a value of 0.76, which is determined by experiments.

By calculating the affine invariant radius, the original circular pattern is deformed into irregular shape pattern. It is shown in Figure 1.

2.2. Multidirectional Rotation Invariant Builder. In order to get the rotation invariant property, existing algorithms need

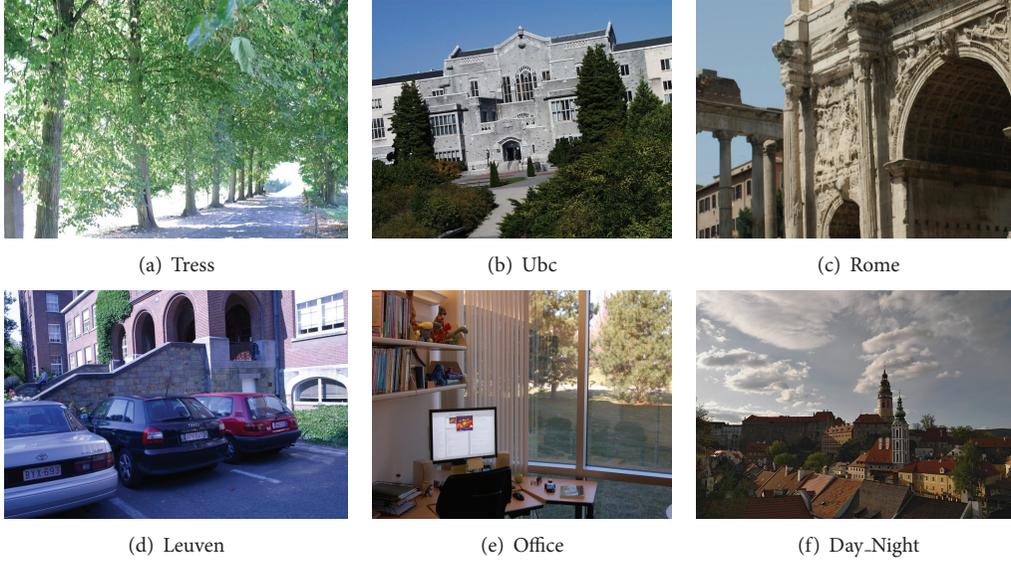


FIGURE 3: Datasets used for evaluation: blur (Trees), JPEG compression (Ubc), rotation (Rome), and brightness changes (Leuven, Office, and Day_Night).

to determine the direction of the local area and rotate the axis. In this process, the defects of these descriptors include

- (1) BRIEF descriptor ignores the rotation; therefore a lot of outliers appear in the case of large-angle rotation.
- (2) The algorithms, like SIFT and ORB, use the rectangle pattern. It is difficult to obtain the local orientation, and they have a high computational complexity.
- (3) The algorithms, like FREAK and BRISK, only consider one local orientation with a pattern, although using the circular pattern.

Based on the above three points and reference of the BRISK and FREAK, the pattern of BAND descriptor is as follows: construct five circles concentric with the feature point, and there are 16 sampling points in each circle homogeneously. Every affine invariant radius of every circle is obtained by the affine invariant radius builder.

Firstly, obtain the locations of sampling points on every circle by affine invariant radius:

$$(x_m, y_m) = \left(x_c + R_M \sin\left(\frac{2\pi p}{P}\right), y_c - R_M \cos\left(\frac{2\pi p}{P}\right) \right), \quad (2)$$

where m is the index of the sampling points surrounding the feature point, M is the number of sampling points, and c is the location of the feature point.

Secondly, according to literature [14], LBP binary string of every circle is

$$\text{LBP}_{v,w} = \begin{cases} 1, & \text{if } g_p > g_c, \\ 0, & \text{else,} \end{cases} \quad (3)$$

where v is the index of the circle, w is the location of one bit of the string, g_p is the gray value at the position p on the circle,

and g_c is the gray value of feature point. Finally, according to the rotation invariant LBP [16], we could obtain five local orientations. The rotation invariant LBP defined by Ojala is $\text{LBP}_{P,R}^{ri}$:

$$\text{LBP}_{P,R}^{ri} = \min\left(\text{ROR}\left(\text{LBP}_{P,R}^{ri}, i\right) \mid i = 0, 1, \dots, P-1\right), \quad (4)$$

where $\text{ROR}(x, i)$ performs a circular bitwise right shift operation on the number x with i times. The pattern of BAND descriptor is shown in Figure 2.

2.3. Simple Random Sampling Pattern Builder. This step is very simple. We get a 256 bits vector by sampling point pairs from the five LBP coding strings randomly and orderly. The bit-vector descriptor is assembled by performing all the comparisons of point pairs (g_a, g_b) , such that each bit L corresponds to

$$L = \begin{cases} 1, & \text{if } g_a > g_b, \\ 0, & \text{else,} \end{cases} \quad (5)$$

$$i = 0, 1, 2, \dots,$$

where g is the gray value of the point sampled from the five LBP strings.

3. Experiments and Analysis

Programming environment is Matlab 2013, Visual Studio 2010, and OpenCV 2.4.4. The datasets come from [17]. All algorithms are running on an Intel(R) Core (TM) i3 of 3.40 GHz. In order to unify the results, all of the results do not remove the mismatches and only show the best results by setting a threshold.

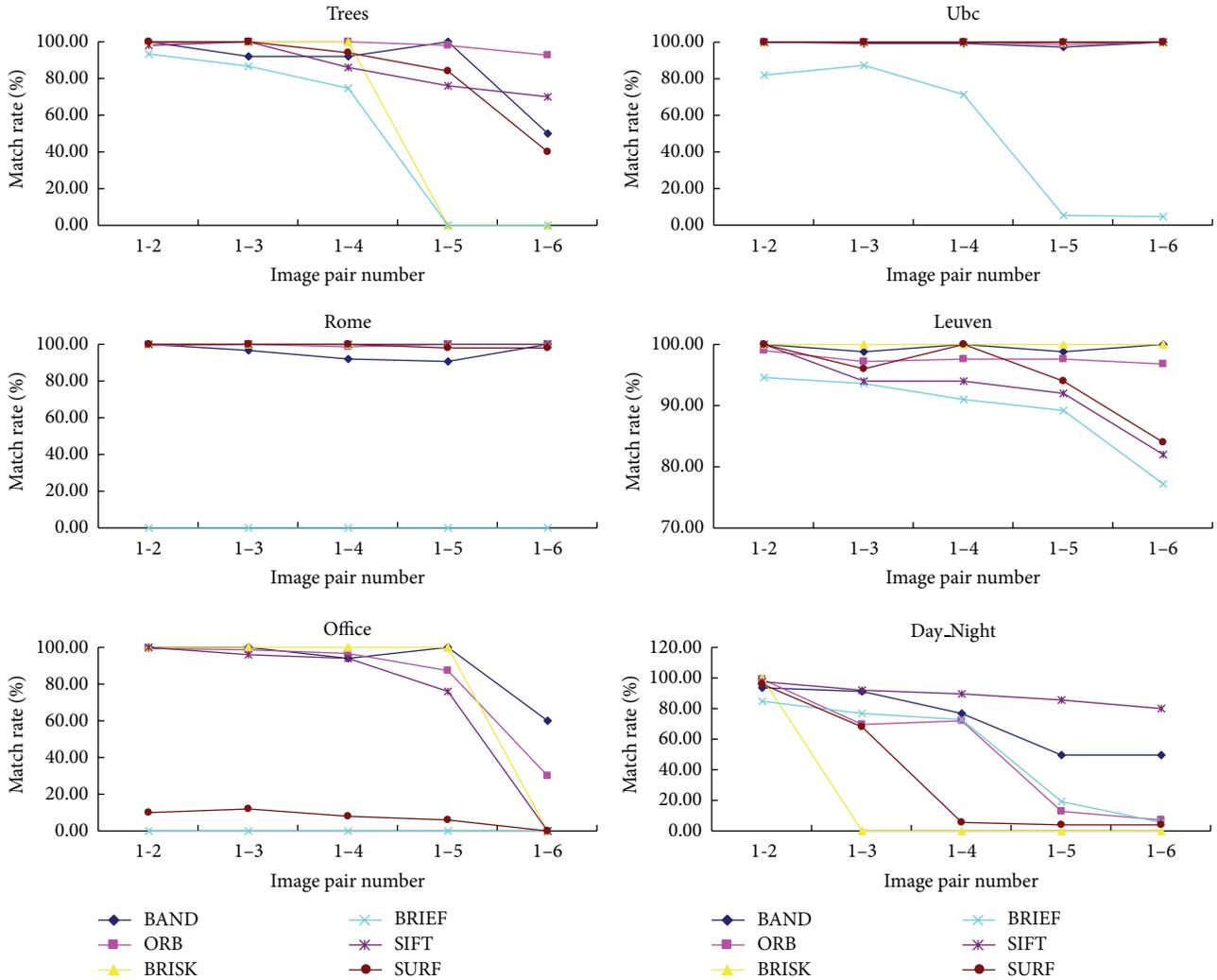


FIGURE 4: Evaluation result shows match rate for BAND, ORB, BRISK, BRIEF, SIFT, and SURF.

3.1. *Single Performance Verification of Local Affine Invariant Descriptors.* Each of the datasets contains a sequence of six images exhibiting an increasing amount of transformation. All comparisons here are performed against the first image in each dataset. Figure 3 shows one image for each dataset analyzed.

The transformations cover blur (Trees), brightness changes (Leuven, Office, and Day_Night), JPEG compression (Ubc), and rotation (Rome). Match rate is defined as a ratio between the number of the correctly matching points and the total number of the matched points. In order to prove the easy integration of our algorithm, we use FAST [18, 19] for detecting the feature points in the sequence of the Trees and the SURF in the others. To make results be comparable, the threshold of Hamming distance for BAND, BRIEF, BRISK, and ORB is 50. We only show the best 50 matched pairs in SURF and SIFT. All of the results are shown in Figure 4.

3.2. *Images Matching for Outdoor Scenes.* We only test single performance of BAND descriptor in Section 3.1. In this

section, we would use 6-pair images showed in Figure 5 to test its comprehensive performance. As it is shown in Figure 5, all of the 6 pairs contain view point changes in different degrees, such as in scale, brightness, and rotation. Besides, all of the images contain many structures which appear repeatedly. For example, the structures in image Church and the image Fountain both are symmetrical, and the buildings in image Brussels, Venice, Semper and Rathaus have similar doors, windows and statues individually. These conditions add extra difficulty to matching.

Figure 6 shows the matching results of the BAND. The matched point numbers and match rates of different descriptors are shown in Figures 7 and 8. To make results be comparable as in Section 3.1, the threshold of Hamming distance about BAND, BRIEF, BRISK and ORB is 50 invariably. We only show the best 50 matched pairs in SURF and SIFT as before.

We use SURF for detecting the feature points, and the threshold is equal to before. The 6-pair outdoor scenes are as shown in Figures 5 and 6.

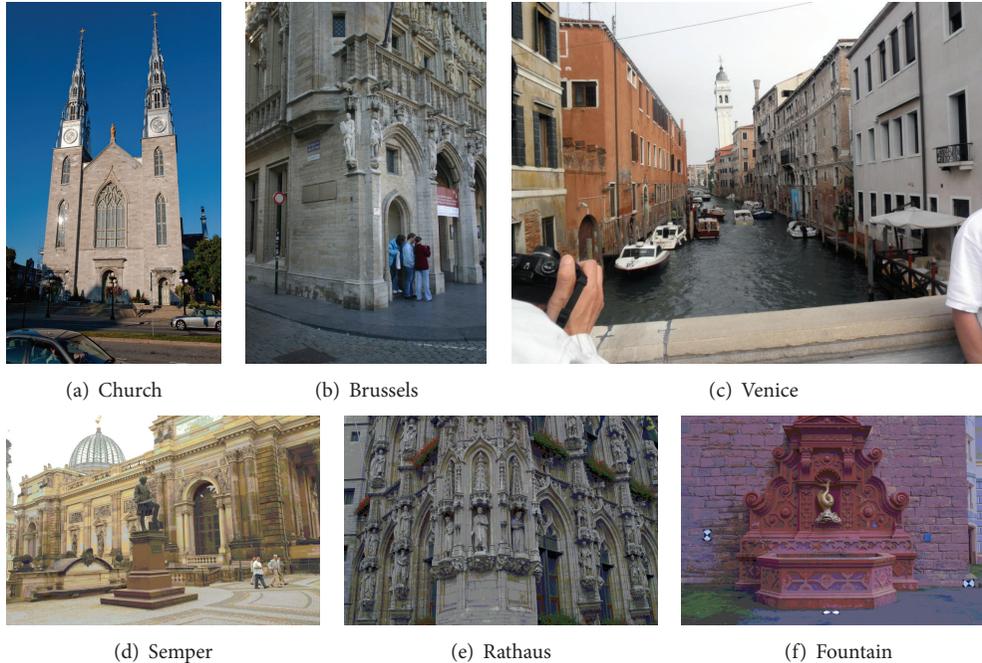


FIGURE 5: The 6-pair outdoor scenes are all the buildings and artifacts.

3.3. Time Consumption Comparison. In Matlab 2013 platform, we compare the time consumption between SIFT and BAND descriptor per point pair. Test result is shown in Table 1.

If we set the time of BAND describing one feature point to be 1, then we would get the time consumption of other descriptors in Table 2 clearly.

3.4. Analysis of Experimental Results. In the single performance verification, BAND is demonstrated as it has a better adaptability than the other descriptors. In 6 experiments, the degree of transformations is increasing. It is noteworthy that the Leuven shows a transformation in brightness from normal to dark gradually. Similarly, the Office shows a transformation in brightness from dark to normal. Differently, the Day_Night shows a transformation in brightness from normal to dark, but some local areas become lighter than surroundings because of some lights, such as lamps, bulbs, and candles. In other words, it changes nonlinearly. Since the similarity is reduced, the direction of every line in Figure 4 shows downward. BRIEF shows the worst performance among all of the descriptors, because it does not note the variety of rotation and scale. In the light experiments, such as the Leuven, the Office, and the Day_Night, BAND demonstrates outstanding performance. BAND can maintain a high rate of correctly matched no matter how seriously changes of lightness. In other experimental conditions, performances of BAND are similar to the mainstream descriptors. In addition, no matter what feature points we use, like FAST and SURF, BAND descriptors can work well, which shows a high adaptability.

In the experiment of matching the outdoor scenes, we can get a lot of matched points by ORB descriptor, but its

TABLE 1: The time consumption comparison between SIFT and BAND descriptor per point pair.

	SIFT	BAND
Description time (ms)	531.1	0.5463
Matching time (ms)	0.093	0.0015

TABLE 2: It shows time consumption comparison of different descriptors.

	Time consumption					
	BAND	ORB	BRISK	BRIEF	SIFT	SURF
Feature description	1	8.5	5.7	8.3	1000	65

correctly matched rate is lower than BAND descriptor. One of its negative effects is that there will be more computation in the process of removing the mismatched points. Although the amount of the matched points we get by BAND is not as much as ORB, the amount is kept at a high level and the correctly matched rate of it is satisfactory. It is better than or as good as the correctly matched rate of other descriptors. The reason is that BAND considers the affine invariance and its pattern, whose shape is variable, can better adapt to the changes of viewpoint.

In the experiment of the time-consuming comparison, description time per point of BAND is as short as 0.1% of the time of SIFT approximately. The reason is that there is no process of establishing the Gauss Pyramids or too much computation of fitting operation. BAND computes the distance between two descriptions by using Hamming distance, but SIFT uses Euclidean distance. Obviously, calculating the Hamming distance has higher efficiency and lower

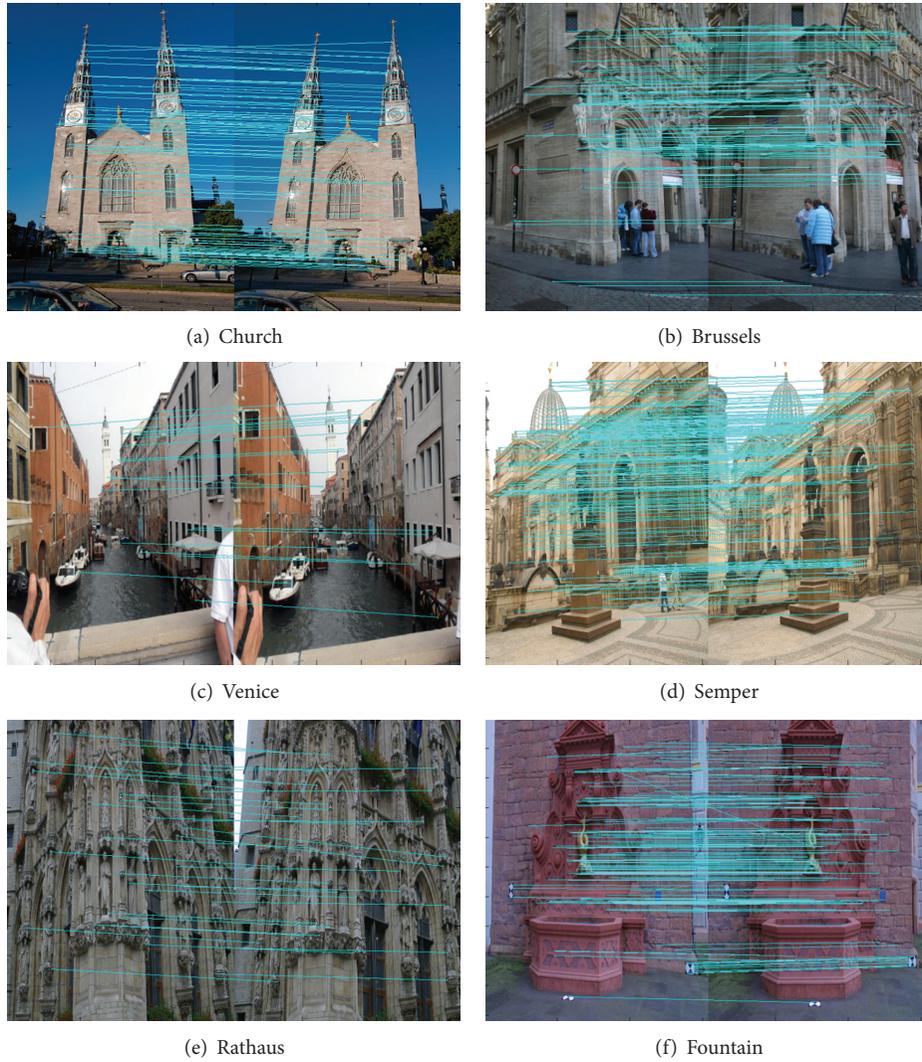


FIGURE 6: The results show the matched points, and there is no process of removing the mismatched.

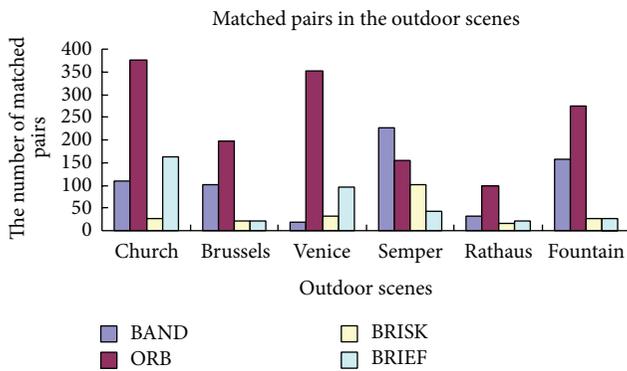


FIGURE 7: This figure shows the numbers of matched point pairs from different binary descriptors.

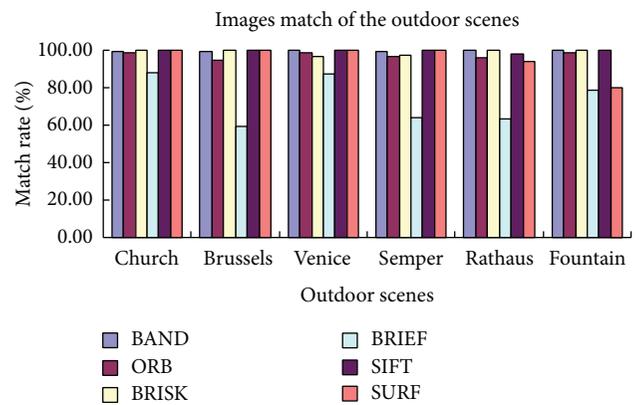


FIGURE 8: This figure shows the match rates of different binary descriptors in the 6 outdoor scenes.

computational complexity. So, together, BAND shows higher efficiency. And the comparison of the time-consuming with other descriptors also shows that BAND is faster. Therefore,

BAND has advantage in the situation that demands high computing speed.

4. Conclusion

Experimental results show that BAND descriptor has significant advantages, such as low computational complexity, well adaptability, and good stability. It makes up for the disadvantages of other descriptors that have high computational complexity and have no affine invariance. BAND has a good matching result in the conditions of rotating, image zooming, noising, lighting, and small-scale perspective transformation. More specifically, it has a moderate number of matched points and a high correctly matched rate. Because of these, on the base of guaranteeing accuracy, BAND could improve the calculation efficiency and meet real-time requirement.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 2, pp. 1150–1157, IEEE, Kerkyra, Greece, September 1999.
- [2] D. G. Lowe, "Local feature view clustering for 3D object recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, vol. 1, pp. I-682–I-688, IEEE, Kauai, Hawaii, USA, December 2001.
- [3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [4] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [5] Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 2, pp. II-506–II-513, IEEE, Washington, DC, USA, July 2004.
- [6] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [7] J. M. Morel and G. Yu, "ASIFT: a new framework for fully affine invariant image comparison," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 438–469, 2009.
- [8] M. Calonder, V. Lepetit, C. Strecha et al., "BRIEF: binary robust independent elementary features," in *Computer Vision—ECCV 2010*, vol. 6314 of *Lecture Notes in Computer Science*, pp. 778–792, Springer, Berlin, Germany, 2010.
- [9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 2564–2571, IEEE, Barcelona, Spain, November 2011.
- [10] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: binary robust invariant scalable keypoints," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 2548–2555, IEEE, Barcelona, Spain, November 2011.
- [11] G. Takacs, V. Chandrasekhar, S. Tsai et al., "Rotation invariant fast features for large-scale recognition," in *Applications of Digital Image Processing XXXV*, vol. 8499 of *Proceedings of SPIE*, San Diego, Calif, USA, October 2012.
- [12] A. Alahi, R. Ortiz, and P. Vandergheynst, "Freak: fast retina keypoint," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 510–517, IEEE, Providence, RI, USA, June 2012.
- [13] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761–767, 2004.
- [14] T. Tuytelaars and L. van Gool, "Matching widely separated views based on affine invariant regions," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 61–85, 2004.
- [15] T. Kadir and M. Brady, "Saliency, scale and image description," *International Journal of Computer Vision*, vol. 45, no. 2, pp. 83–105, 2001.
- [16] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [17] C. Strecha, W. von Hansen, L. van Gool, P. Fua, and U. Thoennessen, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.
- [18] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Computer Vision—ECCV 2006*, vol. 3951 of *Lecture Notes in Computer Science*, pp. 430–443, Springer, Berlin, Germany, 2006.
- [19] E. Rosten, R. Porter, and T. Drummond, "Faster and better: a machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

Research Article

Comparing the State-of-the-Art Efficient Stated Choice Designs Based on Empirical Analysis

Li Tang,¹ Xia Luo,¹ Yang Cheng,² Fei Yang,¹ and Bin Ran²

¹ School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

² Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA

Correspondence should be addressed to Li Tang; tanglitraffic@gmail.com

Received 16 November 2013; Accepted 12 January 2014; Published 20 February 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Li Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The stated choice (SC) experiment has been generally regarded as an effective method for behavior analysis. Among all the SC experimental design methods, the orthogonal design has been most widely used since it is easy to understand and construct. However, in recent years, a stream of research has put emphasis on the so-called efficient experimental designs rather than keeping the orthogonality of the experiment, as the former is capable of producing more efficient data in the sense that more reliable parameter estimates can be achieved with an equal or lower sample size. This paper provides two state-of-the-art methods called optimal orthogonal choice (OOC) and *D*-efficient design. More statistically efficient data is expected to be obtained by either maximizing attribute level differences, or minimizing the *D*-error, a statistic corresponding to the asymptotic variance-covariance (AVC) matrix of the discrete choice model, when using these two methods, respectively. Since comparison and validation in the field of these methods are rarely seen, an empirical study is presented. *D*-error is chosen as the measure of efficiency. The result shows that both OOC and *D*-efficient design are more efficient. At last, strength and weakness of orthogonal, OOC, and *D*-efficient design are summarized.

1. Introduction

Abundant and accurate data is the foundation of study. To date, emerging technologies are largely introduced in data mining and processing [1]. In [2, 3], how to obtain high-quality data with the high-tech in the field of Intelligent Transportation System is deeply discussed. However, the importance of refining data collection technique has not raised researchers' attention in behavior analysis area until last decades, such as travel mode choice and safety in [4, 5]. The purpose of conducting choice experiments is to collect data that can be used to estimate the independent influence of attributes on observed choices. There are two paradigms of choice data: revealed preference (RP) and stated choice (SC) data. Typically, in RP surveys, respondents are asked to recall information about his/her last choice, including alternatives and attribute levels available in real market. Differently, SC experiments present sampled respondents with a number of different hypothetical choice situations, each consisting of a universal but finite set of alternatives defined on a

number of attribute dimensions. Thus, SC data collected on 300 respondents, each of whom is asked to make 8 choices produces a total of 2400 choice observations while RP experiment on the same sample size of respondents collects only 300 observed choices.

SC experiment has been widely used because it can observe choices on alternatives which do not exist in the current market. So analysts are able to predict, for example, the share rate of a newly introduced transportation mode. Another reason of its popularity lies in the ability to provide variability in attributes in a relatively small sample size compared with RP experiment, with which better estimation of influence of each attribute on choice can be achieved. Usually, respondents in a SC experiment will be faced with some "selected" choice situations, considering that making choices among all the possible combinations of attribute levels is too many to accomplish for a single respondent. Thus, how analysts distribute the levels of the design attributes in an experiment plays a big role. It may impact upon not only whether or not an independent assessment of contribution

of each attribute to the choices observed can be determined, but also the ability of the experiment to detect statistical relationships that may exist within the data.

Historically, researchers have relied on orthogonal experimental designs, in which the attributes of the experiment are statistically independent by forcing them to be orthogonal [6]. As such, orthogonal designs theoretically allow for an independent determination of each attribute's influence upon the observed choices. To generate an orthogonal design, usually but not necessarily, the first step is to generate a full factorial design, a design which contains all possible attribute level combinations. Mathematically, a full factorial design will produce $\prod_{k=1}^k L_k$ choice situations, where L_k is the number of levels assigned to attribute k (e.g., a design with 4 attributes, two with 2 attribute levels, one with 3 levels, and one with 4 levels will produce a full factorial design with $2 \times 2 \times 3 \times 4 = 48$ choice situations). The second step is to take a subset of choice situations from the full factorial design, which is known as fractional factorial designs. Randomized, cyclical, Bayesian, and fold over procedures are the common approaches used to generate fractional factorial design in [7–10]. It is noteworthy that, in either a full factorial or a fractional factorial design, orthogonality is kept between two random attributes. The only difference between them is that orthogonality will not be kept in terms of interaction effects (i.e., the influence of two or more attribute columns multiplied together) in fractional factorial design. Here are some simple rules to check whether a fractional factorial design is orthogonal or not:

- (i) Every level of every attribute appears at same times.
- (ii) All possible attribute level combinations of random two attributes appear at same times.

While orthogonal design has long been used in practice, in [11, 12], there is a stream of researchers in recent years doubted the importance of orthogonality in SC data when it is used to estimate discrete choice model, not to mention whether orthogonality can be kept in reality. Orthogonality is important in linear models since it avoids multicollinearity problem and also minimized variance-covariance matrix of the estimated model, in which way standard errors of parameter estimates are also minimized. Unfortunately, discrete choice model is nonlinear; thus the derivation of its parameters' variance-covariance matrix is very different from the way in linear models. Seeing from that, keeping orthogonality of the parameters has little to do with minimizing their standard errors.

Acknowledgment of this fact has led researchers to transfer their efforts to obtain experimental designs that minimize the asymptotic variance-covariance (AVC) matrix of discrete choice models and provide more reliable parameter estimates with an equal or lower sample size. Such designs are called efficient designs. To date, most research has been focused on developing methods to generate efficient designs; comparison and validation of these methods in practice are rarely seen. This paper provides two state-of-the-art methods called optimal orthogonal choice (OOC) and D -efficient design. Their performance on both theoretical efficiency and

TABLE 1: Optimal orthogonal choice design for 2 alternatives with 3 binary attributes.

Choice situation	Alternative 1			Alternative 2		
	X_1	X_2	X_3	X_1	X_2	X_3
1	0	0	0	1	1	1
2	0	1	1	1	0	0
3	1	0	1	0	1	0
4	1	1	0	0	0	1

practical use are compared with the conventional orthogonal design.

The remainder of this paper is organized as follows. In Section 2, OOC and D -efficient design are introduced. In Section 3, SC experiment design using orthogonal, OOC, and D -efficient methods separately is generated. The result in terms of D -errors are presented in Section 4, along with the summarized strength and weakness.

2. Efficient Experimental Design Methods

2.1. Optimal Orthogonal Choice Design. Considering the popularity and convenience of orthogonal design for analysts in practice, there is a stream of researchers in [13, 14] who kept on exploring improved SC design method maintaining orthogonality, which is called optimal orthogonal choice design. The essential idea of OOC design is to maximize the differences of attribute levels across alternatives, so that the parameters can be estimated in the largest extent of variety of attribute levels as well as independently. The basic process of generating an OOC design is as follows.

Step 1. Generate a fractional factorial orthogonal design for alternative 1. N represents the number of choice situations of the design.

Step 2. Choose some systematic changes to get the allocation of attribute levels in alternative 2 from alternative 1. Systematic changes are certain rules to decide how the attribute levels change from alternative 1 and will be discussed in later context.

Step 3. Choose another systematic change to get the allocation of attribute levels in alternative 3 from alternative 1.

Step 4. Keep doing this until all the alternatives are determined.

It will be much easier to understand this method by starting with a binary attribute level design. Again, we assume that L_k is the number of levels assigned to generic attribute k for alternative j , represented by $0, 1, \dots, L_k - 1$. In a design for 2 alternatives and 3 attributes each with 2 levels, an orthogonal design in 4 choice situations for alternative 1 can be firstly generated. Then 0's and 1's in alternative 1 are interchanged in alternative 2. Thus the attribute levels of each attribute are forced to be different across alternatives. The result is shown in Table 1.

TABLE 2: Optimal orthogonal choice design for 3 alternatives with 3 binary attributes.

Choice situation	Alternative 1			Alternative 2			Alternative 3		
	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
1	0	0	0	1	1	0	0	0	1
2	0	1	1	1	0	1	0	1	0
3	1	0	1	0	1	1	1	0	0
4	1	1	0	0	0	0	1	1	1

To generate OOC design for more alternatives, it is necessary to introduce S_k to represent the largest number of different pairs appeared between alternatives for a specific attribute. The equation of S_k is shown as follows, where J stands for the number of alternatives in choice set:

$$S_k = \begin{cases} \frac{(J^2 - 1)}{4}, & l_k = 2, J \text{ odd,} \\ \frac{J^2}{4}, & l_k = 2, J \text{ even,} \\ \frac{(J^2 - (l_k x^2 + 2xy + y))}{2}, & 2 < l_k \leq J, \\ \frac{J(J-1)}{2}, & l_k \geq J. \end{cases} \quad (1)$$

In an example of OOC design for 3 alternatives and 3 attributes each with 2 levels shown in Table 1, we can get $S_k = (J^2 - 1)/4 = 2$ for all the attributes. For instance, in the first choice situation (000, 110, and 001), the attribute levels differ twice for each attribute (i.e., for attribute X_1 , the levels are 010, creating 3 pairs (01, 10, and 00) in which two of them (01, 10) are different).

We can see from Table 1 that the distribution of attribute levels in alternative 2 is obtained by interchanging 0's and 1's in X_1 and X_2 in alternative 1 and that in alternative 3 is obtained by interchanging 0's and 1's in X_3 in alternative 1. These systematic changes can be also described as adding a generator in alternative 1 to get alternative 2 and adding another generator to get alternative 3. The addition is performed in modulo arithmetic according to the number of levels for a specific attribute. Here in the example, $L_k = 2$ for all attributes, thus when a generator 110 is added to the choice situations in alternative 1 in modulo 2 arithmetic like this: $000 + 110 \equiv 110$, $011 + 110 \equiv 101$, and so on, alternative 2 is obtained. Alternative 3 is generated in the same way by adding a generator 001 to alternative 1. Notice that the generators added to alternatives must have a value of $S_k = 2$ (i.e., generators (000, 110, and 001) used in Table 2 can meet the requirement while another generators (000, 100, and 010) cannot not).

Designs for any choice set size with any number of attributes each having any number of levels can be generated in similar way. However, a big limitation of OOC design is that it can only generate designs for generic attributes according to the principle of this method. How alternative-specific attributes distribute across alternatives is rarely discussed

in the literature. Another shortcoming of OOC design is that it may produce a lot of unreasonable combinations by forcing maximum level differences in attributes across alternatives. Answers from responders who have to make decision in such choice situations may not reflect their actual choice statements since the "proper alternative" may not be contained in the questionnaire.

2.2. D-Efficient Design. While people who raise OOC method keep on improving design with remaining orthogonality, another stream of researchers goes straight forward to increasing the statistical efficiency of the design by minimizing the elements of the asymptotic variance-covariance (AVC) matrix of discrete choice models. The AVC matrix can be obtained by taking the negative inverse of the expected second derivatives of the log-likelihood function of the model proved in [15]. To interpret the process of calculating the AVC matrix and the measure of efficiency of a design, here we briefly introduce the most well-known multinomial logit model.

Assume an individual faced with alternative $j = 1, 2, \dots, J$ in choice situation $n = 1, 2, \dots, N$. The utility of an individual for alternative j in choice situation n can be expressed as

$$U_{jn} = V_{jn} + \varepsilon_{jn}, \quad (2)$$

where V_{jn} represents observed part of utility for each alternative j in choice situation n . It is assumed to be a linear additive function of several attributes with corresponding weights. These weights are unknown parameters to be estimated and can be divided into two categories: generic parameters and alternative-specific parameters. The generic parameters and alternative-specific parameters can be denoted by β_k^* , $k = 1, \dots, K^*$, and β_{jk} , $k = 1, \dots, K_j$, respectively, with their associated attribute levels x_{jkn}^* and x_{jkn} for each choice situation n . Thus, the total number of parameters to be estimated equal to $\bar{K} = K^* + \sum_{j=1}^J K_j$. V_{jn} can be expressed as

$$V_{jn} = \sum_{k=1}^{K^*} \beta_k^* x_{jkn}^* + \sum_{k=1}^{K_j} \beta_{jk} x_{jkn}, \quad \forall j = 1, \dots, J, \quad \forall n = 1, \dots, N, \quad (3)$$

where ε_{jn} is the unobserved component, which is independently and identically extreme value type one distributed. The probability P_{jn} that an individual chooses alternative j in choice situation n becomes

$$P_{jn} = \frac{\exp(V_{jn})}{\sum_{j=1}^J \exp(V_{jn})}, \quad \forall j = 1, \dots, J, \quad \forall n = 1, \dots, N. \quad (4)$$

Considering that the most popular way to estimate parameters is maximum likelihood estimation, the log-likelihood function of parameters for a single respondent can be expressed as

$$\begin{aligned} L(\beta^*, \beta) &= \sum_{n=1}^N \sum_{j=1}^J y_{jn} \log P_{jn} \\ &= \sum_{n=1}^N \left[\sum_{j=1}^J y_{jn} \left(\sum_{k=1}^{K^*} \beta_k^* x_{jkn}^* + \sum_{k=1}^{K_j} \beta_{jk} x_{jkn} \right) \right. \\ &\quad \left. - \log \left(\sum_{i=1}^J \exp \left(\sum_{k=1}^{K^*} \beta_k^* x_{ikn}^* + \sum_{k=1}^{K_j} \beta_{jk} x_{ikn} \right) \right) \right], \end{aligned} \quad (5)$$

where y represents the binary outcome of all choice situations. While alternative j is chosen in choice situation n , y_{jn} equals one; otherwise it is zero. Then the AVC matrix can be expressed as the second derivative of the log-likelihood function as follows:

$$\frac{\partial^2 L(\beta^*, \beta)}{\partial \beta_{k_1}^* \partial \beta_{k_2}^*} = - \sum_{n=1}^N \sum_{j=1}^J x_{j k_1 n}^* P_{jn} \left(x_{j k_2 n}^* - \sum_{i=1}^J x_{i k_2 n}^* P_{in} \right), \quad (6)$$

$\forall k_1, k_2 = 1, \dots, K^*$,

$$\frac{\partial^2 L(\beta^*, \beta)}{\partial \beta_{j_1 k_1} \partial \beta_{j_2 k_2}^*} = - \sum_{n=1}^N x_{j_1 k_1 n} P_{j_1 n} \left(x_{j_2 k_2 n}^* - \sum_{i=1}^J x_{i k_2 n}^* P_{in} \right), \quad (7)$$

$\forall j_1 = 1, \dots, J, \quad k_1 = 1, \dots, K_{j_1}, \quad k_2 = 1, \dots, K^*$,

$$\frac{\partial^2 L(\beta^*, \beta)}{\partial \beta_{j_1 k_1} \partial \beta_{j_2 k_2}} = \begin{cases} \sum_{n=1}^N x_{j_1 k_1 n} x_{j_2 k_2 n} P_{j_1 n} P_{j_2 n}, & \text{if } j_1 \neq j_2; \\ - \sum_{n=1}^N x_{j_1 k_1 n} x_{j_2 k_2 n} P_{j_1 n} (1 - P_{j_2 n}), & \text{if } j_1 = j_2. \end{cases}$$

$\forall j_1 = 1, \dots, J, \quad k_i = 1, \dots, K_{j_i}. \quad (8)$

Equations (6)–(8) represent functions that allow generic and alternative-specific parameters. In the case where only generic parameters exist, only (6) remains, and when there are only alternative-specific parameters, (8) remains. In addition, if there are M identical respondents, these second derivatives are multiplied by M .

Let $(\bar{\beta}^*, \bar{\beta})$ denote the true values of the parameters. The Fisher information matrix I is defined as the expected values of the second derivative of the log-likelihood function:

$$I(\bar{\beta}^*, \bar{\beta}) = M \cdot \frac{\partial^2 L(\bar{\beta}^*, \bar{\beta})}{\partial \beta \partial \beta'}. \quad (9)$$

Hence, the AVC matrix can be expressed as a $\bar{K} \times \bar{K}$ matrix that is equal to the negative inverse of the Fisher information matrix [16]:

$$\Omega = - \left[I(\bar{\beta}^*, \bar{\beta}) \right]^{-1} = - \frac{1}{M} \left[\frac{\partial^2 L(\bar{\beta}^*, \bar{\beta})}{\partial \beta \partial \beta'} \right]^{-1}. \quad (10)$$

Rather than working with each element within the AVC matrix directly, a preferred measure within the literature is D -error, calculated by taking the determinant of the AVC matrix and scaling this value by the number of parameters \bar{K} . It is common to assume that a single respondent (i.e., $M = 1$) represents all respondents, an assumption consistent with the multinomial logit (MNL) model form [8, 17]. Designs that aim at minimizing D -error are called D -efficient designs.

Since the calculation of D -error involves the values of parameters, approaches to determine D -error have been improved in recent years. In our empirical study later, we use D_p -error as the statistic to measure the efficiency of experimental designs. To calculate D_p -error, nonzero priors are needed. It can be expressed as follows:

$$D_p\text{-error} = - \left[\det \frac{\partial^2 L(\bar{\beta}^*, \bar{\beta})}{\partial \beta \partial \beta'} \right]^{-1/\bar{K}}. \quad (11)$$

According to the D -efficient method, we can figure out that, for a given sample size, attempts to minimize D -error statistic will directly lead to the minimization of the AVC matrix. Meanwhile, by taking the square root of the diagonal elements (including M) of the AVC matrix, the minimization of asymptotic standard errors is achieved. Thus the asymptotic standard error of the parameter estimates will diminish in terms of statistical significance from each additional respondent added to a survey.

Figure 1(a) reveals that the standard error decreases by increasing sample size at the beginning of a given design X^I . Exceeding a certain limit, enlarging sample size has little impact on demising standard error. On the other side, Figure 1(b) shows that investing a more efficient design X^{II} can lead to larger decreases in standard error. In other words, a smaller sample size may be satisfied at a certain level of standard error when using efficient designs rather than common ones.

3. Empirical Study

In this section, three experimental designs which used different methods (orthogonal, OOC, and D -efficient) are generated. They will be used to obtain trip mode choice data on a corridor connecting two large business districts in Chengdu. Shawan conference and exhibition center and Jinsha station are chosen as the origin and destination for the survey. Three alternatives are involved: car, taxi, and bus. A typical multinomial logit model in transportation will be formulated and serve as the basis of most of the analyses in

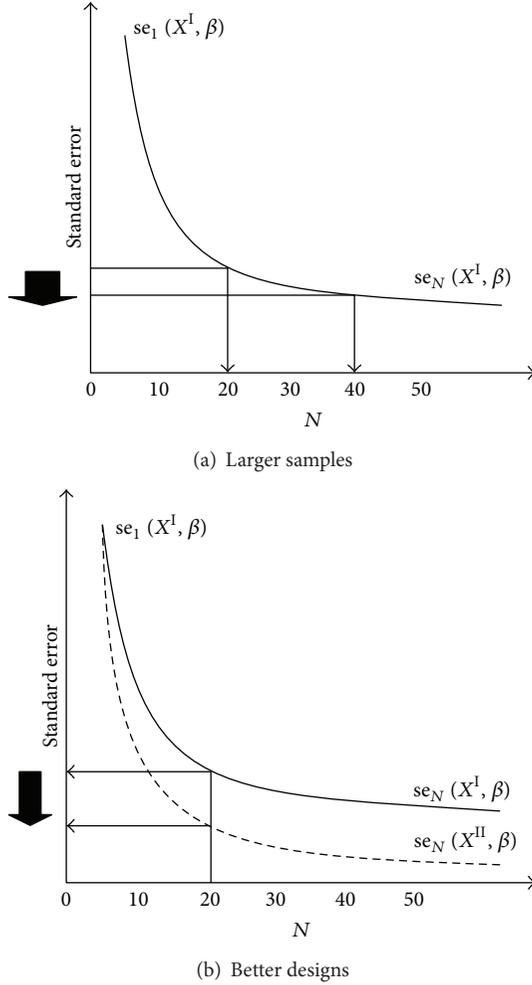


FIGURE 1: Comparison of investing in larger sample sizes versus more efficient designs.

the subsequent section. The observed part of utility of every alternative is expressed as follows:

$$\begin{aligned}
 V^{\text{car}} &= \beta_0^{\text{car}} + \beta_1 \text{TT}^{\text{car}} + \beta_2 \text{TC}^{\text{car}}, \\
 V^{\text{taxi}} &= \beta_0^{\text{taxi}} + \beta_1 \text{TT}^{\text{taxi}} + \beta_2 \text{TC}^{\text{taxi}}, \\
 V^{\text{bus}} &= \beta_1 \text{TT}^{\text{bus}} + \beta_2 \text{TC}^{\text{bus}},
 \end{aligned} \quad (12)$$

where TT represents travel time. TC represents travel cost. For car users, TC equals the fuel cost. For taxi users, TC equals the money paid for the trip. For bus users, TC equals the ticket price. Seeing from (12), parameters for TT and TC are generic across three alternatives. Thus, four parameters are going to be estimated in total (two of them are alternative-specific constant, which has nothing to do with any attribute). The attribute levels and prior information about parameters are given in Table 3 based on previous study results as well as to preserve realistic estimates for the private and public transport alternatives.

In order to obtain better estimation of parameters, three levels are set for each attribute for maximum variation as

TABLE 3: Prior parameter values and attribute levels for case study.

(a)					
Prior parameter values					
β_0^{car}	β_0^{taxi}	β_1	β_2		
1	-0.6	-0.45	-0.9		
(b)					
Attribute level					
Car		Taxi		Bus	
TT ^{car}	TC ^{car}	TT ^{taxi}	TC ^{taxi}	TT ^{bus}	TC ^{bus}
(min)	(RMB)	(min)	(RMB)	(min)	(RMB)
15	4	15	16	25	1
20	5	20	18	30	2
30	7	30	22	38	3

much as possible. The values of TT^{car} , TT^{taxi} , and TT^{bus} are measured in free, normal, and congested traffic flow. The values of TC^{car} are calculated as the kilometers between the origin and destination by the consumed oil price under 7.3 RMB/liter, 7.7 RMB/liter, and 8 RMB/liter. The values of TC^{taxi} are measured in free, normal, and congested traffic flow. The value of TC^{bus} is based on the current price and plus/minus 1 RMB. The number of choice situations (i.e., 18) is selected such that attribute level balance can be achieved. Obviously, this number is too large for a single respondent. Thus, we introduce a block variable to divide the design into smaller parts (i.e., here we block the design into three parts so that six choice situations are provided to a single respondent). Each block is not orthogonal by itself, but in combination with other blocks. Attribute level balance is maintained as much as possible in each block.

We generate three different (attribute level balanced) designs with 18 choice situations assuming the above MNL model, using the software Ngene 1.1.1. The design results are shown in Table 4 as well as D -error value for each design.

4. Results and Discussion

As expected, the two efficient designs produce lower D -error value (0.118313 and 0.114612 for the OOC design and D -efficient design, resp.), while orthogonal design produces higher D -error value (0.194724), seeing from Table 4. The D -error of the orthogonal design is 1.64 times greater than the D -error value of the OOC design and 1.69 times greater than D -efficient design. This suggests that, on average, the asymptotic standard errors of the parameter estimates using the orthogonal design will be 1.28 to 1.31 times larger than the efficient designs. Clearly, the efficient designs are able to provide more reliable parameter estimates than orthogonal design.

On the other hand, comparing the two efficient designs, D -efficient outperforms OOC in terms of the D -error value (0.114612 versus 0.118313). Furthermore, since the OOC method can only generate designs with generic attributes, the use of it is largely limited. In a word, with high statistical

TABLE 4: Experimental designs for empirical study.

Choice situation	Car		Taxi		Bus		Block
	TT ^{car} (min)	TC ^{car} (RMB)	TT ^{taxi} (min)	TC ^{taxi} (RMB)	TT ^{bus} (min)	TC ^{bus} (RMB)	
Orthogonal design for MNL model (D -error = 0.194724)							
1	15	4	15	16	25	1	1
2	20	5	20	18	30	2	1
3	30	7	30	22	38	3	1
4	15	5	20	22	38	1	1
5	20	7	30	16	25	2	1
6	30	4	15	18	30	3	1
7	30	7	20	18	25	1	2
8	15	4	30	22	30	2	2
9	20	5	15	16	38	3	2
10	20	7	15	22	30	1	2
11	30	4	20	16	38	2	2
12	15	5	30	18	25	3	2
13	30	5	30	16	30	1	3
14	15	7	15	18	38	2	3
15	20	4	20	22	25	3	3
16	20	4	30	18	38	1	3
17	30	5	15	22	25	2	3
18	15	7	20	16	30	3	3
Orthogonal optimal design for MNL model (D -error = 0.118313)							
1	15	4	20	18	38	3	1
2	30	7	15	16	30	2	1
3	15	5	20	22	38	1	1
4	20	4	30	18	25	3	1
5	30	5	15	22	30	1	1
6	20	7	30	16	25	2	1
7	20	7	30	16	25	2	2
8	30	5	15	22	30	1	2
9	15	4	20	18	38	3	2
10	20	4	30	18	25	3	2
11	15	5	20	22	38	1	2
12	30	7	15	16	30	2	2
13	15	7	20	16	38	2	3
14	20	5	30	22	25	1	3
15	30	4	15	18	30	3	3
16	30	4	15	18	30	3	3
17	20	5	30	22	25	1	3
18	15	7	20	16	38	2	3
D -efficient design for MNL model (D -error = 0.114612)							
1	30	5	20	18	25	1	1
2	20	7	15	16	38	2	1
3	30	4	15	22	25	3	1
4	15	5	30	16	30	3	1
5	15	5	30	16	38	2	1
6	30	5	20	22	25	1	1
7	20	4	15	22	30	2	2
8	15	7	30	22	38	1	2
9	15	7	20	22	38	1	2
10	30	5	15	16	30	3	2

TABLE 4: Continued.

Choice situation	Car		Taxi		Bus		Block
	TT ^{car} (min)	TC ^{car} (RMB)	TT ^{taxi} (min)	TC ^{taxi} (RMB)	TT ^{bus} (min)	TC ^{bus} (RMB)	
11	15	5	30	18	38	2	2
12	20	4	20	18	25	3	2
13	20	7	30	16	25	1	3
14	20	4	20	18	30	2	3
15	30	7	15	16	30	3	3
16	15	7	30	18	38	1	3
17	20	4	20	18	30	3	3
18	30	4	15	22	25	2	3

TABLE 5: Comparison of orthogonal, OOC, and *D*-efficient design.

Method	Advantage	Disadvantage
Orthogonal	(i) It is the most widely used method and easy to construct or obtain (ii) There are no correlations between attribute levels; thus it allows for an independent estimation of the influence of each attribute on choice	(i) There are too many choice situations/questions for a single respondent (ii) Orthogonally it is hard to maintain in actual design: subsets replicated unevenly, introducing sociodemographic variable and allocation bias of the implausible choice situation (iii) It may contain “useless” choice situations
Optimal orthogonal choice	(i) Attribute level differences are maximized (ii) Choice situations will be reduced as well as attaining the design’s orthogonality	(i) It can only generate designs for generic attributes; the rules for setting up alternative-specific attributes are not clear right now (ii) Unreasonable combinations of attribute levels may appear; thus the “real choice” of respondents is hard to capture
<i>D</i> -efficient	(i) The smaller the asymptotic standard errors achieved, the smaller the width of the confidence intervals observed around the parameters estimates will be (ii) <i>t</i> -Radios will be maximized thus producing more reliable study results and analyst is able to minimize the sample size	(i) In general not orthogonal (not that important) (ii) Advanced knowledge of the parameter estimates is needed (iii) It needs more computation power

efficiency and wide applicability, *D*-efficient design achieves the best performance.

Further, most research focused on developing one of these experimental design methods by far. Though there may be a few of discussions about strength and shortcomings of every method separately, comparisons are rarely found in theory or practice area [12, 18, 19]. Here, we conclude the advantages and disadvantages of these three methods in Table 5. Also, a popularity rate is given as a reference of their applications in the field.

5. Conclusion

The SC experiment has been generally regarded as an effective method for discrete choice analysis, especially for newly introduced alternatives. The high cost on survey forces researchers to find more efficient design methods to obtain better estimation on parameters instead of investing a larger sample size. Though orthogonal design has been used as the

major experimental design method, orthogonality is not that important in the nonlinear discrete choice models. In this paper, we provide two state-of-the-art efficient designs: OOC and *D*-efficient design. By comparing orthogonal, OOC, and *D*-efficient design in both theory and practice, we find that efficient designs are more capable of producing more efficient data in the sense that more reliable parameter estimates can be achieved with an equal or lower sample size. The generation process requires the assumption of prior parameter estimates and model structure to construct AVC matrix. The result suggests a move away from orthogonal designs for SC experiments towards *D*-efficient designs, which make relative discrete choice models being more fit with such data.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study is supported by the 2014 Doctoral Innovation Funds of Southwest Jiaotong University, the Fundamental Research Funds for the Central Universities (no. A0920502051307-03), Specialized Research Fund for the Doctoral Program of Higher Education (no. 20130184110020), National Natural Science Foundation of China (no. 51178403), and Sichuan Province Science and Technology Support Project (no. 2011FZ0050).

References

- [1] H. Tan, G. Feng, J. Feng, W. Wang, Y. Zhang, and F. Li, "A tensor-based method for missing traffic data completion," *Transportation Research Part C*, vol. 28, pp. 15–27, 2013.
- [2] H. Tan, J. Feng, G. Feng, W. Wang, and Y.-J. Zhang, "Traffic volume data outlier recovery via tensor model," *Mathematical Problems in Engineering*, Article ID 164810, 8 pages, 2013.
- [3] H. Tana, B. Chenga, J. Fenga, G. Fenga, W. Wang, and Y. J. Zhangb, "Low-n-rank tensor recovery based on multi-linear augmented lagrange multiplier method," *Neurocomputing*, vol. 119, pp. 144–152.
- [4] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [5] W. Wang, X. Jiang, S. Xia, and Q. Cao, "Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation," *Safety Science*, vol. 48, no. 10, pp. 1248–1262, 2010.
- [6] J. J. Louviere, D. A. Hensher, and J. D. Swait, *Stated Choice Methods: Analysis and Applications*, Cambridge University Press, Cambridge, UK, 2000.
- [7] P. Galilea and J. D. D. Ortúzar, "Valuing noise level reductions in a residential location context," *Transportation Research D*, vol. 10, no. 4, pp. 305–322, 2005.
- [8] D. S. Bunch, J. J. Louviere, and D. Anderson, "A comparison of Experimental Design Strategies for Choice-based Conjoint Analysis with Generic-attribute Multinomial Logit Models," Working Paper, Graduate School of Management, University of California, Rickey Davis, Ala, USA., 1994.
- [9] J. J. Louviere and G. Woodworth, "Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data," *Journal of Marketing Research*, vol. 20, pp. 350–367, 1983.
- [10] J. Huber and K. Zwerina, "The importance of utility balance in efficient choice designs," *Journal of Marketing Research*, vol. 33, no. 3, pp. 307–317, 1996.
- [11] J. M. Rose and M. C. J. Bliemer, "Constructing efficient stated choice experimental design," ITLS Working Paper ITLS-WP-05-07, 2005.
- [12] M. C. J. Bliemer, J. M. Rose, and D. A. Hensher, "Efficient stated choice experiments for estimating nested logit models," *Transportation Research B*, vol. 43, no. 1, pp. 19–35, 2009.
- [13] D. J. Street and L. Burgess, "Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments," *Journal of Statistical Planning and Inference*, vol. 118, no. 1-2, pp. 185–199, 2004.
- [14] L. Burgess and D. J. Street, "Optimal designs for choice experiments with asymmetric attributes," *Journal of Statistical Planning and Inference*, vol. 134, no. 1, pp. 288–301, 2005.
- [15] D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, P. Zarembka, Ed., pp. 105–142, Academic Press, New York, NY, USA, 1974.
- [16] D. McFadden, "The choice theory approach to market research," *Marketing Science*, vol. 5, no. 4, pp. 275–297, 1986.
- [17] J. M. Rose and M. C. J. Bliemer, "Stated preference experimental design strategies," in *Transport Modeling Handbooks in Transport*, D. A. Hensher and K. Button, Eds., vol. 1, chapter 8, Elsevier Science, Oxford, UK, 2nd edition, 2007.
- [18] J. M. Rose, M. C. J. Bliemer, D. A. Hensher, and A. T. Collins, "Designing efficient stated choice experiments in the presence of reference alternatives," *Transportation Research B*, vol. 42, no. 4, pp. 395–406, 2008.
- [19] D. J. Street, L. Burgess, and J. J. Louviere, "Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments," *International Journal of Research in Marketing*, vol. 22, no. 4, pp. 459–470, 2005.

Research Article

Decision of National and Provincial Highway Asphalt Pavement Structure Based on Value Engineering

Yingwei Ren and Jingsong Shan

*Shandong Provincial Key Laboratory of Civil Engineering Disaster Prevention and Mitigation,
Shandong University of Science and Technology, Qingdao 266590, China*

Correspondence should be addressed to Jingsong Shan; cyhsjs@163.com

Received 16 November 2013; Revised 3 January 2014; Accepted 12 January 2014; Published 18 February 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Y. Ren and J. Shan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is important that decision of asphalt pavement structure requires overall considerations of the performance and financial investment. To have asphalt pavement structure fulfilling good reliability, the asphalt pavement structure decision was researched based on value engineering theory. According to the national and provincial highway investigation data in Shandong Province during the last decade, the asphalt pavement performance attenuation rules of traffic levels and asphalt layer thicknesses were developed, and then the road performance evaluation method was presented. In addition, the initial investments, the costs of road maintenance, and middle-scale repair in a period were analyzed. For the light traffic and medium traffic example, using the value engineering method, the pavement performance and costs of which thickness varies from 6 cm to 10 cm were calculated and compared. It was concluded that value engineering was an effective method in deciding the asphalt pavement structure.

1. Introduction

Under the effect of traffic load and natural factors, the asphalt pavements gradually suffer damages in different forms, such as ruts, cracks, and pot holes. As the damages increasing, the functional performance of asphalt pavements decreases, which affects the traffic safety and comfort. In order to maintain a relatively high level of service performance, measures such as road maintenance and periodic maintenance are needed.

When a road is built, the functional performance is good at the beginning with little damage. At this stage, the cost of maintenance is low. As the damages increase with time, the functional performance of asphalt pavements decrease continuously, which requires more maintenance costs to keep the asphalt pavements in good condition [1, 2]. Therefore, the choice of the asphalt pavement structure involves the factor of pavement long-term function as well as the cost during the pavement life cycle. Thus, it is an economic and technical problem. If the initial investment is large, the function is better in the later period and the maintenance

cost is low. By contrast, if pavement structure with low initial investment is used, the maintenance cost will be high in the later period. Therefore, the relationship between function and cost should be considered in pavement design. If good functional performance of pavement structure is to be achieved without excessive costs, decision index should be adopted for comprehensive technical and economic evaluation.

Value engineering is the systematic application of recognized techniques that identify the function of the product or service, establish a monetary value for that function, and provide the necessary function reliably at the lowest possible cost [3]. In recently years, value engineering has been researched and widely practiced in the construction industry and become an integral part in the development of many projects [4–7]. But it was few which value engineering was used to decide asphalt pavement structure.

This paper makes choices in asphalt pavement structure by using value engineering method in the engineering economics and tries to provide guidance for highway administration department in decision making.

2. Theory of Value Engineering

Value engineering (VE) is a method to achieve necessary product function with the lowest possible life-cycle cost. Expression of value engineering is as follow:

$$V = \frac{F}{C}, \quad (1)$$

where V is value; F is function; C is cost.

Figure 1 shows the relationship between function and cost. C_1 indicates the initial construction cost and C_2 indicates the cost in the process of use. The relationship between C_1 and C_2 can be seen from this figure. Product function will increase either by increasing C_1 or decreasing C_2 . Product function will decrease either by decreasing C_1 or increasing C_2 . The purpose of value engineering is to check and give consideration to both costs and to figure out the optimum point of function so as to minimize the overall cost (C) [8].

According to value engineering theory, the methods of improving the value are as follows:

- (1) reducing the cost while improving the function;
- (2) keeping the function unchanged while reducing the cost;
- (3) keeping the cost unchanged while increasing the function;
- (4) increasing the cost slightly while increasing the function greatly;
- (5) reducing the function slightly while reducing the cost greatly. The reduced function should be within the users' permission.

As is shown in formula (1), the pavement is regarded as a product with relatively long life time and if reasonable pavement structure is to be achieved, function and cost should be taken into account, so as to achieve high function at relatively low cost. Therefore, applying the value engineering method to choose the asphalt pavement structure requires determinations of both functional performance evaluation and the cost calculation method, which are shown, respectively, in the sections below.

3. Performance Evaluation Method

3.1. Pavement Performance Model. The asphalt pavement performance can be expressed by single index, such as crack rate, and amount of ruts. It can also be expressed by comprehensive index such as PCI (Pavement Condition Index) and RQI (Riding Quality Index). A single index can directly show the severity and development of a single damage. So the maintenance measures can be established with more pertinence. Comprehensive index, synthesized with many single indexes, can reflect the service condition within the whole road at a certain time. This paper uses PCI (Pavement Condition Index) to evaluate the functional performance of pavement.

The asphalt performance is affected by many factors, such as the asphalt pavement layer thickness, the thickness and

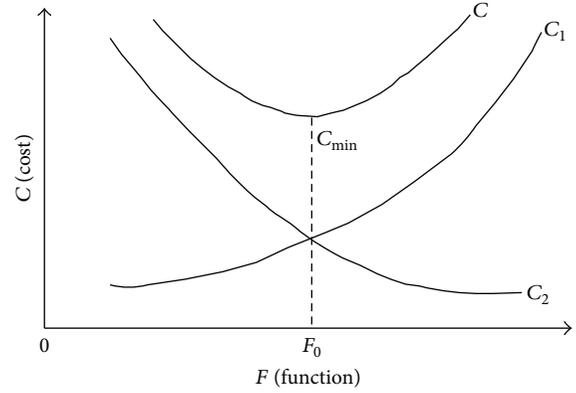


FIGURE 1: Relationship between function and cost.

strength of base layer, the subgrade strength, the natural environment, and the automobile load. And the quality of construction also has a direct impact on pavement performance. Therefore, the asphalt pavement performance evaluation is a complicated task. And asphalt pavement performance evaluation is related to highway investment decision, design, and later maintenance, so it is a necessary task. In the 1960s, American Association of State Highway and Transportation Officials (AASHTO) established the pavement performance attenuation equation according to the experimental data and Pavement Service Index (PSI), which is one of the first road performance models [9, 10]. The model is

$$PSI = PSI_0 - (PSI_0 - PSI_t) \left(\frac{ESAL_t}{ESAL_0} \right)^\beta. \quad (2)$$

Based on investigation, the PSI_0 at the initial state of the pavement and the PSI_t at the critical state as well as the accumulated equal single axle load $ESAL_t$ near the collapse state are firstly determined in this method. And the PSI value is determined according to the accumulated equal single axle load $ESAL_t$ at a certain time.

During the past decades, many people started to do related research on the expression forms of pavement function. Professor Sun Lijun put forward the expression form of Pavement Condition Index (PCI), as is shown in formula (3), which can simulate the pavement function attenuation law through the changes in α and β value [11]. Factors such as the pavement layer thickness, the surface deflection, the accumulated equal single axle load, and the natural environment should be taken into consideration when determining the α and β value, thus to make it become a comprehensive asphalt function model [12–14]. Based on the national and provincial highway research data in Shandong Province, the value of α and β is established with this model in this paper:

$$PCI = PCI_0 \left\{ 1 - \exp \left[- \left(\frac{\alpha}{y} \right)^\beta \right] \right\}. \quad (3)$$

The newly built and rebuilt asphalt pavement layer thicknesses in China are mostly 6 cm, 8 cm, and 10 cm, with 7 cm and 9 cm accounting for the minority. When deciding the

TABLE 1: Questionnaire of asphalt pavement condition of national and provincial highway.

Number	Asphalt pavement layer thickness	Base layer type and thickness	Damages start time	Years of intermediate maintenance	Large vehicle volume	Notes
1						
2						
3						

α and β value, asphalt pavement is divided by thickness, namely 6 cm, 8 cm, and 10 cm. And functional performances are counted and analyzed under different thickness groups. In each group, traffic load is considered as the main factor. Because of inadequate construction budget, observation stations are not equipped with axle load weighing appliances, so only traffic flow can be obtained according to vehicle types in different months. According to vehicle classification, large- and medium-sized trucks and motor buses have largely effects on pavement damages than other vehicles; thus they are considered in this research. The data collected are listed in Table 1.

3.2. The Solving Process of α and β Values. The determination of α and β is very important to expressing pavement function through formula (3). According to the national and provincial highway research data in Shandong Province, formula (4) can be obtained by inputting PCI of different pavements (a_1 and a_2 represent different years) in formula (3):

$$PCI_1 = PCI_0 \left\{ 1 - \exp \left[- \left(\frac{\alpha}{a_1} \right)^\beta \right] \right\}, \quad (4)$$

where PCI_1 represents PCI value after a_1 years and PCI_2 represents PCI value after a_2 years. Supposing the initial pavement condition is good and the PCI is 100, then formula (3) can be simplified into formula

$$\begin{aligned} \ln A_1 &= - \left(\frac{\alpha}{a_1} \right)^\beta, \\ \ln A_2 &= - \left(\frac{\alpha}{a_2} \right)^\beta, \end{aligned} \quad (5)$$

where $A_1 = 1 - PCI_1/100$, $A_2 = 1 - PCI_2/100$.

By taking logarithm on both sides of the formula (5), formula (6) is obtained:

$$\begin{aligned} a &= \frac{\ln(\ln(1/A_1)) \times \ln a_2 - \ln(\ln(1/A_2)) \times \ln a_1}{\ln(\ln(1/A_1)) - \ln(\ln(1/A_2))}, \\ \beta &= \frac{\ln(\ln(1/A_1)) - \ln(\ln(1/A_2))}{\ln a_2 - \ln a_1}. \end{aligned} \quad (6)$$

α and β can be obtained through formula (6). When the asphalt pavement layer thickness is 10 cm, α and β curves are drawn in Figure 2 considering different daily large vehicle volume. After comparison in different fitting methods, the power exponent form is used, as the fitting curve is consistent

with the variation trend of measured data. The relationship is established between α , β and the daily large vehicle volume per lane, which is listed in formula

$$\begin{aligned} \alpha &= 98.6Q^{-0.369}, \\ \beta &= 4.469Q^{-0.267}. \end{aligned} \quad (7)$$

Using the same method, relations between α and β under different asphalt pavement layer thicknesses (6 cm and 8 cm) can be established. Take asphalt pavement layer thickness (h) as a variable and use h to express coefficients and exponents in the relation of α and β . Formula (8) is as follows:

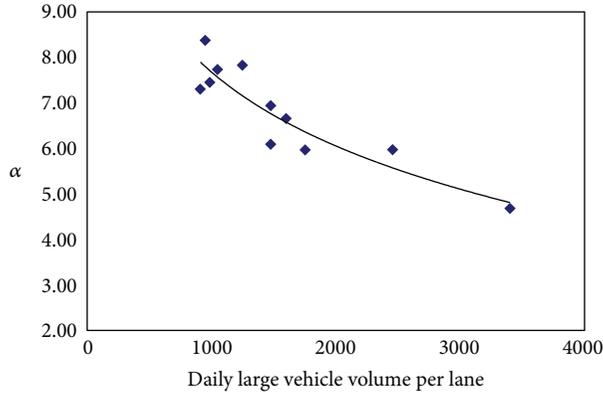
$$\begin{aligned} \alpha &= 1431h^{-1.164}Q^{0.051h-0.868}, \\ \beta &= 0.171e^{0.323h}Q^{-0.0465h+0.2}. \end{aligned} \quad (8)$$

For instance, supposing the daily large vehicle volume per lane is 1000, α and β can be obtained firstly through formula (8), then putting them into formula (3) and PCI value of the road will be got. The PCI curve with years is drawn in Figure 3 considering the thickness variations of asphalt layer. As can be seen from the figure, when the asphalt pavement layer thickness changes from 6 cm to 10 cm, the age of PCI down to 60 increases from 5 to 9 years. This implies that under the traffic level, the deteriorating rate is quicker when asphalt thickness is thinner, and to improve pavement property, asphalt thickness should be increased appropriately.

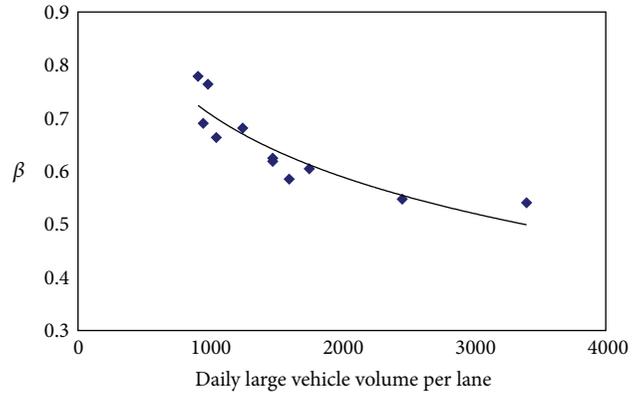
3.3. Performance Evaluation Method of Asphalt Pavement Structure. According to value engineering, pavement function index is needed when analyzing the value of different pavement schemes. The curve in Figure 3 can express the functional performance attenuation law of asphalt pavement with different layer thicknesses, but it cannot directly make comprehensive comparison through this PCI curve. Therefore, pavement performance is expressed by the area under the PCI curve in Figure 4. Before calculating the area under the PCI curve, the lowest acceptable level is determined firstly according to pavement rank and function requirements, and then, integration method is used to calculate the area S_d .

4. Cost Calculation

After construction completion, the highways need repeated maintenance, overlay and rebuilt during the life time. It is obvious that highway engineering needs constant investment on maintenance during more than ten years or even decades,



(a) Relationship between α and daily large vehicle volume per lane



(b) Relationship between β and daily large vehicle volume per lane

FIGURE 2: Thickness of asphalt pavement layer is 10 cm.

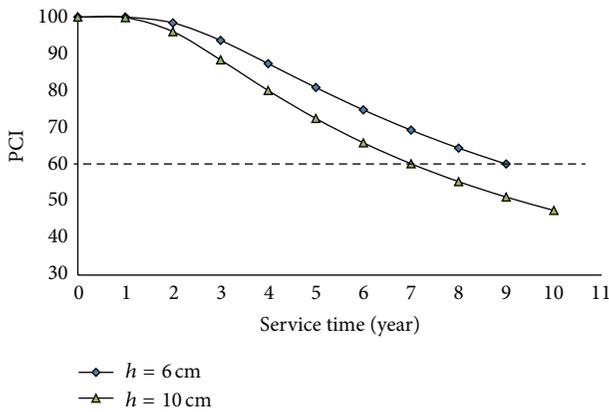


FIGURE 3: Regularity of PCI with years.

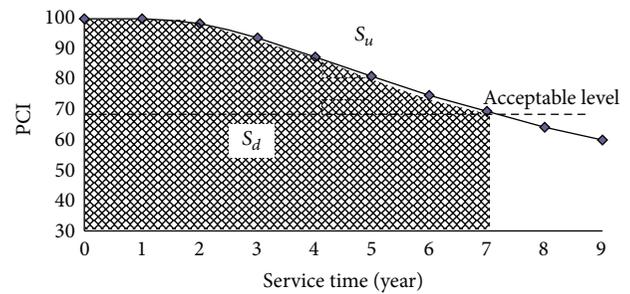


FIGURE 4: Area under PCI curve.

which should not be underestimated. When considering pavement cost, except initial construction cost, the maintenance cost should be considered integrally. In the 1960s, American Association of State Highway and Transportation Officials (AASHTO) put forward the concept of life-cycle cost. In 1986 and 1993, AASHTO Pavement Design Guide was published twice, both of which actively advocated the adoption of life-cycle cost analytical method. In 1998, American Association of State Highway and Transportation Officials issued the standard, which also advocated the adoption of life-cycle cost analytical method when making decisions on highway investment [15–17]. In 2008, transportation department in the state of Mississippi conducted an investigation with 21 states. Among them, 18 states, including Alabama and New Jersey, applied life-cycle cost analytical method when making decisions on highway investment. In China, Professor Yao Zukang along with his research group did some research in light of China’s conditions, and the results have not been applied in decision making of highway structures yet.

Life-cycle cost includes initial construction cost, maintenance cost, overlay rebuilt cost, and user cost. As our country fall behind developed counties in the users’ cost research

field, there is no suitable calculation model for users’ cost. In this paper, users’ cost is not included in life-cycle cost. So life-cycle cost includes road management department’s investment during the life cycle, namely, initial construction cost, maintenance cost, and overlay rebuilt cost.

(1) *Maintenance Cost.* The calculation of maintenance cost adopts the computation model by Tongji University, as is show in formula

$$MC_i = a + b \times 10^{-6} (100 - PCI_i) \times AADT_i, \quad (9)$$

where $AADT_i$ is ADT (Average Daily Traffic) in the year of i ; PCI_i is road condition index in the year of i ; and MC_i is maintenance cost in the year of i (Yuan/m²).

(2) *Salvage Value.* At the end of the analysis period, if the pavement functional performance is still above the lowest acceptable level and roads can be passed through, the current value left is called the salvage value. As is shown in formula (10), the salvage value is modeled as a linear regression function, in which the salvage value reduces linearly with age, without being considered in the case of the nonlinear reduction of the pavement functional performance. Therefore, in this paper, the computation strategy is conducted in which the area under the PCI curve is computed to indicate the salvage value based on formula (11), as is shown in Figure 5.

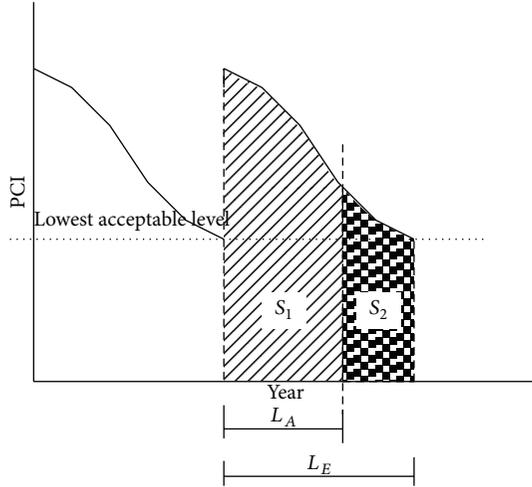


FIGURE 5: Pavement salvage value.

Consider

$$S_v = \left(1 - \frac{L_A}{L_E}\right) C_r, \quad (10)$$

where S_v is salvage value; C_r is the last intermediate maintenance overlay cost in analysis period; L_E is years between the last time maintenance and the end of life time; L_A is years between the last maintenance time and the end of analysis period.

Consider

$$S_v = \left(1 - \frac{S_1}{S_1 + S_2}\right) C_r, \quad (11)$$

where S_1 is area under curve L_A and S_2 is area under curve L_E .

(3) *Present Value of Cost.* Various costs in the analysis period are discounted into present value according to a certain discount rate. Therefore, economic efficiency of different schemes can be compared through the uniform present value. Capital flow in road investment is shown in Figure 6.

One has

$$P_0 = C_0 + \sum_{i=1}^n f_i (MC_i + RC_i) - f_n S_v, \quad (12)$$

where P_0 is total cost of present value; C_0 is the initial construction cost; MC_i is the maintenance cost in the year of i ; RC_i is the intermediate maintenance overlay cost in the year of i ; S_v is salvage value; f_i, f_n is discount rate in the year of i and in the year of n , respectively.

5. Pavement Structure Decision Based on Value Engineering

According to the proceeding discussion, the analysis of pavement functional performance is vital when making decisions of pavement structure based on value engineering. PCI curve

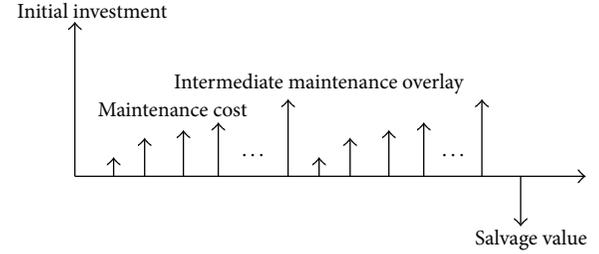


FIGURE 6: Capital flow in road investment.

is closely related to pavement structure function evaluation, daily maintenance cost, intermediate maintenance time, and salvage value calculation. Therefore, PCI curve should be established in the first place when making decisions of pavement structure based on value engineering, and on this basis, cost is further analyzed. Figure 7 is the flow chart of value calculation.

Take the example of light traffic and medium traffic of national and provincial asphalt pavement in Shandong. Supposing daily large vehicle volume per lane is 600 vehicles in light traffic which has average daily traffic 6000 vehicles, large vehicle volume per day per lane is 600 vehicles in medium traffic which has average daily traffic 10000 vehicles. Suppose the asphalt pavement layer thickness is 6 cm to 10 cm and the base layer is double-layered cement stabilized crushed stone base. Comparison of different asphalt pavement layer thicknesses is made according to value engineering, which is followed below.

Based on the PCI formula above, the PCI curve of this case where the asphalt pavement layer thicknesses are 10 cm, 8 cm, and 6 cm is drawn in Figure 8, respectively. The analysis period is 15 years, during which asphalt pavement of 6 cm needs twice intermediate overlay, while asphalt pavement of 8 cm and 10 cm need only once. The area under PCI curve is calculated through integration method, so as to express the function of different asphalt pavements.

Initial investment is determined according to calculation method of life-cycle cost and based on asphalt pavement price in Shandong in 2010. Supposing the price inflation rate is 5% and the discount rate is 10%, the total daily maintenance cost on a yearly basis and the intermediate maintenance costs are calculated in different asphalt pavement layer thicknesses and are converted into present value. When the asphalt pavement layer thickness is 10 cm under light traffic, the calculation process of costs is listed in Table 2. Other costs with different thicknesses are calculated in the same way and are listed in Table 3.

As can be seen from Table 3, from the perspective of function, asphalt pavement's performance is best when the layer thickness is 10 cm under light traffic, which is followed by 8 cm and 6 cm. However, from the perspective of life-cycle cost, asphalt pavement with layer thickness of 10 cm is the most expensive and it is relatively cheaper when the layer thickness is 8 cm and 6 cm. This is caused by the high initial investment (layer thickness is 10 cm) and the money saved by maintenance investment and intermediate maintenance cost

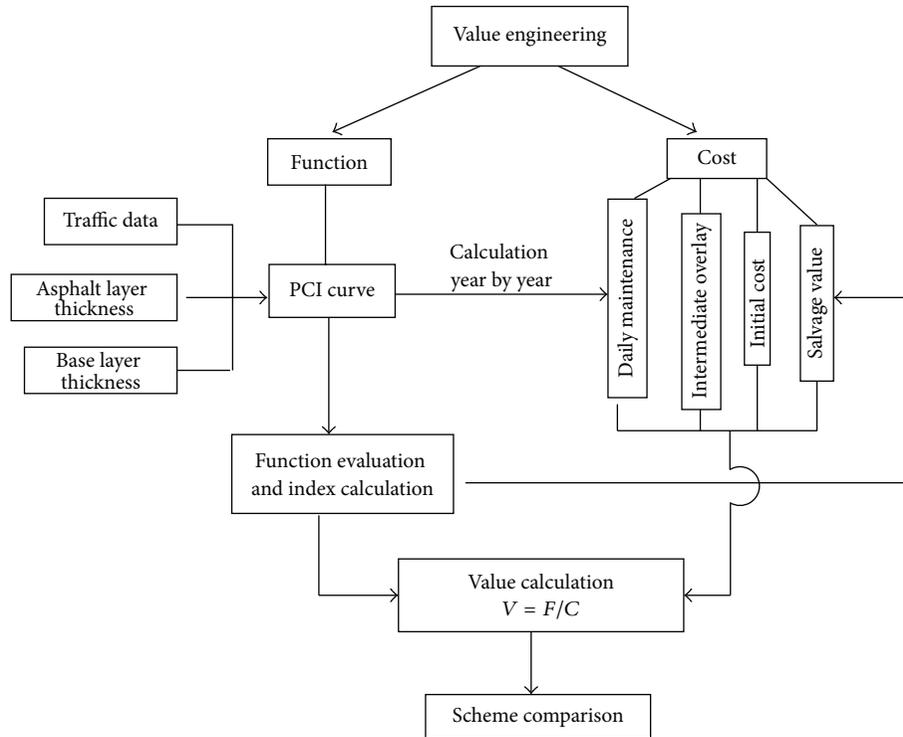


FIGURE 7: Value engineering calculation procedure.

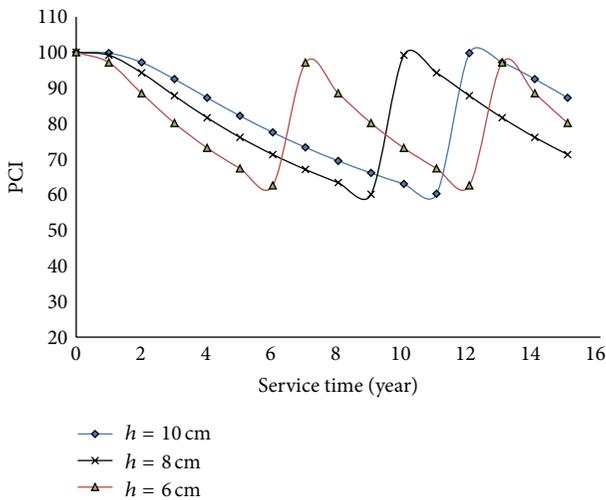


FIGURE 8: PCI curves of different asphalt pavement layer thicknesses.

cannot make up for the high initial cost. So in the case of light traffic, the asphalt pavement is the most suitable when the layer thickness is 8 cm.

The functional performance is the best when the layer thickness is 10 cm under medium traffic, which is followed by 8 cm and 6 cm. The life-cycle cost is relatively high when the layer thickness is 10 cm and 6 cm, the life-cycle cost is relatively low when the layer thickness is 8 cm, which illustrates the fact that thin layer thickness of asphalt pavement

can accelerate the damage speed so as to largely increase the maintenance cost. It can be concluded that the value is lowest when the asphalt pavement layer thickness is 6 cm and it is economically more reasonable when the asphalt pavement layer thickness is 8 cm or 10 cm. With the increase in traffic, namely, when the daily large-sized vehicles volume per lane goes to over 1200, asphalt pavement has better functional performance with layer thickness over 10 cm.

6. Conclusions

In this paper the theory of value engineering is applied to the decision-making process of the national and provincial highway asphalt pavement structures to find the best point of pavement performance and economic investment. The decision of asphalt pavement structure based on value engineering is reliable and economic. The main achievements are as follows.

(1) According to national and provincial highway investigation data in Shandong Province during the last decade, the relationship is established between α , β and the two factors, namely, the asphalt pavement layer thickness and the daily large-sized vehicle volume per lane. The attenuation law of national and provincial highway asphalt pavement is easily analyzed with these two parameters α and β .

(2) Based on the PCI attenuation curve, the pavement performance is expressed by the area under the PCI curve and the calculation method of pavement performance is established, which can be applied to pavement performance

TABLE 2: Life-cycle cost (Yuan/m²).

Service life period	PCI	Initial investment	Maintenance investment	Intermediate maintenance and overlay	Salvage value	Present value of maintenance	Present value of intermediate maintenance overlay	Present value of salvage value
0	100.00	100						
1	99.79		1.01			0.94		
2	97.18		1.17			1.00		
3	92.47		1.45			1.15		
4	87.23		1.77			1.30		
5	82.17		2.07			1.41		
6	77.50		2.35			1.48		
7	73.29		2.60			1.52		
8	69.50		2.83			1.53		
9	66.09		3.03			1.52		
10	63.01		3.22			1.49		
11	60.24		3.39			1.45		
12	99.79		1.01	71.8		0.40	28.5	
13	97.18		1.17			0.43		
14	92.47		1.45			0.49		
15	87.23		1.77		-41.29	0.56		-13.02
	Total present value	100				16.67	28.5	-13.02
	Life-cycle cost					132.15 Yuan/m ²		

TABLE 3: Calculation of value.

Number of large-sized vehicles per lane per day		<i>h</i> = 10 cm	<i>h</i> = 8 cm	<i>h</i> = 6 cm
600	function (area covered by PCI)	1265	1239	1213
	Life-cycle cost	132.15	122.27	122.94
	value (function/cost)	9.57	10.13	9.87
1000	function (area covered by PCI)	1228	1190	1173
	Life-cycle cost	153.28	145.67	158.77
	value (function/cost)	8.01	8.17	7.39

evaluation and comparison between different pavement structures.

(3) In the example of light traffic and medium traffic of national and provincial highway, value engineering theory is applied to analyze the technicality and economic efficiency of asphalt pavement with different layer thicknesses (10 cm, 8 cm, and 6 cm). It can be concluded that under light traffic, the value is the highest when the layer thickness is 8 cm and the value is the lowest when the layer thickness is 10 cm. Under light traffic, the value is maximized in the case of the asphalt pavement layer thickness at 8 cm and minimized at 10 cm. However, under medium traffic, the value is maximized when the asphalt pavement layer thickness equals 9 cm and minimized when it is 6 cm. Through the analysis of this example, it can be indicated that using engineering value

theory is an effective way in deciding reasonable pavement structures.

(4) When deciding pavement structures, the schemes with the highest value or closer to the highest value should be selected and it is not appropriate to choose the schemes with the highest function or the lowest life-cycle cost. With consideration of the project investment conditions, schemes with the lowest economic cost as well as satisfying the pavement functions should be chosen.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was funded by “Shandong Provincial Natural Science foundation (no. ZR2011EEQ027)” and “Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University (no. 201002)”. The authors gratefully acknowledge their support.

References

- [1] T. F. Fwa and K. C. Sinha, “Pavement performance and life-cycle cost analysis,” *Journal of Transportation Engineering*, vol. 117, no. 1, pp. 33–46, 1991.
- [2] Federal Highway Administration, *Life-Cycle Cost Analysis in Pavement Design-in Search of Better Investment Decisions*, Pavement Division Interim Technical Bulletin, FHWA, US DOT, Washington, DC, USA, 1998.
- [3] Q. S. Liu and W. S. Shen, *Value Engineering Management and Practice*, Beyond Business Administration Consultant Co., 1995.
- [4] L. D. Miles, *Techniques of Value Analysis and Engineering*, McGraw-Hill, New York, NY, USA, 2nd edition, 1972.
- [5] A. Omigbodun, “Value engineering and optimal building projects,” *Journal of Architectural Engineering*, vol. 7, no. 2, pp. 40–43, 2001.
- [6] S. Assaf, O. A. Jannadi, and A. Al-Tamimi, “Computerized system for application of value engineering methodology,” *Journal of Computing in Civil Engineering*, vol. 14, no. 3, pp. 206–214, 2000.
- [7] X. Mao, X. Zhang, and S. M. Abourizk, “Enhancing value engineering process by incorporating inventive problem-solving techniques,” *Journal of Construction Engineering and Management*, vol. 135, no. 5, pp. 416–424, 2009.
- [8] D. L. Younker, *Value Engineering Analysis and Methodology*, Marcel-Dekker, New York, NY, USA, 2003.
- [9] AASHTO, *AASHTO Guide For Design of Pavement Structures. Part II*, Washington, DC, USA, 1993.
- [10] SHRP, “Evaluation of the AASHTO design equations and recommended improvements,” Tech. Rep. SHRP-P-394, Washington, DC, USA, 1994.
- [11] S. Lijun and L. Liping, “Performance-based whole-life design method for asphalt pavement,” *Journal of Tongji University*, vol. 31, pp. 832–837, 2003.
- [12] S. Labi, *Impact of highway pavement maintenance activities [Ph.D. thesis]*, School of Civil Engineering, Purdue University, West Lafayette, Ind, USA, 2001.
- [13] S. Labi and K. C. Sinha, “Grassroots perspectives of pavement preventive maintenance—a state-of-practice survey in Indiana,” *International Journal of Pavements*, vol. 1, no. 3, pp. 44–57, 2002.
- [14] S. Labi and K. C. Sinha, “Measures of short-term effectiveness of highway pavement maintenance,” *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 673–683, 2003.
- [15] FHWA, “Life-cycle cost analysis in pavement design-in search of better investment decisions,” Tech. Rep. FHWA-SA-98-079, Federal Highway Administration, U.S. Department of Transportation, Washington, DC, USA, 1998.
- [16] FHWA, *Life-Cycle Cost Analysis Primer*, US Department of Transportation, Washington, DC, USA, 2002.
- [17] G. Lamptey, M. Z. Ahmad, S. Labi, and K. C. Sinha, “Life-cycle cost analysis for INDOT pavement design procedures,” Tech. Rep. FHWA/IN/JTRP-2005/28, Joint Transportation Research Program, 2005.

Research Article

Impact of Road Bends on Traffic Flow in a Single-Lane Traffic System

Zeng Junwei,¹ Qian Yongsheng,¹ Xu Dejie,¹ Jia Zhilong,¹ and Huang Zhidan²

¹ School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

² School of Mechatronic Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Correspondence should be addressed to Qian Yongsheng; qianyongsheng@mail.lzjtu.cn

Received 26 September 2013; Revised 5 December 2013; Accepted 19 December 2013; Published 4 February 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Zeng Junwei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Taking the characteristics of road bends as a research object, this work proposes the cellular model (CA) with road bends based on the NaSch model, with which the traffic flow is examined under different conditions, such as bend radius, bend arc length, and road friction coefficient. The simulation results show that, with the increase of the bend radius, the peak flow will be continuously increased, and the fundamental diagram will become more similar to that of the classic NaSch model; the smaller the bend radius is, the easier it is for the occurrence of blockage; for different bend lengths, all the corresponding traffic flows show that the phenomenon of go-and-stop and the bends exert slight inhibitory effect on traffic flow; under the same bend radius, the inhibition effect of the bends on the traffic flow will be weakened with the increase of the friction coefficient.

1. Introduction

The contradiction between the road line and the limitation of the actual terrain (topography, towns and villages, natural lakes and rivers, etc.) should be fully considered in expressway. In addition, these factors should be taken into account comprehensively to ensure driving comfort, alignment continuity, easy operating, and road safety. Therefore, such elements as straight line, transition curve, circular curve, superelevation, and longitudinal grade should be taken into full consideration in the road line design so as to meet the design requirements to the greatest scale. It appears to be particularly an urgent issue to research the traffic flow characteristics on bends, to establish models for the simulation of vehicles on bends, and to analyze the impact of bends on road traffic flow.

The theory of cellular automata has been widely used in the field of road traffic flow. The CA models for traffic flow mainly include NaSch model, FI model, and BML model. It is originally based on the Wolfram cellular automata models [1]. NaSch model is a one-dimensional CA model, which can successfully simulate the go-and-stop traffic wave as well as the traffic flow transformation from free phase to congestion

one. As to different factors, a variety of valuable traffic flow CA models have been established based on NaSch model [2]. Ishibashi and Fukui investigated dynamic effect of deceleration in advance on traffic flow [3]. Biham et al. researched the influence of headway space on traffic flow [4]. Qian et al. proposed the cellular automaton model that considers lane-control and speed-control to research the characteristics of highway traffic flow [5]. Yang et al. use the cellular automaton model to study the electric vehicle's electricity energy consumption under the different conditions [6]. Tang et al. proposes a cellular automaton (CA) model for dynamic simulations under mixed traffic conditions [7–10]. Ge et al. put some new CA model for intelligent transportation system, which makes the experiences for CA model of trains system [11, 12]. Li et al. propose a cellular automata model to simulate the traffic flow to analyze how the length of speed-limited section, train running time interval, and the speed-limit value affect the traffic flow. The decrease of the length of speed-limited section, the moderate increase of train running time interval, and the increase of the speed-limit value can improve the green light running time of the trains [13–16].

Currently, research work of traffic flow on road bends are limited. They are mainly conducted from the viewpoint

of bend alignment combined with road line design. In 1995, Rao studied the alignment design of road bends [17], where the limit value of the bend radius, the tangent length of the transition curve, and the curvature change were discussed emphatically. Bentley and Gallagher explored the setup of the transition curve and made a summarized analysis on the bend alignment design of the US based on horizontal and vertical curve [18]. Dabbour et al. examined the vertical curve of the bends nonlinearly, proposed a road safety evaluation method, and analyzed the values of vertical curves for different alignment types [19]. Persaud et al. presented the guidance law for the hazard detection of road bends, pointing out that it is a low-cost traffic safety improvement measure with obvious effect to set effective bend hazard warnings [20].

Research works of road bends in China today are mainly conducted from the viewpoint of alignment design. Mu and Yang introduced the concept of possible velocity into road alignment design, providing a theoretical basis for the selection of highway design technical indicators by using possible velocity prediction method [21]. Liang and Xue examined the effect of curvature, arc length, and the road surface friction coefficient on traffic flow and reproduced the phenomenon of go-and-stop in actual traffic conditions [22]. Wang et al. revealed the characteristics and laws of running speed on different types of roads and its relationship with horizontal curve radius through the analysis of different types of road running speed prediction models and pointed out horizontal curve design values for all levels of roads [23].

In summary, many domestic and foreign scholars have largely improved the study of road traffic flow or make a good progress at the aspect of CA model, but there are few studies on bend road by CA model. This work conducts further studies on the characteristics of traffic flow on road bends on the basis of the previous studies and makes supplement and perfection to the relevant theories, such as improved algorithm of NaSch model which is used on road bends. It will make a good development for the CA using on the traffic flow.

2. Characteristics of Road Bends

2.1. Characteristics of Road Bends Alignment Elements. The horizontal alignment includes straight line, circular curve, and transition curve. When traveling on straight line section, the vehicle is easy to control with better comfort. Thus the driver will be able to travel at high desired speed with little fluctuation. The transition curve located between straight line and circular curve plays a transition role; when traveling on transition curve section, the vehicle will accelerate when entering straight road section, will decelerate when entering circular curve section, and will enter the next section with appropriate speed. The length and radius of the circular curve have direct impacts on vehicle speeds: the smaller the radius, the worse the stability and the harder it is to manipulate the vehicle. As a result, the running speed will be reduced generally. On a circular curve section, the vehicle

will be traveling at a constant speed and will be accelerated or decelerated at the end of the horizontal curve. When the radius of the circular curve is greater than the critical value, the impact of curvature on vehicle speed will disappear, and the circular curve can be treated as a straight line theoretically.

The vertical alignment mainly includes longitudinal gradient and vertical curve. The length and grade of the longitudinal gradient have a great impact on vehicle speed: under normal circumstances, the greater the length and grade, the lower the vehicle speed. Vertical curve is a circular curve connecting two longitudinal gradients, the purpose of which is to ensure smooth transition from one longitudinal gradient to the other; compared to the longitudinal gradient, a vertical curve has less impact on the vehicle speed.

A curved ramp is the connection of horizontal curve and longitudinal gradient, in which the radius of the horizontal curve $r \in [125, 1000]$ and the grade of the longitudinal gradient $i \in [-6\%, 2\%]$. The curved ramp is a road accident blackspot, and its design quality must be critical enough to ensure traffic safety.

2.2. Characteristics of Traffic Flow on Road Bends. Road bends are the special sections of an expressway and are quite different from the straight-line sections. The characteristics of traffic flow on road bends can be listed as follows.

(1) *Significant Speed Difference.* When traveling on road bends, due to the factors such as a driver's age, driving experience and his personality, and his familiarity with the road, together with the vehicle's performance, the vehicle speed on road bends varies considerably and the speed difference is significant between vehicles of different types and different models, and the mutual interaction between vehicles will be naturally serious.

(2) *Longitudinal Gradient Exerts Greater Impact on Vehicles.* In expressway design, the occurrence of longitudinal gradient exerts direct impact on large trucks: in climbing, the speed of the trucks is significantly lower, making its travel quite difficult; in downhill, the trucks are accelerated and are hard to decelerate, which is likely to be prone to traffic accidents and will have great impact on the vehicle operation and the road capacity.

(3) *Degradation in the Vehicle Handling Performance.* The friction between vehicle tires and road surface is used as the centripetal force needed for vehicle traveling on road bends. The centripetal force will lead to the lateral offset of vehicles, thereby increasing the requirements of handling performance in direction. The higher the vehicle speed is, the more difficult it is to manipulate the vehicle.

(4) *Reduced Driving Comfort.* When traveling on road bends, the lateral force coefficient increases, resulting in the reduction of the vehicle stability; when the lateral force is too large, the occupants will be discomforted; if the speed is too large, a rollover accident will possibly occur.

3. CA Model of Single-Lane Bends

3.1. Model Establishment. A CA model of a bend on a single-lane is established in this work, as shown in Figure 1: the lane of this model is composed of a total of L (set $L = 7000$ in this work) cells, and the length of each cell is 1 m; namely, the actual length of the lane in the model is 7 km. In the initial moment, vehicles are equally spaced on the lane. The length of each vehicle is L_{car} ; vehicle speed on the bend is V_{safe} ; the maximum velocity for the vehicles $V_{\text{max}} = 35$ cells/s = 126 km/h; curve length is L_{bend} . The bend is set at the intermediate position of the lane, the vehicle deceleration is b , the friction coefficient of the lane is μ , the gravity acceleration is g , and the curve radius is r . Periodic boundary condition can be adopted in this model; that is, when a vehicle running out of the L th cell, it will enter the system again from the first cell. It can be inferred from Figure 1 that the lane can be divided into straight-line section, transition section, and the bend section. The revolution laws are different for a vehicle when it is on different sections. On the straight-line section, because the driver is in a relaxed state, his driving behaviors such as sudden shift or change lanes are unlikely to occur. On the transition section, a vehicle may be decelerated until entering the bend due to the fact that a driver must adjust himself to the vehicle state according to his driving experience and the vehicle's running speed, and the sharp deceleration may occur at a higher vehicle speed, leading to the increased uncertainty on road safety. After entering the bend, the vehicle needs to turn around the bend, and the driver is normally driving cautiously, together with the case where vehicle speed is typically reduced before entering the bend. Therefore, the vehicle will be accelerated to the safe speed after entering the bend to pass through as quickly as possible. Different evolutionary rules are developed for specific driving behaviors on different sections, which can reflect the driving state more realistically.

Before the enactment of the model operating rules, the following assumptions are made in this paper accordingly.

- (1) The external environment is good, without interference of pedestrians or horizontal traffic, and the effect of weather and other factors is ignored hereby.
- (2) On the straight-line section, the vehicles are running at the desired speed as possible, and vehicle speed will not increase after the vehicle reached its maximum speed.
- (3) Vehicle speed is decelerated on the transition section before entering the bend.
- (4) The speed of vehicles on the circular curve is not higher than the safe speed; namely, there is no accident caused by driving with too higher speed.
- (5) No significant difference in vehicle performance is considered in this work; namely, the effects of acceleration performance, being empty, and with load are all ignored hereby.

3.2. Model Operating Rules. Based on the classic NaSch model, considering the characteristics of driving behavior on

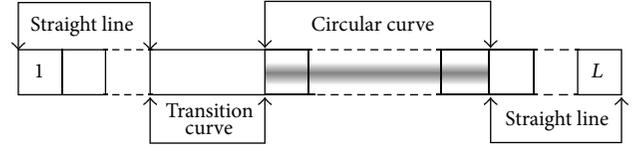


FIGURE 1: Structure diagram of the road bend.

the bend, the effects of road bend conditions (curve radius r , curve arc length s , and road surface friction coefficient μ) and vehicle characteristics on the vehicle speed and latency probability, the evolution rules of the model in this paper are summarized accordingly, which include the evolution rules of normal sections, transition sections and bend sections, which can be detailed below.

(1) Evolution Rules of Normal Sections

Step 1 (acceleration). Consider

$$V_n(t) = \min(V_n(t) + 1, V_{\text{max}}). \quad (1)$$

Step 2 (deceleration). Consider

$$V_n(t) = \min(V_n(t), \text{gap}_n(t)). \quad (2)$$

Step 3 (randomization deceleration with probability P10). Consider

$$V_n(t) = \max(V_n(t) - 1, 0). \quad (3)$$

Step 4 (location update). Consider

$$X_n(t) = X_n(t) + V_n(t). \quad (4)$$

(2) Evolution Rules of Transition Sections

Step 1 (acceleration). If $V_n(t) < V_{\text{expect}}$, then vehicle will be accelerated with probability P2 as

$$V_n(t) = V_n(t) + D1. \quad (5)$$

Step 2 (deceleration). If $V_n(t) > V_{\text{expect}}$, then vehicle will be decelerated with probability P3 as

$$V_n(t) = V_n(t) - D2. \quad (6)$$

Here $V_{\text{expect}} = (V_n^2 - V_{\text{safe}}^2)/2b$ is the desired running speed on the transition section, which is considered as reference to reduce the sharp shift of vehicle speed. When the vehicle speed is less than V_{expect} , the vehicle will be accelerated (the acceleration is $D1$) with probability $P2$; when the vehicle speed is larger than V_{expect} , the vehicle will be decelerated (the acceleration is $D2$) with probability $P3$.

Step 3 (safe deceleration). Consider

$$V_n(t) = \min(V_n(t), \text{gap}_n(t)). \quad (7)$$

Step 4 (deceleration with probability P11). Consider

$$V_n(t) = \max(V_n(t) - 1, 0). \quad (8)$$

Step 5 (location update). Consider

$$Xn(t) = Xn(t) + Vn(t). \quad (9)$$

(3) Evolution Rules of Bend Sections

Step 1 (acceleration). If $Vn(t) < V_{\text{safe}}$, then vehicle will be accelerated with probability $P4$ as

$$Vn(t) = \min(Vn(t) + 1, V_{\text{safe}}). \quad (10)$$

Here this means that the vehicle will be accelerated with probability $P4$ when the vehicle speed is less than the safe speed.

Step 2 (forced deceleration). If $Vn(t) > V_{\text{safe}}$, then the vehicle must be decelerated

$$Vn(t) = V_{\text{safe}}. \quad (11)$$

When the vehicle enters the bend, if the vehicle speed is greater than the safe speed, to ensure driving safety, the vehicle speed must be decelerated to the safe speed; otherwise, the vehicle may be skidded off the lane because of the excessive centrifugal force.

Step 3 (safe deceleration). Consider

$$Vn(t) = \min(Vn(t), \text{gap}_n(t)). \quad (12)$$

Step 4 (randomization deceleration with probability $P12$). Consider

$$Vn(t) = \max(Vn(t) - 1, 0). \quad (13)$$

Step 5 (location update). Consider

$$Xn(t) = Xn(t) + Vn(t). \quad (14)$$

Here $Vn(t)$ is the speed of vehicle n at time t ; $Xn(t)$ is the location of vehicle n at time t ; $\text{gap}_n(t) = Xn(t+1) - Xn(t) - L_{\text{car}}$ is the distance between vehicle n and the front vehicle at time t ; V_{safe} is the safety speed (also the maximum speed) on the bend; $P10$, $P11$, and $P12$ are the randomization deceleration probability for normal sections, transition sections, and bend sections, respectively; $P2$ is the deceleration probability when the vehicle speed is larger than the desired speed on the transition section; $P3$ is the acceleration probability when the vehicle speed is less than the desired speed on the transition section; $P4$ is the acceleration probability when the vehicle speed is less than the maximum speed on the bend section.

4. Simulation Results

In this section, we want to conclude the rule of effects of road bend conditions and vehicle characteristics on the vehicle speed and latency probability. For the model building, we give the following parameter for the simulation.

Set $L_{\text{car}} = 7$ cells, $g = 10 \text{ m/s}^2$, $P10 = 0.15$, $P11 = 0.2$, $P12 = 0.1$, $P2 = 0.3$, $P3 = 0.1$, $P4 = 0.2$, $D1 = 2$ cells, $D2 = 1$ cells, simulation steps $T_{\text{step}} = 20000$, and Bendpoint = $L/2$ (the bend is located on the middle of the lane). The average value of 20 simulation results is adopted to eliminate the impact of randomness of the results.

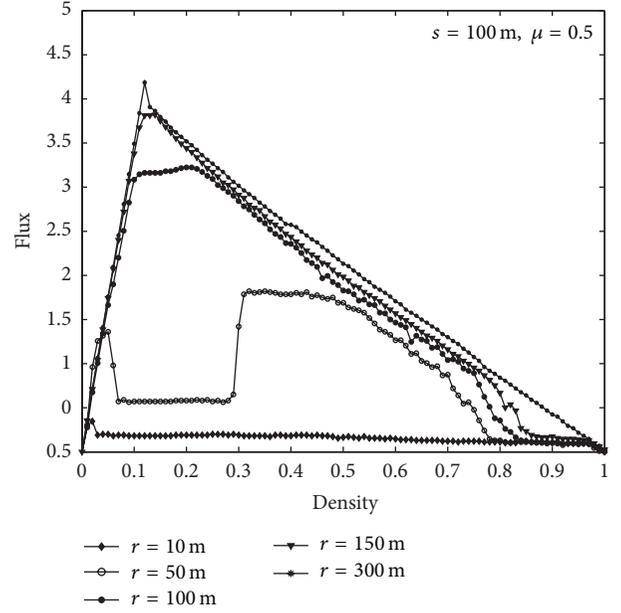


FIGURE 2: The fundamental diagram under different radii when $s = 100$, $\mu = 0.5$.

4.1. The Effect of Bend Radius on Traffic Flow. Road bend is the critical factor among all the road alignment elements, which is determined by its radius. Therefore, the effect of bend radius on traffic flow is discussed in this paper firstly, and the fundamental diagram is shown in Figure 2.

In this paper, set the curve length $s = 100$ m, the road surface friction coefficient $\mu = 0.5$, and the curve radius $r = 10$ m, 50 m, 100 m, 150 m, and 300 m. As can be seen from Figure 2, with the increase of the bend radius, the peak flow will be continuously increased, and the fundamental diagram will become more similar to that of the classic NaSch model; when the bend radius is greater than 150 m, the density-flow diagram is close to the classic NaSch model; when the radius is greater than 300 m, the density flow is consistent with the classic NaSch model; for relatively small bend radius, a “plateau” that does not change with density will appear, and the smaller the radius, the longer the “plateau” phase; when the bend radius is not so small (greater than 50 m), the density-flow diagram is close to the classic NaSch model on both sections separated by the proposed “plateau.”

Firstly, the appearance of the “plateau” is because of the inhibition effect of the road bend. The smaller the bend radius, the smaller the maximum vehicle’s speed on the bend, resulting in a blockage around the bend. The phenomenon of go-and-stop before entering the bend hinders the transformation from crowded state to free-flow, leading to a long “plateau” phase. In this case, the bend can be regarded as a “bottleneck” in the road, which is similar to toll stations or other facilities on an expressway. For small density, the traffic flow will be increased with the increase of the density; for large density, the traffic flow will be decreased with the increase of the density. For small density, the interaction between vehicles is weak, so the vehicles can remain in free

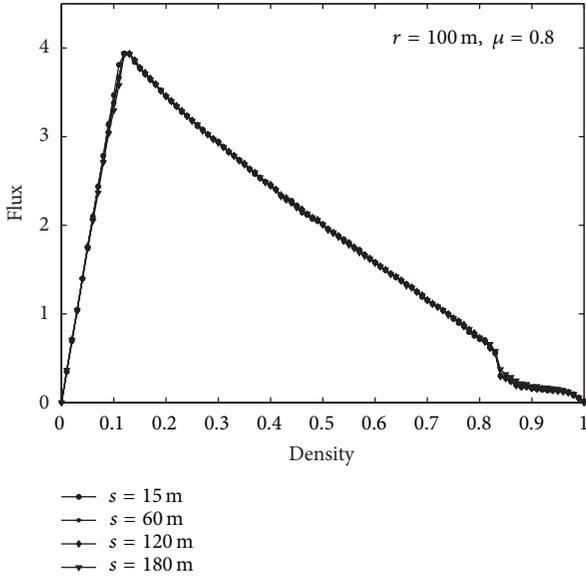


FIGURE 3: The fundamental diagram under different bend arc lengths when $r = 100\text{ m}$, $\mu = 0.8$.

state before and after passing through the bend, and the “bottleneck” effect was not so significant. For large density, the interaction between vehicles is quite strong, and the vehicles are in a congested state, so the inhibition effect of the bend is secondary compared to the interaction between the vehicles. Therefore, the main factor affecting the traffic flow is density, and the fundamental diagram exhibits a linear relationship.

Figures 3 and 4 show the density-flow fundamental diagrams under different bend arc lengths ($s = 15\text{ m}$, 60 m , 120 m , and 180 m). As can be seen from Figure 3, the difference between the fundamental diagram and the classic NaSch model is small under different bend arc lengths. For small density, the longer the length, the smaller the traffic flow; for large density, the fundamental diagrams are basically overlapped. In Figure 4, a “plateau” that does not change with density appears in the fundamental diagrams under different bend arc lengths, indicating the effect of bend arc length on traffic flow, will be smaller than that of the bend radius.

4.2. The Effect of Road Surface Friction Coefficient on Traffic Flow. Road surface friction coefficient plays an important role in highway engineering design, which not only determines the road passing capacity but also has an important influence on the vehicle driving safety. The value of friction coefficient is determined by the following factors: tire size, tire pressure, wheel load, road construction, road conditions, and the vehicle’s speed, as well as the service time of the road surface.

Figure 5 shows the density-flow fundamental diagrams under different friction coefficient ($\mu = 0.2, 0.5, 0.7, 0.8$, and 0.9). As can be seen from the figure, the fundamental diagrams under different friction coefficient exhibit the same trend with fundamental diagrams under different radii,

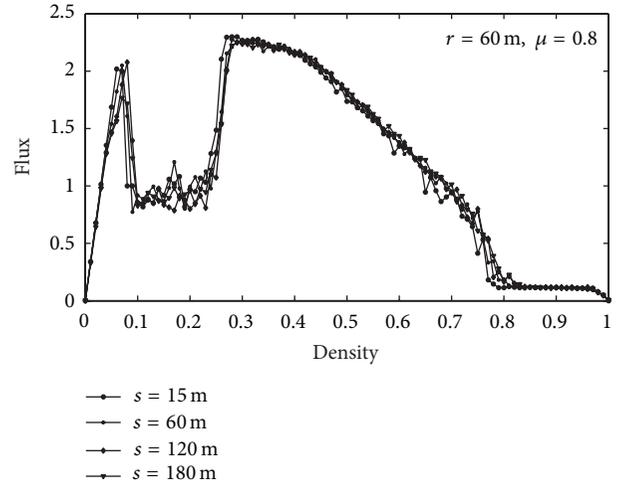


FIGURE 4: The fundamental diagram under different bend arc lengths when $r = 60\text{ m}$, $\mu = 0.8$.

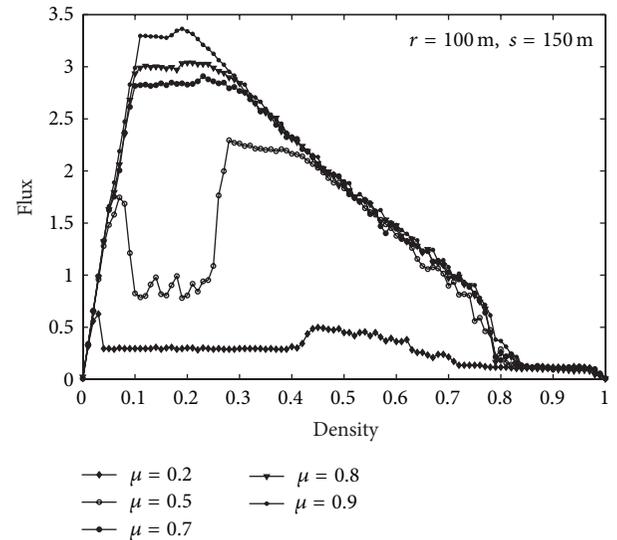


FIGURE 5: The fundamental diagram under different friction coefficients when $r = 100\text{ m}$, $s = 150\text{ m}$.

indicating that the effect of friction coefficient is consistent with the impact of radius. Under the same bend radius, the larger the friction coefficient, the greater the limit speed V_{safe} and the weaker the inhibition and “bottleneck” effect of the bend; therefore, the road passing capacity and the traffic flow are improved with the increase of the friction coefficient.

5. Conclusions

Road bend is an important part of the road. In this paper, we want to find the structure and characteristics of road bends; based on this, a single-lane CA model with bends is established accordingly. Using this model, the traffic flow is explored under different bend radii; bend arc length, and road friction coefficient. Simulation results show that a bend has

an inhibitory effect on traffic flow, in which the impact of bend radius and road surface friction coefficient is more significant than that of the bend arc length. The results can help readers to better understand the effects of road bends on traffic flow and help traffic engineers to reasonably design the bend radius.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the Humanities Social Sciences Programming Project of the Ministry of Education of China nos. 10YJA630126, 12YJC630200, 12YJC630100, the State Social Science Fund Project no. 11CJY067, and the Natural Science Foundation of Gansu Province, China, no. 1107RJYA070 and 1208RJZA164.

References

- [1] S. Wolfram, "Statistical mechanics of cellular automata," *Reviews of Modern Physics*, vol. 55, no. 3, pp. 601–644, 1983.
- [2] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," *Journal de Physique I*, vol. 2, pp. 2221–2229, 1992.
- [3] Y. Ishibashi and M. Fukui, "Traffic flow in 1D cellular automaton model including cars moving with high speed," *Journal of the Physical Society of Japan*, vol. 65, no. 6, pp. 1868–1870, 1996.
- [4] O. Biham, A. A. Middleton, and D. Levine, "Self-organization and a dynamical transition in traffic-flow models," *Physical Review A*, vol. 46, no. 10, pp. R6124–R6127, 1992.
- [5] Y. S. Qian, W. J. Li, J. W. Zeng, M. Wang, J. W. Du, and X. P. Guang, "Cellular automaton models of highway traffic flow considering lane-control and speed-control," *Communications in Theoretical Physics*, vol. 56, no. 4, pp. 785–790, 2011.
- [6] S. C. Yang, C. Deng, T. Q. Tang, and Y. S. Qian, "Electric vehicle's energy consumption of car-following models," *Nonlinear Dynamics*, vol. 71, no. 1-2, pp. 323–329, 2013.
- [7] T. Q. Tang, H. J. Huang, and S. G. Zhao, "A signal light model and its stability analysis," *International Journal of Modern Physics B*, vol. 24, no. 28, pp. 5613–5623, 2010.
- [8] T. Q. Tang, P. Li, Y. H. Wu, and H. J. Huang, "A macro model for traffic flow with consideration of static bottleneck," *Communications in Theoretical Physics*, vol. 58, no. 2, pp. 300–306, 2012.
- [9] T. Q. Tang, Y. Li, and H. J. Huang, "The effects of taxi on traffic flow," *International Journal of Modern Physics C*, vol. 20, no. 10, pp. 1537–1546, 2009.
- [10] T. Q. Tang, H. J. Huang, and H. Y. Shang, "A dynamic model for the heterogeneous traffic flow consisting of car, bicycle and pedestrian," *International Journal of Modern Physics C*, vol. 21, no. 2, pp. 159–176, 2010.
- [11] H. X. Ge, H. B. Zhu, and S. Q. Dai, "Cellular automaton traffic flow model considering intelligent transportation system," *Acta Physica Sinica*, vol. 54, no. 10, pp. 4621–4626, 2005.
- [12] H. X. Ge and R. J. Cheng, "The "backward looking" effect in the lattice hydrodynamic model," *Physics A*, vol. 387, no. 28, pp. 6952–6958, 2008.
- [13] H. L. Zhou, Z. Y. Gao, and K. P. Li, "Cellular automaton model for moving-like block system and study of train's delay propagation," *Acta Physica Sinica*, vol. 55, no. 4, pp. 1706–1710, 2006.
- [14] Y. P. Fu, Z. Y. Gao, and K. P. Li, "The characteristic analysis of the traffic flow of trains in speed-limited section for fixed-block system," *Acta Physica Sinica*, vol. 56, no. 9, pp. 5165–5171, 2007.
- [15] K. P. Li, Z. Y. Gao, and B. Ning, "Modeling fixed-block railway signaling system using cellular automata model," *International Journal of Modern Physics C*, vol. 16, no. 11, pp. 1793–1801, 2005.
- [16] K. P. Li, Z. Y. Gao, and L. X. Yang, "Modeling and simulation for train control system using cellular automata," *Science in China E*, vol. 50, no. 6, pp. 765–773, 2007.
- [17] S. K. Rao, "Unit curve for design of highway transitions," *Journal of Transportation Engineering*, vol. 121, no. 2, pp. 169–175, 1995.
- [18] J. B. Bentley and W. E. Gallagher, "Review of highway curve design," *Journal of the Institution of Highway Engineers*, vol. 18, no. 11, pp. 7–16, 1971.
- [19] E. Dabbour, S. Easa, and K. Raahemifar, "Optimum vertical curves for highway profiles using nonlinear optimization," in *Proceedings of the Annual Conference: Canadian Society for Civil Engineering*, pp. 2833–2841, 2008.
- [20] B. Persaud, R. A. Retting, and C. Lyon, "Guidelines for identification of hazardous highway curves," *Transportation Research Record*, no. 1717, pp. 14–18, 2000.
- [21] H. Mu and S. W. Yang, "Alignment design of curved section of highway," *Journal of Chang'an University (Natural Science Edition)*, vol. 27, pp. 26–29, 2007.
- [22] Y. J. Liang and Y. Xue, "Study on traffic flow affected by the road turning," *Acta Physica Sinica*, vol. 59, no. 8, pp. 5325–5331, 2010.
- [23] M. Wang, J. W. Zeng, Y. S. Qian, and W. J. Li, "Properties of train traffic flow in a moving block system," *Chinese Physics B*, vol. 21, no. 7, Article ID 070502, 2012.

Research Article

Multiagent-Based Simulation of Temporal-Spatial Characteristics of Activity-Travel Patterns Using Interactive Reinforcement Learning

Min Yang,¹ Yingxiang Yang,² Wei Wang,¹ Haoyang Ding,¹ and Jian Chen¹

¹ School of Transportation, Southeast University, Si Pai Lou No. 2, Nanjing 210096, China

² Department of Civil and Environment Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA

Correspondence should be addressed to Min Yang; transtar2002@163.com

Received 6 September 2013; Revised 6 November 2013; Accepted 9 December 2013; Published 30 January 2014

Academic Editor: Geert Wets

Copyright © 2014 Min Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a multiagent-based reinforcement learning algorithm, in which the interactions between travelers and the environment are considered to simulate temporal-spatial characteristics of activity-travel patterns in a city. Road congestion degree is added to the reinforcement learning algorithm as a medium that passes the influence of one traveler's decision to others. Meanwhile, the agents used in the algorithm are initialized from typical activity patterns extracted from the travel survey diary data of Shangyu city in China. In the simulation, both macroscopic activity-travel characteristics such as traffic flow spatial-temporal distribution and microscopic characteristics such as activity-travel schedules of each agent are obtained. Comparing the simulation results with the survey data, we find that deviation of the peak-hour traffic flow is less than 5%, while the correlation of the simulated versus survey location choice distribution is over 0.9.

1. Introduction

Over the few last decades, activity-based approaches has become the main theme in transportation demand modeling, taking the place of trip-based approaches. Trip-based approach has several drawbacks: trip generation is fixed and independent of the transportation system; travel demand is generated from the need of activity participation; and the space and temporal relationship of all trips and activity patterns is ignored. Such drawbacks brought activity-based approach into transportation demand modeling.

The first activity-based approaches began in the 1970s [1–3]. Those pioneering studies explored choices and constraints in travel demand. Since that time, activity-based modeling has flourished. Various methodologies have been introduced and they can be classified into three categories.

The first category is utility-maximizing model (or econometric model) which suggests that individuals seek to maximize their cumulative utilities when performing activities. Those models link individual or household's sociodemographics, transportation policies, and other environmental

factors to their activity and travel patterns. Econometric models ranging from discrete choice models (such as multinomial logit and nested logit mode) to hazard duration models remain to be a powerful approach in activity-travel analysis [4–7].

The second category is computational process model (CPM) which focuses on using context-dependent choice heuristics to model individual's decision process. A computational process model is a set of condition-action rules that specify how a decision is made. One precursor in CPM is the time-space prism method. Hägerstrand [3] introduced the three-dimensional space-time models. In such models limited resources of time and space became constraints on each individual's behavior alternatives [8]. The techniques used in more recent studies include decision trees, neural networks, and Bayesian networks [9–11].

The combination of the above two approaches leads to hybrid models. Hybrid models concentrate on the integration of econometric models and CPM. Decision-tree is combined with parametric modeling [12]; random utility maximization

is incorporated into activity scheduling model [13]. New algorithms such as reinforcement learning are also introduced into the field.

Reinforcement learning integrates the concepts of reward (utility) maximization and context-dependent choice heuristics. The applications of reinforcement learning include robotics, game theory, dispatching system, and financial trading [14–17]. Tan used reinforcement learning to formalize an automated process for determining stock cycles by tuning the momentum and the average periods. The total experimental results from the five stocks are able to beat the market by about 50 percentage points [18]. Lahkar and Seymour studied reinforcement learning in a population game. Agents in a population game revise mixed strategies using the cross rule of reinforcement learning [17]. In addition, formulation of economic dispatch as a multistage decision making problem is carried out using reinforcement learning by Jasmin et al. [15]. Applying reinforcement learning in transportation demand modeling has several advantages. First, the imitation of human learning through trial and error interactions with a dynamic environment helps to explain behavioral mechanisms [19]. The RL mechanism is distinguished from other computational cognitive mechanisms by its emphasis on learning by an individual from direct interaction with individual's decision environment in the presence of an explicit goal and feedback and without relying on any exemplary supervision. Secondly, it does not need an expert-system to inform it what selection is right and what is wrong. Thirdly, it could react to unforeseen events and take both long-term learning and short-term dynamics into account. Among the first attempts, Charypar and Nagel built the basic model of activity time plans using q-learning and got quite realistic results [20]. This model was then modified to allocate both time and location choice of activity-travel pattern [21]. Because q-learning generally takes a long time to converge and the curse of dimensionality occurs when the problem gets complex, q-learning was combined with the regression tree method to form a new algorithm called q-tree [22].

The above-mentioned researches show several aspects that need further development.

- (i) In most of the reinforcement-learning-based studies, though the format of reward function has been scrutinized, the rewards are based on assumption values and are hard to be acquired from survey data, so that the result is hard to be put into practical use.
- (ii) In many of the multiagent systems, “multi” means several components of the system such as road, intersection, and traveler rather than multiple travelers. Interactions of travelers are neglected.
- (iii) The result analysis is often limited within individual activity-travel schedule. Macroscopic characteristics such as traffic flow distribution are often ignored.

In this study we propose an interactive reinforcement learning algorithm in which individuals not only receive information from the environment, but also give feedback to the environment. We did this by adding road congestion degree, which is determined by travelers' decisions, to the

algorithm. The dynamic environment is a medium that passes the influence of one traveler's decision to others. The self-organization effect shown through this mechanism makes the system reach a dynamic equilibrium. This algorithm not only ensures rationality of each single traveler's behavior, but also obtains aggregated temporal-spatial traffic features such as traffic flow distribution and the distribution of activity locations. We also seek a compromise between the well-established theoretical reward function form and the quality of data we could truly get from practical surveys. The simplified reward function makes the algorithm immediately applicable.

The rest of this paper is organized as follows. Section 2 introduces the algorithm of modified multiagent-based q-learning. Section 3 is devoted to the analysis and calculation of the survey data. Section 4 shows the temporal-spatial simulation results of Shangyu city's traffic system. Section 5 concludes the findings of this paper and discusses future research directions.

2. Multiagent-Based Q-Learning Method

2.1. Reinforcement Learning. Multiagent system focuses on the analysis of several agents' dynamic and complex collective behavior. Because multiagent system has no global control and each agent may get incomplete information, the system must learn repetitively to improve the performance. Reinforcement learning is a major method of this kind. Kaelbling et al. [19] define reinforcement learning as the problem faced by an agent that must learn behavior through trial and error interactions in a dynamic environment. Moreover, the consequences of actions change over time and depend on the current and future state of the environment. Reinforcement learning has the potential to deal with this uncertainty through continuous observations of the environment and through consideration of indirect and delayed effects of actions.

Basic concepts concerning reinforcement learning include the following.

- (i) Agent: in this paper, an agent means a traveler.
- (ii) State: a vector (activity, start time, duration, location, and congestion degree) denotes an agent's state. The vector is denoted as $(a, s, d, l, \text{ and } vc)$ for brief.
- (iii) Location: the unit of location is traffic zone which is an area that has multifunctions including leisure, shopping, and working.
- (iv) Activity: activities include home, work, maintenance, and leisure.
- (v) Action: there are 4 actions, staying at current activity or move to one of the other 3 activities. The same as the way activities are represented; actions are denoted as $h, w, s, \text{ and } l$ for brief.
- (vi) Duration and start time: time variables should be discrete in q-learning. The unit of time slot is 15 min, which divides a day into 96 slots. 24 pm is connected with 0 am. Because the number of state should be

finite, the longest duration of an activity is limited to 24 hours. Hence, both duration and start time could be represented as a number from 1 to 96.

- (vii) Policy: it means how an agent's action may bring it from one state to another.
- (viii) Reward function: it is defined as the immediate feedback an action brings.
- (ix) Value function: it shows the total feedback an action may bring both immediately and afterward.
- (x) Q-value: the Q-value of an action a , given a state s , denotes the expected utility of an agent taking action a in state s .
- (xi) Congestion degree: in order to show influence of agents' behavior on the environment, a variable of congestion degree vc is added to the state of environment. Because environment is represented by discrete variables in q-learning, vc should also be a discrete variable. It is defined as $vc = \text{ceil}(5 \cdot (v/c))$. MATLAB function $\text{ceil}()$ rounds $r(s_t, a_t)$ to the nearest integer towards infinity. V is the traffic volume of a given OD pair and C is the capacity of the OD pair. Because the capacity is hard to be measured directly, we assume that each OD pair's capacity is 4000 considering the total population and the size of Shanguyu city.

Reinforcement learning tasks are generally treated in discrete time steps. At each time step t , the agent observes the current state s_t and chooses a possible action a_t to perform, which leads to its succeeding state $s_{t+1} = \delta(s_t, a_t)$. The environment responds by giving the agent a reward $r(s_t, a_t)$. These rewards can be positive, zero, or negative. It is probable that these preferable rewards come with a delay. In other words, some actions and their consequential state transitions may bring low rewards in short-term, while it will lead to state-action pairs later with a much higher reward.

For this reason, the task of the agent is to learn a policy π according to the state S and the action A to receive the maximal accumulative rewards. Given a random policy π from a random state s_t , the accumulative reward can be formulated as follows:

$$V^\pi(s_t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots = \sum_{i=0}^{\infty} \gamma^i r_{t+i}, \quad (1)$$

where r_{t+i} represents the scalar reward received i steps in the future and γ is the discounting factor. The agent only receives the immediate reward if γ is set to zero.

2.2. Q-Learning Algorithm. The agent needs to learn the optimal policy $\pi^*(s)$ that maximizes the accumulative reward. Unfortunately, it is required that the knowledge of immediate reward function r and state transition function δ are known in advance. In reality, however, it is usually impossible for the agent to predict in advance the exact outcome of applying a random action to a random state. In other words, the domain knowledge is probably not perfect. q-learning is then devised to select optimal actions even when the agent has no knowledge about the reward and state functions.

We define \widehat{Q} as the estimation of true Q-value. The q-learning algorithm maintains a large table with entries to each state-action pair. When it starts, the value of $\widehat{Q}(s, a)$ is initially filled with random numbers. The agent repeatedly observes its current state s , chooses a possible action a to perform, and determines its immediate reward $r(s, a)$ and resulting new state $\delta(s, a)$. The $\widehat{Q}(s, a)$ value is then updated according to the following rule:

$$\widehat{Q}(s, a) \leftarrow r(s, a) + \gamma \max_{a'} \widehat{Q}(s', a'). \quad (2)$$

That is to say, the \widehat{Q} -value of the current state-action pair is refined based on its immediate reward and the \widehat{Q} -value of its next state. The agent can reach a globally optimal solution by repeatedly selecting the action that maximizes the local values of Q for the current state.

This is only a brief introduction of q-learning and detailed introduction could be found in reference [20]. The process can be described as follows:

- (1) initialize the Q-values,
- (2) select a random starting state s which has at least one possible action to select from,
- (3) select one of the possible actions. This action leads to the next state,
- (4) update the Q-value of the state-action pair according to the update rule above,
- (5) go back to Step 3 if the new state has at least one possible action, if not, go to Step 2.

2.3. Reward Function. Previous researchers in this domain constructed their reward functions based on activity start time, duration, length of travel, and so on [20, 22]. This method is adopted by us and our reward function contains the following parts.

2.3.1. Reward Based on Attraction Degree of Zones. In this paper a location is a zone that has multiple land use functions. In reality, people sometimes prefer to travel for a long time downtown to go shopping because the land use characteristics make downtown more attractive. To quantify this, the reward based on attraction degree of zones is added to the reward function. It is only for maintenance and leisure activities because home and work have fixed locations. We assume the more maintenance activities are conducted in a zone, the higher attraction degree this zone has. This also applies for leisure activities.

Consider

$$\text{attract}_{i,j} = \frac{n_{i,j} - n_{i,\text{avg}}}{n_{i,\text{max}} - n_{i,\text{avg}}}, \quad (3)$$

where $n_{i,j}$ is the number of leisure activities or maintenance activities conducted in zone j , i is the activity type, n_{avg} is the average leisure or maintenance activities conducted among all zones, and n_{max} is the maximum leisure or maintenance activities conducted among all zones. The reward is $r_{\text{attract}(i,j)} = 50 * \text{attract}_{i,j}$.

2.3.2. Reward Based on Activity Duration. When an agent conducts an activity and the duration is within a reasonable range, it should get a fairly large accumulative reward. When the duration is less than the expected value, the marginal benefit is positive, while if the duration is more than the expected value, the marginal benefit is negative.

Consider

$$r_{\text{duration}(i)} = \begin{cases} 100 + \frac{d_{\min(i)} - d}{d_{\min(i)}}, & (d < d_{\min(i)}), \\ 50, & (d_{\min(i)} \leq d < d_{\text{avg}(i)}), \\ -50, & (d_{\text{avg}(i)} < d \leq d_{\max(i)}), \\ -100, & (d > d_{\max(i)}), \end{cases} \quad (4)$$

where $d_{\min(i)}$, $d_{\text{avg}(i)}$, and $d_{\max(i)}$ represent the reasonable minimum, maximum, and average duration of activity i . They are, respectively, the 5%, 50%, and 95% percentile duration of activity i in the survey data.

2.3.3. Reward Based on Activity Start Time. Each activity's start time distribution is calculated using the survey data. To make the distribution curve more smooth in order to diminish the effect of randomness, we use polynomial functions (use C to denote) to fit the curve. Then function C is normalized.

Consider

$$r_{\text{start time}(i)} = C_i(s), \quad (5)$$

where i represents the type of activity, while s is the start time of the activity. The range of s is (1, 96).

2.3.4. Reward Based on Travel Time. Some scholars define travel-time-based reward as $r_{\text{travel}} = -c * (bt)^\alpha$ [23]. This form is adopted by us, but it needs some modifications because the influence of congestion degree is taken into account. t is no longer a fixed value decided by the length between zones, but it relates to the congestion degree of the OD pair. We use the widely accepted impedance function in China [21]:

$$U = \begin{cases} U_0 \left(1 - \frac{0.6V}{C}\right), & \frac{V}{C} < 0.9, \\ \frac{U_0}{7.4V/C}, & \frac{V}{C} \geq 0.9, \end{cases} \quad (6)$$

where U is the actual speed, while U_0 is free flow speed. t_0 is the free flow travel time. Actual travel time t could be defined as $t = t_0 * U_0/U$.

2.4. Flow-Chart of Calculation. When q-learning is applied in this paper, the process described below could be shown in Figure 1. The whole process is separated into 3 steps.

Step 1 is to utilize travel diary survey data to extract typical activity patterns and form different kinds of agents according to their activity patterns. Also utilizing the survey data, the reward function for different kinds of agents is calculated.

Step 2 is to estimate the value function (in this algorithm: Q-values) through trial and error until the Q-value matrix converges.

Step 3 is to add agents on the network and then use the Q-values to decide the activity-travel schedule of each agent. In the end, temporal-spatial characteristics of the simulation result and each agent's activity-travel schedule are calculated and recorded.

Taking congestion degree into account could enable interactions among agents and let them cooperate and compete in the environment. In the simulation of a network, a number of agents are set on the network and their states are initialized. Then at each time step, agents decide their actions by choosing an action that brings maximum Q-value one by one according to the congestion degree and other aspects of the environment. Their actions would in turn influence the congestion degree therefore would influence other agents' actions. In this way, all agents' activity-travel schedules could be decided.

3. Data Analysis and Process

3.1. Data Survey. We utilized the travel diary survey data from Shangyu city conducted in 2006. The survey includes individual/household sociodemographics and travel records. Travel records include trip starting and ending times, origin and destination, mode used, and trip purpose. Trip purpose is divided into nine categories, including work, school, official business, shopping, socializing-recreation, serving passengers, personal business, returning home, and returning to work. Among these purposes, work, school, and official business are named commute activity or simply work. Shopping, serving passengers, and personal business are called maintenance activities. Socializing-recreation is called leisure activity. Maintenance and leisure activities generally are named none-working activities. Hence, the 9 categories of activities could be divided into 4 types: work, maintenance activity, leisure activity, and staying at home.

Shangyu city has a population of 204,900. 4,101 residents from 1,564 households are surveyed. After deleting the incorrect statistics, data from 3,368 people are used, representing 82.1% of the people surveyed. 486 students account for 14.4% of the valid data. Because students' activity-travel schedules are rather fixed and the main focus of this paper is on working and none-working groups, the students' data are not considered. Thus, the data obtained from the remaining 2,883 people, accounting for 85.6% of all the valid data, are used for the analysis.

3.2. Typical Activity Patterns. The first step of processing valid survey data is to extract typical activity patterns. A tour is defined as the travel from home to one or more activity locations and back to home again [4]. An activity pattern here is defined as all tours an individual conducted in a single day. In the valid data, 10 of the patterns are shared by more than 20 samples. We call these 10 activity patterns typical activity patterns and the description of them can be seen in Table 1. They take up 2397 of the 2882 valid samples. Agents could be

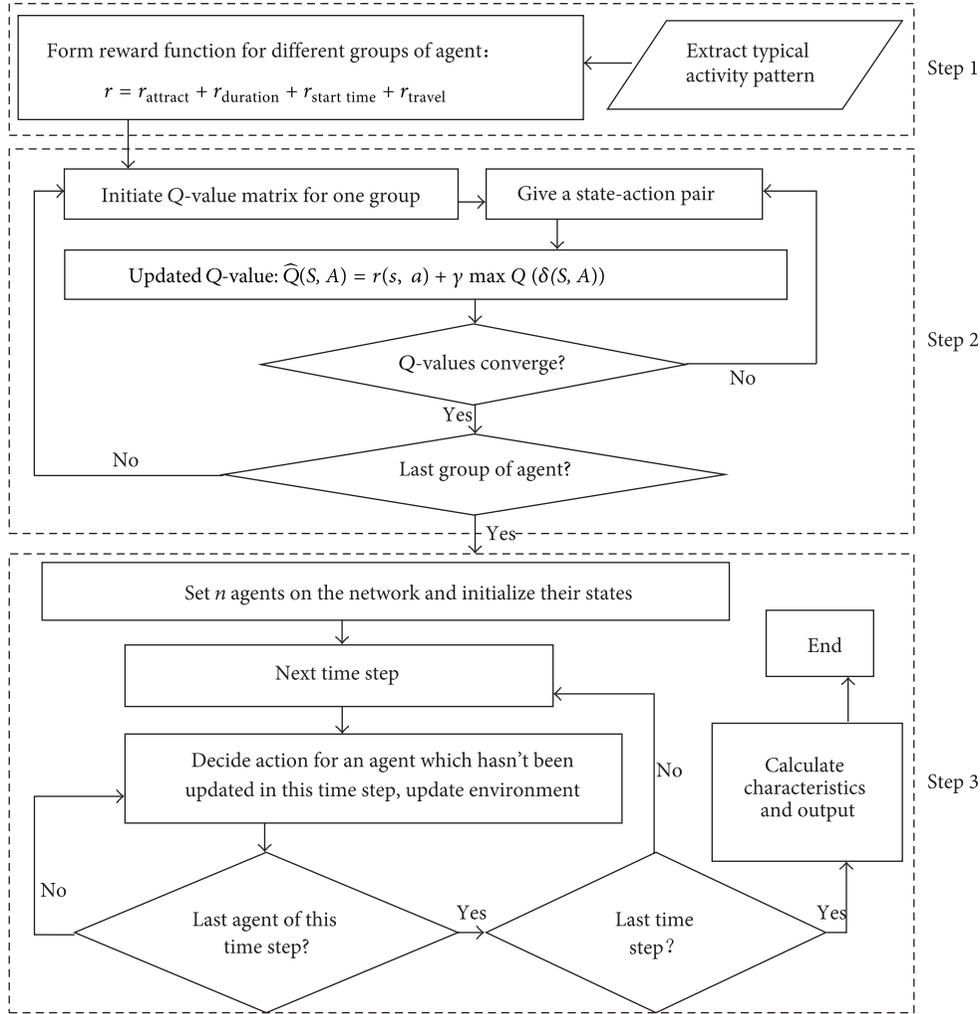


FIGURE 1: Three Steps of multiagent-based q-learning simulation.

classified according to their activity patterns. We take these 10 typical patterns to form 10 types of agents. Patterns which include working activity are called commuting patterns, and others are called none-working patterns. The characteristics of these 10 patterns are described as in Table 1 (the 4 activities are written as h, w, s, and l for brief).

3.3. Reward Function Calculation. The reward function has been constructed in Section 2.3. The paragraphs below show the values of parameters used in the reward function, calculated from the survey data. Furthermore, ten different types of agents have their own parameters, respectively, though the functional forms are the same.

3.3.1. Attraction Degree of Zone. The attraction degrees of zones are listed in Figure 2, next to it is the traffic zone division of Shangyu city. Because these degrees are decided by land use characteristics of different zones, to different groups of people the attraction degrees are the same.

It is quite clear that zones 2, 8, and 13 are the center of leisure activity, while zones 3 and 5 are the center of

maintenance activity. This result corresponds to the land use characters of Shangyu because these zones are in the center of Shangyu.

3.3.2. Reward Based on Duration. To make the results more realistic, we calculate the rewards based on duration of 10 typical activities patterns according to the definition in Section 2.3. The unit of these parameters is 15 min. The relatively small value of standard deviation shows that people who belong to the same group share much similarity in behavior, at least in the duration of activity.

Where are the statistics?

3.3.3. Reward Based on Activity Start Time. The process of calculating this reward has been stated in Section 2.3. Use polyfit function in MATLAB to fit every activity's start time distribution of each group into smooth curves.

In Figure 3 min is not 0.

The start time-duration-reward graphs of the four activities are shown in Figure 3.

TABLE 1: Description of typical activity patterns.

Activity pattern	Number of samples	Ratio	Description
hwh	1069	37.1%	Simple work pattern with only primary tour
hwsh	26	1.0%	Having other stops when getting off work, with only primary tour
hwhsh	44	1.5%	With a secondary tour, primary tour is simple work pattern
hwhwh	568	19.7%	Work tour with home-based subtour
hswsh	26	1.0%	With a subtour during work
hsh	357	12.4%	Simple maintenance tour
hlh	111	3.9%	Simple leisure tour
hshlh	32	1.1%	With both maintenance and leisure tours, the prior one is a maintenance tour
hshsh	105	3.6%	With two maintenance tours
hlhsh	59	2.0%	With both maintenance and leisure tours, the prior one is a leisure tour

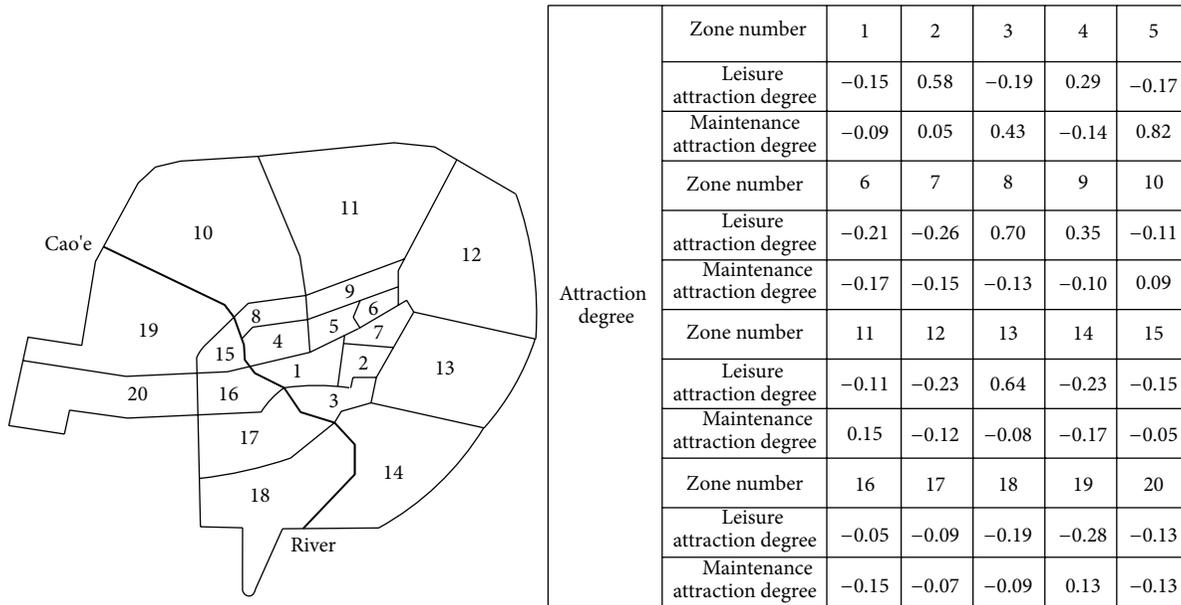


FIGURE 2: Attraction degrees of zones and traffic zone division of Shangyu city.

4. Simulate Temporal-Spatial Features of Multipleagents

4.1. *Assumptions and Preparations for the Simulation.* To simulate traffic conditions in Shangyu, the first step is to expand the number of agents from the size of the sample to the proportion of population these types of agent take up in Shangyu. By calculation, the 2397 samples in the survey should be expanded to a population of 145684 people. Apart from the already existed data of 2397 people, we need to establish 143287 people's attribute data. Each people's attributes include activity pattern and home and work locations (if this person works). To make the distribution of each attribute in the newly established data the same as the survey data, the procedure of establishing one person's attributes could be as follows.

(1) Randomly generate a natural number from 1 to 2397. The activity pattern of this people will equal to that of the number i people in the survey data.

(2) Likewise, the attribute of home and work locations can be decided by randomly choosing one from the 2397 survey samples.

Initialize each agent's state and simulate 1000 time steps. We take the last 96 time step to analyze. Both each individual agent's activity-travel schedule and spatial-temporal characteristics are analyzed.

4.2. *Simulation Results of Activity-Travel Schedules.* Each agent's activity-travel schedule in one day is recorded. We randomly choose one agent from each pattern and show his/her activity-travel schedule in a day (from 0 am to 24 pm). The

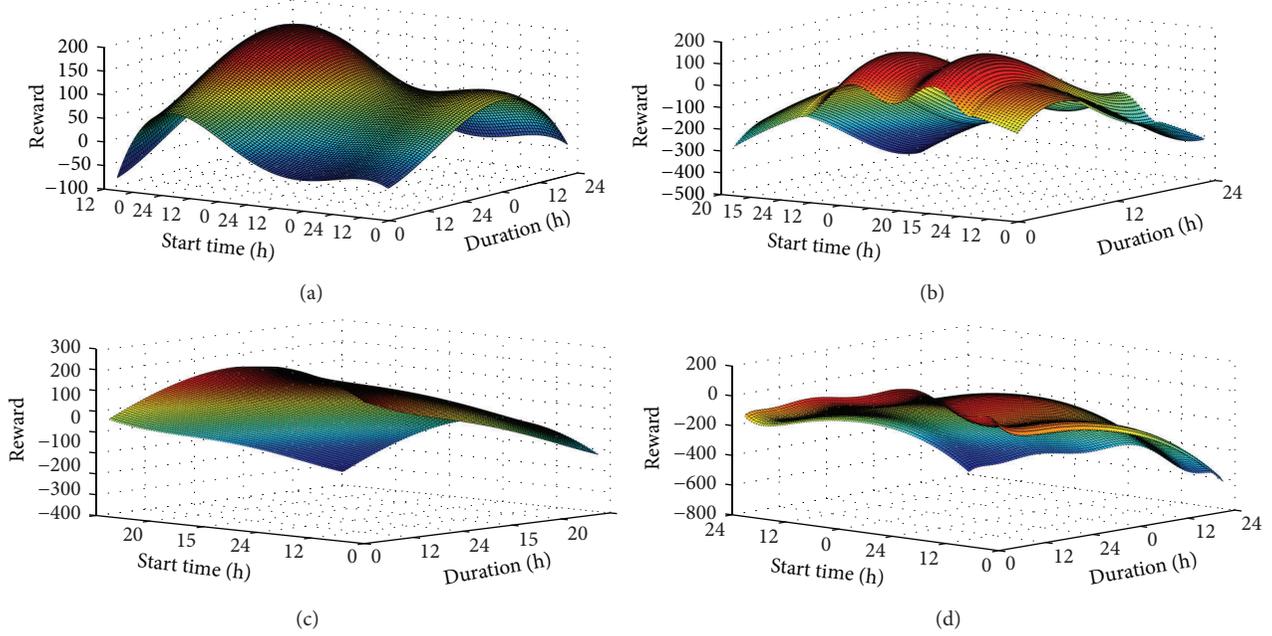


FIGURE 3: (a) Home reward function. (b) Working reward function. (c) Leisure reward function. (d) Shopping reward function.

TABLE 2: Activity-travel schedules.

Agent type	Activity-travel schedule
hwh	h (00:00–08:00, 05) w (08:30–16:45, 08) h (17:30–24:00, 05)
hwsh	h (00:00–07:00, 12) w (07:30–17:00, 11) s (17:15–18:15, 11) h (18:45–24:00, 12)
hwhsh	h (00:00–07:45, 01) w (08:15–15:30, 02) h (15:45–19:00, 01) s (19:15–19:45, 03) h (20:00–24:00, 01)
hwhwh	h (00:00–06:30, 11) w (07:30–11:15, 07) h (12:00–13:00, 11) w (13:45–17:30, 07) h (18:30–24:00, 11)
hwswh	h (00:00–07:15, 04) w (07:45–11:45, 06) s (12:15–12:45, 05) w (13:15–16:30, 06) h (17:15–24:00, 04)
hlh	h (00:00–05:45, 03) l (06:00–07:00, 02) h (07:30–24:00, 03)
hlhsh	h (00:00–04:45, 02) l (05:15–06:15, 08) h (06:45–07:00, 02) s (07:30–08:00, 03) h (08:30–24:00, 02)
hsh	h (00:00–06:30, 15) s (07:00–07:30, 05) h (08:00–24:00, 15)
hshlh	h (00:00–06:00, 04) s (06:30–06:45, 08) h (07:15–07:30, 04) l (08:00–08:30, 08) h (09:00–24:00, 04)
hshsh	h (00:00–05:15, 17) s (06:30–07:00, 03) h (08:00–16:30, 17) s (17:15–17:45, 03) h (18:30–24:00, 17)

result is shown in Table 2. The first part in each parenthesis is activity time and the second part is activity location. The table shows that no abnormal sequence, such as staying at

one activity for too long or conducting activities in improper time, occurs in these 10 examples. One flaw is that to avoid the morning peak of commute agents, the none-working agents’ trips are generally a little bit earlier than the peak shown by the survey.

Having activity-travel schedules of all agents, we could move our analysis further to macroscopic temporal-spatial characteristics of the traffic.

4.3. Temporal Characteristics of the Simulation Result. The traffic flow distributions of this paper’s algorithm and the traditional algorithm which has not taken interactions between agents are compared in Figure 4. Both methods show apparent morning and evening peak. But in the traditional method environment is static, which means one agent’s action will not affect other agents’ choices; it is natural that agents of the same attributes all do the same activity at the same time and zone. Therefore, the traditional method’s distribution of flow is ladder-like, which means peak hour flow is very large.

By comparison, because the congestion degree is taken into account, in this paper’s method, some agents avoid traveling in the rush hour because it will lead to lower rewards. As a result, the peak hour flow is much lower. Even agents of the same attributes would have different activity-travel schedules because the environment is dynamic. Thus, the behavior of the whole population is not isolated but has interactions.

Traffic flow distributions of the 2397 samples’ survey result and their corresponding agents’ simulation result are shown in Figure 5. The simulation result matches the survey result well. Their peak hour flow deviation is less than 5%.

To show the features of different patterns’ traffic flow distribution, we could mark different traffic patterns’ flows with different colors as is shown in Figures 6(a) and 6(b).

TABLE 3: Comparison of two methods' PHR values.

OD pair	Morning peak		Evening peak	
	Traditional method	New method	Traditional method	New method
(3, 11)	32.25%	17.93%	27.79%	11.42%
(5, 16)	24.43%	11.34%	31.64%	15.03%
(4, 17)	32.73%	21.52%	46.19%	20.49%
(5, 11)	25.50%	8.65%	28.61%	15.01%
(9, 10)	22.18%	16.79%	18.77%	10.12%
(9, 5)	46.23%	18.11%	17.98%	12.38%

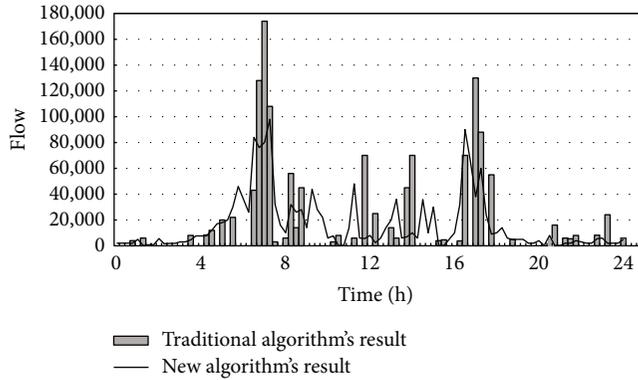


FIGURE 4: Comparison of traffic flow distribution between new method and survey data (flipped with Figure 5).

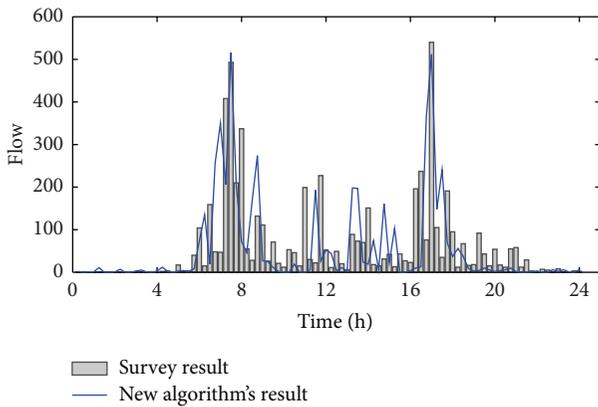


FIGURE 5: Comparison of traffic flow distribution between traditional method and new one.

Figure 6(a) shows the flow distribution of the 5 commute patterns. It shows clear morning and evening peaks, at about 7 am and 6 pm, respectively. Compared with Figure 3 we could find out that these commuting patterns, especially pattern hwh and hwhwh, account for a large percentage of morning and evening peaks' flow. The peak at noon is caused by pattern hwhwh agents who go home at noon. On the whole, pattern hwh and hwhwh are the determinants of commuting patterns' flow distribution, and other commuting patterns have too few people to influence the trend.

Figure 6(b) shows none-working agents' traffic flow distribution, which is totally different from commuting agents': there is no such dominant pattern. On the contrary, all none-working patterns contribute to the formation of figure's shape. Two peaks of the flow are all in the morning, at about 5 am and 9 am, respectively. The survey result shows that 42.9% of the none-working groups are retired people in Shangyu. In China the elderly usually like to go out to do some exercises early in the morning and food markets usually open very early; this explains why both the survey result and the simulation result show that none-working people's travel peak is in the morning. Agents of pattern hlh tend to go out early at 5 am, while the flow of pattern hsh almost distributes evenly from 5 to 10. In China, most people tend to stay at home in the evening, especially the none-working people so there is not much traffic in the evening as in Figure 6(b).

Table 3 shows two methods' comparison of peak hour ratio (PHR). Peak hour ratio is defined as the ratio of peak hour flow and the traffic flow of a whole day. In China, the measured PHR is often between 10% and 15%. Because there are too many OD pairs, the table listed the results of 6 OD pairs which have the largest traffic flow as representative. For all OD pairs, the original method's average PHR is 30.5% and the result of the new method is 16.2%. It is clear that the latter is closer to reality.

4.4. Spatial Characteristics of the Simulation Result. In the traditional method, because congestion degree is not taken into account and attraction degree's effect is quite distinct, all agents conduct their maintenance and leisure activities at the zone that has maximum attraction degree: all the 20094 leisure activities are conducted in zone 8, while all the 70433 maintenance activities are conducted in zone 5. We need to mention that because Shangyu is a small city and the distances between zones are not very long; the influence of distances between OD pairs is subtle. After taking into account congestion degree, the choice of location is much more dispersed. Agents would choose to conduct their activities in other zones which have lower attraction degrees when center zones are crowded. Finally, 5845 leisure activities are conducted in zone 8, which accounts for 29.0% of all leisure activities. 22392 maintenance activities are conducted in zone 5, accounting for 31.7% of all maintenance activities.

The choice of activity zones is shown in Figure 7. The survey data's activity location distribution is calculated and then it is extended the same proportion that the samples are

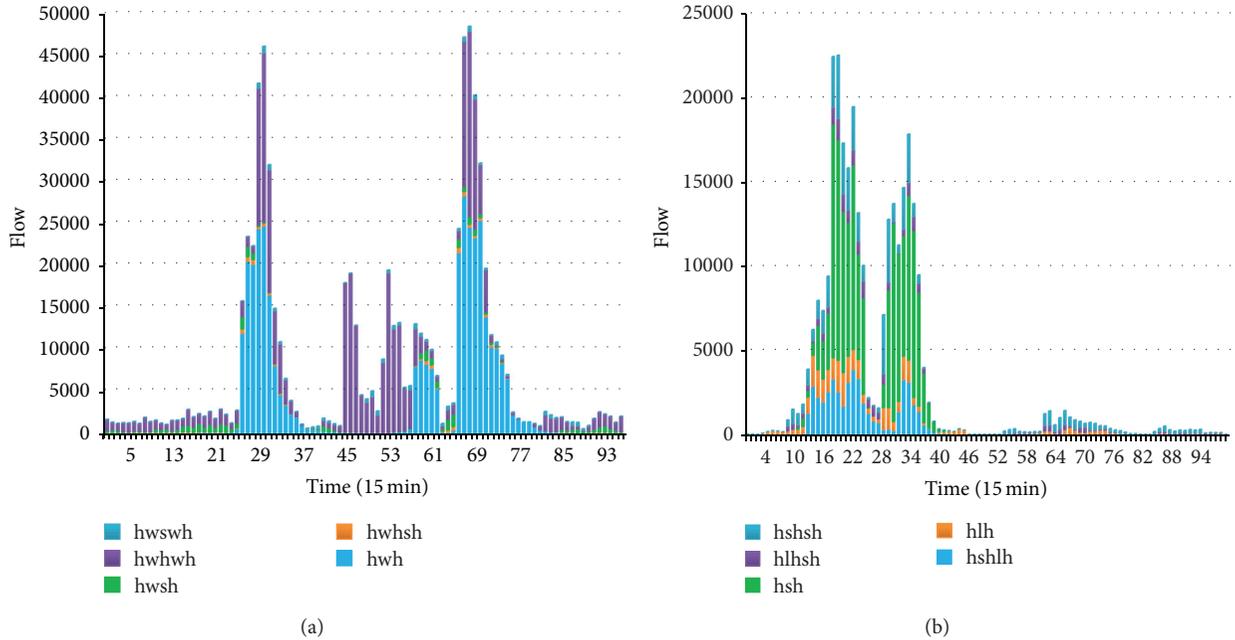


FIGURE 6: (a) Commute agents' traffic flow distribution. (b) None-working agents' traffic flow distribution.

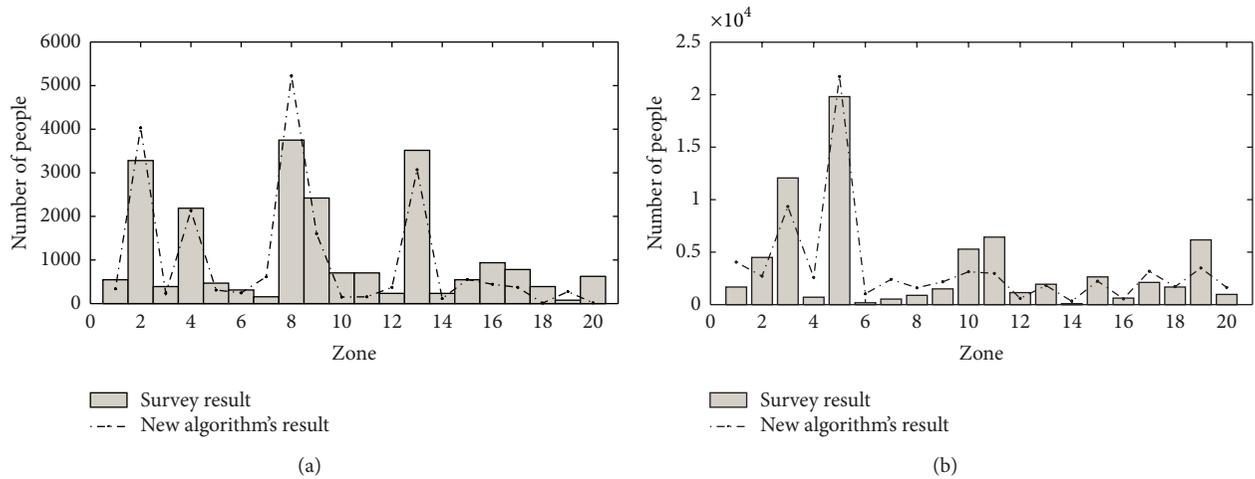


FIGURE 7: (a) Choice leisure activity location. (b) Choice of maintenance activity location.

extended to show how the 145684 people's choice of location would be like according to the survey data.

It is compared with the simulation result and the figure shows that the simulation result is quite close to the extended survey result. The correlation coefficient between the survey data's leisure activity location distribution and the simulation result is 0.921. And the correlation coefficient between survey data's and the simulation result's maintenance location distribution is 0.902.

5. Conclusions and Future Directions

In this paper we use a modified multiagent-based reinforcement learning algorithm to simulate the traffic condition of Shangyu city. Both the spatial-temporal features of the

entire population and the activity-travel schedule of single individuals are analyzed. The main findings are listed as follows.

- (i) This paper's method takes the congestion degree between OD pairs into account, which enables agents' actions to influence the environment. Thus, agents' actions have interactions with each other. Because of this interaction, both the spatial-temporal features of the entire population and the activity-travel schedule of single agent are close to actual situations.
- (ii) Because in this paper agents are no longer separated individuals but an integrity that interacts with each other, the spatial-temporal features of the whole population, such as traffic flow distribution, PHR factor,

and location choice distribution, could be calculated, which is rarely seen in previous research in this field.

- (iii) Survey data are utilized throughout the whole process, including the setting of traffic zones, extraction of typical activity patterns, formation of agents, and reward functions. The utilization of the survey data makes the simulation result closer to the actual situation in Shangyu; therefore, the simulation result has practical meanings and could be further utilized in transportation planning and management. For example, it could be used in TDM policy effect analysis.
- (iv) Data used in this paper come from the survey of a typical small city in east China. Both the survey data and the simulation results have distinct Chinese characteristics. For example, maintenance and leisure activity are conducted mostly in the morning and people tend to stay at home in the evening; commuting groups have few leisure and maintenance activities during weekdays. These features provide materials for future research of Chinese traffic.

The above mentioned analysis of the simulation result shows that this paper's simulation method could better reflect actual traffic conditions. Both the macroscopic spatial-temporal features and the microscopic activity-travel schedule render this method valid. The veracity of the simulation result and the utilization of survey data enable this method to better service practical transportation planning and management.

Because of the limitations of the survey data and the algorithm, several aspects of the research can be improved in the future.

- (i) Route choice in the current model is simplified. The travelers "jump" directly from the origin to the destination, while the influence on the intermediate regions is neglected.
- (ii) In this paper, the reward function contains four different parts; they are, respectively, based on attraction degree of zones, activity start time, duration, and travel time. When accumulated, the weights of them are considered to be equal. However, in reality, these factors have different effects on people when they make the decision on their trips. So one future direction is to calculate these weights according to the survey data, making the simulation results more accurate.
- (iii) Road impedance varies greatly according to the type of traffic mode, since different modes have different occupation rates of roads and their speed are also different. As a result, it is better to take traffic mode of each agent into consideration when calculating congestion degree.
- (iv) Reaction to uncertain events is a special characteristic of reinforcement learning. In this paper we are focusing on the most probable or the "average" state of the system. But it is also interesting to explore

how the agents would react to radical changes of the environment and how do they interact with each other under this circumstance.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the National Basic Research 973 program (2012CB725400) and Natural Science Foundation of China (51378120 and 51338003). Also Fundamental Research Funds for the Central Universities and Foundation for Young Key Teachers of Southeast University are appreciated.

References

- [1] M. Fried, J. Havens, and M. Thall, "Travel behavior—a synthesized theory," Final Report, NCHRP, Transportation Research Board, 1977.
- [2] F. S. Chapin, *Human Activity Patterns in the City*, John Wiley & Sons, New York, NY, USA, 1974.
- [3] T. Hägerstrand, "What about people in regional Science?" *Papers of the Regional Science Association*, vol. 24, no. 5, pp. 6–21, 1970.
- [4] J. L. Bowman and M. E. Ben-Akiva, "Activity-based disaggregate travel demand model system with activity schedules," *Transportation Research A*, vol. 35, no. 1, pp. 1–28, 2001.
- [5] F. S. Koppelman and C.-H. Wen, "The paired combinatorial logit model: properties, estimation and application," *Transportation Research B*, vol. 34, no. 2, pp. 75–89, 2000.
- [6] M. E. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press, London, UK, 1985.
- [7] C. R. Bhat, "A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity," *Transportation Research B*, vol. 30, no. 3, pp. 189–207, 1996.
- [8] P. M. Jones, *Understanding Travel Behavior*, University of Oxford, Transport Studies Unit, Oxford, UK, 1983.
- [9] M. G. Karlaftis and E. Vlahogianni, "Statistical methods versus neural networks in transportation research: differences, similarities and some insights," *Transportation Research C*, vol. 19, no. 3, pp. 387–399, 2011.
- [10] T. A. Arentze and H. J. P. Timmermans, "A learning-based transportation oriented simulation system," *Transportation Research B*, vol. 38, no. 7, pp. 613–633, 2004.
- [11] T. Arentze, F. Hofman, and H. Timmermans, "Reinduction of Albatross decision rules with pooled activity-travel diary data and an extended set of land use and cost-related condition states," *Transportation Research Record*, vol. 1831, pp. 230–239, 2003.
- [12] T. Arentze and H. Timmermans, "Parametric action decision trees: incorporating continuous attribute variables into rule-based models of discrete choice," *Transportation Research B*, vol. 41, no. 7, pp. 772–783, 2007.

- [13] K. M. Nurul Habib, "A random utility maximization (RUM) based dynamic activity scheduling model: application in weekend activity scheduling," *Transportation*, vol. 38, no. 1, pp. 123–151, 2011.
- [14] B. Fernandez-Gauna, J. M. Lopez-Guede, and M. Graña, "Transfer learning with partially constrained models: application to reinforcement learning of linked multicomponent robot system control," *Robotics and Autonomous Systems*, vol. 61, no. 7, pp. 694–703, 2013.
- [15] E. A. Jasmin, T. P. Imthias Ahamed, and V. P. Jagathy Raj, "Reinforcement learning approaches to economic dispatch problem," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 4, pp. 836–845, 2011.
- [16] A. Agung and F. L. Gaol, "Game artificial intelligence based using reinforcement learning," *Procedia Engineering*, vol. 50, pp. 555–565, 2012.
- [17] R. Lahkar and R. M. Seymour, "Reinforcement learning in population games," *Games and Economic Behavior*, vol. 80, pp. 10–38, 2013.
- [18] Z. Tan, C. Quek, and P. Y. K. Cheng, "Stock trading with cycles: a financial application of ANFIS and reinforcement learning," *Expert Systems with Applications*, vol. 38, no. 5, pp. 4741–4755, 2011.
- [19] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: a survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, 1996.
- [20] D. Charypar and K. Nagel, "Q-learning for flexible learning of daily activity plans," *Transportation Research Record*, vol. 1935, pp. 163–169, 2005.
- [21] W. Wang, *Transportation Engineering*, Southeast University Press, Nanjing, China, 2000.
- [22] M. Vanhulsel, D. Janssens, and G. Wets, "Calibrating a new reinforcement learning mechanism for modeling dynamic activity-travel behavior and key events," in *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, Transportation Research Board, Washington, DC, USA, 2007.
- [23] D. Janssens, Y. Lan, G. Wets, and G. Chen, "Allocating time and location information to activity-travel patterns through reinforcement learning," *Knowledge-Based Systems*, vol. 20, no. 5, pp. 466–477, 2007.

Research Article

Topological Effects and Performance Optimization in Transportation Continuous Network Design

Jianjun Wu,¹ Xin Guo,² Huijun Sun,² and Bo Wang³

¹ State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

² Moe Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China

³ Beijing Municipal Commission of Transport, Beijing 100073, China

Correspondence should be addressed to Jianjun Wu; jjwul@bjtu.edu.cn

Received 21 July 2013; Revised 3 December 2013; Accepted 19 December 2013; Published 21 January 2014

Academic Editor: Geert Wets

Copyright © 2014 Jianjun Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Because of the limitation of budget, in the planning of road works, increased efforts should be made on links that are more critical to the whole traffic system. Therefore, it would be helpful to model and evaluate the vulnerability and reliability of the transportation network when the network design is processing. This paper proposes a bilevel transportation network design model, in which the upper level is to minimize the performance of the network under the given budgets, while the lower level is a typical user equilibrium assignment problem. A new solution approach based on particle swarm optimization (PSO) method is presented. The topological effects on the performance of transportation networks are studied with the consideration of three typical networks, regular lattice, random graph, and small-world network. Numerical examples and simulations are presented to demonstrate the proposed model.

1. Introduction

The network design problem (NDP) involves the optimal decision on the expansion of an urban street and highway system in response to a growing demand for travel. It has emerged as an important area for progress in handling effective transport planning, because the demand for travel on the roads is growing at a rate faster than our urban transport systems can ever hope to accommodate, while resources available for expanding the system capacity remain limited. The objective of NDP is to optimize a given system performance measure such as to minimize total system travel cost, while accounting for the route choice behavior of network users [1]. The decisions made by road planners influence the route choice behavior of network users, which is normally described by a network user equilibrium model. Mathematically, the bilevel programming is a good technique to describe this hierarchical property of the NDP with an equilibrium constraint. Generally the upper level problem is to minimize

the total system cost and the lower level problem is to characterize the UE traffic flow pattern ([2, 3]).

Studies have been overwhelmingly focused on the continuous network design problem (CNDP) and substantial achievements in algorithmic development have been made. Abdulaal and LeBlanc [4] formulated the CNDP under deterministic user equilibrium (DUE) as a bilevel programming model and the Hooke-Jeeves heuristic algorithm was also introduced. As an application, Friesz et al. [5] used a simulated annealing approach to solve the multiobjective equilibrium network design problem as a single level minimization problem. Marcotte [6] transferred the CNDP into a single level equivalent differentiable optimization problem. Meng et al. [7] reformulated the CNDP under the DUE constraints into an equivalent single level continuously differentiable problem by virtue of a marginal function tool. Gao and Song [3] combined the concept of reserve capacity with the continuous equilibrium network design problem and proposed a globally convergent algorithm to solve the CNDP.

In the planning of road works, there should be awareness about the impacts of the increased capacity of a link on the whole network. Because the improvement of road capacity will attract the limitation of budget, the prioritization for road maintenance, repair, and contingency planning should be considered carefully in NDP. Therefore, increased efforts should be made on links that are more critical to the system. It would be thus helpful to model and evaluate the vulnerability and reliability of the transportation network when the network design is processing.

In a series of papers, the measures to quantifiable efficiency/performance of a network have been developed. For example, Latora and Marchiori discussed the network performance issue by measuring the “global efficiency” in a weighted network as compared to that of the simple non-weighted small-world network [8–10]. In a weighted network, the network is characterized not only by the edges that connect different nodes, but also by the weights associated with different edges in order to capture the relationships between different nodes.

The flow on a network is an additional important indicator of network performance as well as network vulnerability. Indeed, flows represent the usage of a network and which paths and links have positive flows and the magnitude of these flows are relevant in the case of network disruptions [11]. There are few papers to date that consider network flows in assessing network performance. The results in Zhu et al. [12] are notable since they demonstrate empirically through an application to the airline network of China how a measure with flows and costs outperforms existing measures in yielding more realistic results in terms of, for example, which cities are critical and their rankings in the network.

Recently, Jenelius et al. [13] proposed several link importance indicators and applied them to the road transportation network in northern Sweden. Murray-Tuite and Mahmassani [14] also focused on identifying indices for the determination of vulnerable links in transportation networks. Qiang and Nagurney [11] proposed a unified network performance measure that can be applied to assess the importance of either links or nodes or both in the case of either fixed or elastic demands.

Although previous work for the traffic network design has been focused on minimizing the total system cost or maximizing the total profits of the network, little attention has been given to the structure factors underlying vulnerabilities and the robustness. In the traffic network, if the Hub link is attacked, the whole network may be broken down soon. Therefore, the purpose of network designs is not only to decrease the travel cost of the system, but also to increase the robustness and reduce the vulnerabilities. Different with most previous works on CNDP, the performance measure is adopted to test the network design strategies in this paper. Our results give a better understanding of the network structure effects on its performance.

This paper is organized as follows. A bilevel model of CNDP is given in Section 2. Further solution algorithm is developed in Section 3. And Section 4 gives a numerical example to illustrate the model. Finally, conclusions and further considerations are presented in Section 5.

2. A Bilevel Programming Model for CNDP

The transportation CNDP can be represented as a leader-follower game where the transportation planning departments are leaders, and the users who can freely choose the path are the followers [15]. It is assumed that the transportation planning managers can influence, but cannot control, the users’ path-choosing behaviors. The users make their decision in a user optimal manner. Thus, this interactive game can be represented as a bilevel programming problem.

The transportation planners, the upper level, decide the capacity of each road to maximize the system performance based on traffic flows. However, the lower level reflects the choice behaviors of drivers with user equilibrium assignment. In this model, the system performance and link flows are considered all together.

Notations

A : the set of arcs (links);

R and S : the sets of vertices which represent origins and destinations, respectively;

x_a : the total flow on link a , $a \in A$, and A is the link set of the network;

ψ_a : the continuous capacity increase of link a ;

K_{rs} : the set of path between r and s ;

r : the origin node, $r \in R$;

s : the destination node, $s \in S$;

$t_a(\cdot)$: the link travel time (or cost) function which is continuously differentiable and convex for fixed ψ_a . Generally, we use the following form for $t_a(\cdot)$:

$$t_a(x_a, \psi_a) = t_{0a} \left[1 + \eta_a \left(\frac{x_a}{k_a + \psi_a} \right) \right]^4, \quad (1)$$

where η_a are parameters and k_a is the capacity of link a [16]. t_{0a} is the free-flow of link a ;

q_{rs} : the total traffic demand between origin r and destination s ;

h_k^{rs} : flows on path k connecting r and s ;

$\delta_{a,k}^{rs}$: path/link incidence variables.

2.1. The Lower Level User Equilibrium Assignment. It is worth emphasizing that the network design problem must be solved with the network flow pattern constrained to be a user equilibrium problem. In general, improvement of road network characteristics will definitely induce changes in traffic flow over the network. More importantly, addition of a new road segment or capacity enhancement to a congested network, without considering the response of network users, may actually increase network-wide congestion. This well-known phenomenon has been demonstrated by the ostensible Braess paradox. Therefore, prediction of traffic patterns via a comprehensive behavior model is essential to the network design process.

Traditionally, CNDP models hypothesize that the demand is given and fixed, and the users' route choice is characterized by the user equilibrium assignment problem. Let A be the set of arcs (links); R and S are the sets of vertices which represent origins and destinations, respectively. The user equilibrium problem with fixed demand can be formulated as follows [16]:

$$(L) \quad \min T(\mathbf{x}, \boldsymbol{\psi}) = \sum_{a \in A} \int_0^{x_a(\boldsymbol{\psi})} t_a(v, \psi_a) dv, \quad (2)$$

$$\text{s.t.} \quad \sum_k h_k^{rs} = q_{rs}, \quad \forall r \in R, s \in S, \quad (3)$$

$$h_k^{rs} \geq 0, \quad \forall r \in R, s \in S, k \in K_{rs}, \quad (4)$$

$$x_a = \sum_r \sum_s \sum_k h_k^{rs} \delta_{a,k}^{rs}, \quad \forall a \in A. \quad (5)$$

In this model, the users at the lower level are assumed to follow the user-equilibrium principle of Wardrop under the given network. Constraints (3), (4), and (5) are conservative, definitions, and nonnegativity of the flow constraints.

2.2. The Upper Level Model

2.2.1. The Performance Measure of Network. Recently, the vulnerability of a network has attracted many interests in the urban traffic system. As an important performance index, it can be used to assess the efficiency of a network in the case of either fixed or elastic demands and capture flow information and behavior, allowing one to determine the criticality of various nodes (as well as links) through the identification of their importance and ranking [11]. Therefore, this measure can be used to assess the vulnerability of a network to disruptions and is a more general one since it not only considers the topology characteristics, but also captures the flows on the network through the disutility, costs, and the demands.

The network performance/efficiency measure Z , for a given network topology G and the equilibrium (or fixed) demand vector q , is defined as follows [11]:

$$Z = \frac{\sum_{rs \in W} (q_{rs}/c_{rs})}{n_{rs}}, \quad (6)$$

where n_{rs} is the number of OD pairs in the network and q_{rs} and c_{rs} denote the equilibrium (or fixed) demand and the equilibrium disutility or the shortest path for OD pair rs , respectively. W is the set of OD pair.

In this paper, it is an optimized objective in the upper level model. The network planners of the upper level are assumed to make the decisions about the improvement of link capacities and investments in order to maximize the performance of the whole system in the range of budgets formulated by the government. Therefore, the upper level of the bilevel model (U) can be presented as follows:

$$(U) \quad \max Z = \frac{\sum_{rs \in W} (q_{rs}/c_{rs})}{n_{rs}}, \quad (7)$$

$$\text{s.t.} \quad \sum_{a \in A} G_a(\psi_a) \leq B, \quad (8)$$

$$\psi_a \geq 0; \quad \forall a \in A, \quad (9)$$

$$c_k^{rs} = \sum_{rs \in W, k \in K_{rs}, a \in A} \delta_{a,k}^{rs} t_a(x_a, \psi_a), \quad (10)$$

$$\forall rs \in W, \quad k \in K_{rs},$$

$$c_{rs} = \min \{c_k^{rs}\}, \quad \forall rs \in W, k \in K_{rs}, \quad (11)$$

where ψ_a is the continuous capacity increase of link a ; $G_a(\psi_a)$ is the investment function of link $a \in A$; B is the total investment budget; c_k^{rs} is the cost of path k of OD pair rs .

The objective of upper level model is to maximize the performance. Constraint (8) ensures that the total investment cost will not exceed the total budget. Constraint (9) is the nonnegativity of the decision variables. Constraint (10) expresses the relationship between the link and path cost. Constraint (11) is the shortest path cost between OD pair rs .

3. Solution Algorithm of the Bilevel Model

3.1. Particle Swarm Optimization (PSO). Since the bilevel programming is a NP-hard problem, in general, it is difficult to solve with optimization algorithms. Although many solution algorithms such as the sensitivity analysis based on an algorithm (SAB) [1, 3] have been developed, the SAB method has to calculate the inverse of the matrix and only can be used under many assumptions (e.g., assume that the lower level functions are second order continuous differentiable and the lower level has a unique solution for fixed upper level variables), which limits the application in engineering. Recently, the intelligent algorithms are designed to solve the bilevel programming including the Gene algorithm, Ant algorithm, and particle swarm optimization (PSO) [17]. Compared with other algorithms, PSO has some advantages: (1) the algorithm is easy to implement and there are few parameters to adjust; (2) all of the particles have a strong ability to find the optimized solution; (3) it need not any special conditions about functions and can be used to solve all kinds of bilevel models. Currently, PSO has been widely used in function optimization, neural networks, fuzzy systems control, and other applications. In recent years, the research and practice show that PSO has fast convergence at high solution quality, robustness, and other advantages in terms of optimization objectives. Therefore, PSO algorithm based on solution procedure is used in this paper.

Bird flocking optimizes a certain objective function. Each particle knows its personal best position ($pbest$) so far and the global best position ($gbest$) of the swarm among all the $pbests$. Each particle tries to modify its position using the current velocity and the distance from $pbest$ and $gbest$ [18].

In this paper, a particle position is denoted by a feasible link capacity enhancement $\psi_i = (\dots, \psi_{(i)a}, \dots)$. The velocity $V_i = (\dots, v_{(i)a}, \dots)$ of the i th particle corresponds to the search direction. Therefore, the position of i th particle is used to represent the feasible decision scheme. By comparing the fitness value, which denotes the value of upper objective function, the optimal positions $pbest_i^n$ and $gbest^n$ of i th particle

and the whole swarm particle at time n are obtained, respectively. Then the velocity and the position of i th particle are updated at time $n + 1$ as follows:

$$V_i^{n+1} = \omega V_i^n + c_1 R_{i1}^n (pbest_i^n - \psi_i^n) + c_2 R_{i2}^n (gbest^n - \psi_i^n), \quad (12)$$

$$\psi_i^{n+1} = \psi_i^n + \chi V_i^{n+1}, \quad (13)$$

where $i = 1, 2, \dots, M$; M is the swarm's size; χ is a constriction factor used to control and constrict velocities; ω is the inertia weight; c_1 and c_2 are two positive constants, called the cognitive and social parameter, respectively; R_{i1}^n and R_{i2}^n are random numbers uniformly distributed within the interval $[0, 1]$. The parameters $c_1 R_{i1}^n$ and $c_2 R_{i2}^n$ provide randomness that renders the technique less predictable but more flexible [17]. In general, the performance of each particle is measured according to fitness function.

Fourie and Groenwold [19] proposed a dynamic inertia and maximum velocity reduction method. That is, if no more improved solutions are obtained after several iterations, set the inertia weight $\omega^{n+1} = \alpha \omega^n$ and maximum allowed velocity $V_{\max}^{n+1} = \beta V_{\max}^n$ ($0 < \alpha, \beta < 1$) to fine-tune local search.

3.2. The Algorithm for Solving the Bilevel Model of CNDP. The major steps of the PSO method for solving the CNDP are summarized as follows.

Step 1 (initialization). Initialize parameters $\omega^0, V_{\max}, \alpha, \beta$ ($0 < \alpha, \beta < 1$), the dynamic delay period h , and the swarm size M which corresponds to feasible solutions and evaluate the initial velocity of every particle. The initial position ψ_i^0 and velocity V_i^0 are given. Let time step $n = 0$.

Step 2 (calculating fitness). For each particle, we translate it into a feasible solution. Then, we solve the corresponding lower level problem by Frank-Wolfe method [16] and evaluate the upper level objective in (7), calculating the particle's fitness.

Step 3 (updating). For each particle, update its $pbest$ with the position relevant to the smallest fitness which it so far experienced. For the swarm, update the $gbest$ with the position relevant to the smallest fitness among all the $pbest$. If no improvement in the $gbest$ occurs after h iterations, set $\omega^{n+1} = \alpha \omega^n, V_{\max}^{n+1} = \beta V_{\max}^n$. Then, update the velocities using (12) and truncate them to the nearest integers.

Step 4 (generating new position). Generate the new position with (13).

Step 5 (checking feasibility). All the new positions are tentative particles. Discard the new position i ($i = 1, 2, \dots, M$) if it does not satisfy constraints (8) and (9). Leave it at the previous position and go to Step 6.

Step 6 (termination). If the iteration number reaches the maximum iteration number or other termination criteria, then stop. Get the last $gbest$ as the optimal solution. Otherwise, $n++$, go to Step 2.

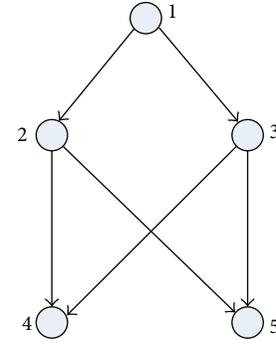


FIGURE 1: Test network.

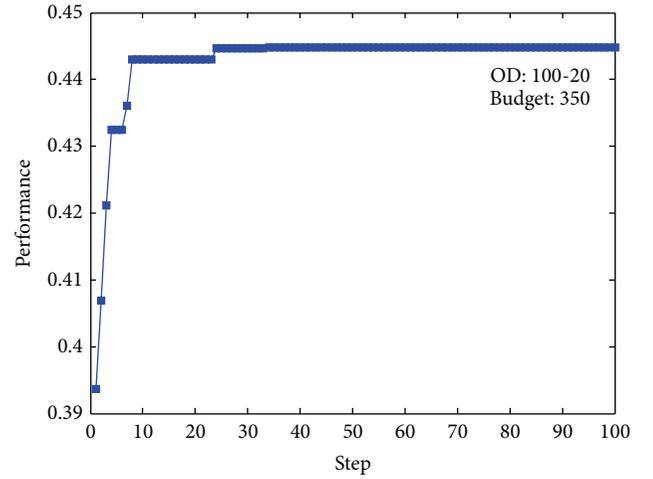


FIGURE 2: Illustration of the convergences for the algorithm.

TABLE 1: Data for the test network.

Link	t_{0a}	η_a	d_a	k_a
1 → 2	1.0	0.15	0.5	40
1 → 3	2.0	0.15	0.5	40
2 → 4	1.0	0.15	0.5	40
2 → 5	1.5	0.15	0.5	40
3 → 4	1.0	0.15	0.5	40
3 → 5	1.5	0.15	0.5	40

$$t_a(x_a, \psi_a) = t_{0a} [1 + \eta_a (x_a / (k_a + \psi_a))]^4$$

4. Numerical Examples

Example 1 (a simple network form). To illustrate the effectiveness of the proposed approach, consider the 5-node and 6-link graph given in Figure 1 [11]. There are two OD pairs from origin 1 to destinations 4 and 5, respectively. Table 1 presents the functional forms of the travel and investment costs, as well as the parameter values for each link, used in this numerical test.

The investment function is given as $G_a(\psi_a) = 1.5d_a(\psi_a)^2$, where d_a is a parameter of the function. In this example, assume that the largest iteration number $n = 1000$. Other

TABLE 2: Parameters used in Example 1.

Parameters	M	α, β	h	V_{\min}	V_{\max}	ω^0	c_1	c_2	χ
Values	10.0	0.98	15.0	0.05	0.2	0.8	2.1	0.9	1.0

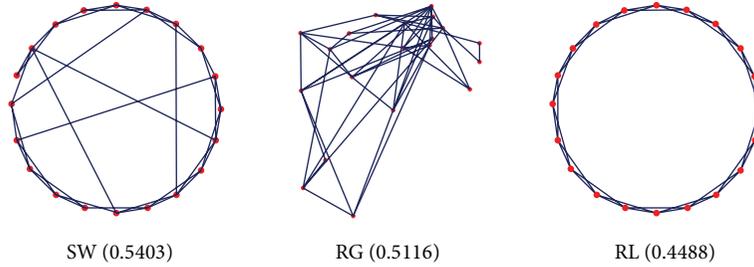


FIGURE 3: Three different network topologies.

TABLE 3: Comparison results of different budgets.

Budget	Capacity					
	1 → 2	1 → 3	2 → 4	2 → 5	3 → 4	3 → 5
150	4.60	2.46	3.12	2.41	1.72	1.26
250	5.73	2.97	4.42	3.28	2.96	0.63
350	6.18	4.45	4.85	0.11	4.03	1.58

TABLE 4: TSC of the network with different OD demands and budgets.

Budget	TSC				
	100-20	100-40	100-60	100-80	100-100
150	432.44	584.10	865.89	1213.67	1853.30
250	431.25	579.45	793.35	1159.37	1636.84
350	410.91	560.11	743.75	1112.40	1509.32

parameters are shown in Table 2. Table 3 presents the comparisons of different budgets with OD demand 100-80 (the two numbers denote the demand of OD pairs 1-4 and 1-5).

Tables 4 and 5 give the variation of total system cost (TSC) and performance (PER) with different budgets and OD demands. As we know, with the increase of budgets, TSC and PER could be improved by using the bilevel decision methods, which can avoid paradox effectively. This has been clearly reflected in Tables 4 and 5. The result indicates that the larger budget can alleviate traffic congestion and improve the performance efficiently. In addition, a positive fact is that the large OD demand will deteriorate the traffic.

To illustrate the efficiency of the solution algorithm further, a convergence test is given in the numerical example. Figure 2 shows the convergence of PSO in the case of budget 350 and OD demand 100-20. The result indicates that the algorithm can converge to a steady state after 30 steps.

Example 2 (an extended example considering network topologies). The structure properties on traffic networks have attracted a tremendous amount of recent interest which

TABLE 5: PER of the network with different OD demands and budgets.

Budget	Performance				
	100-20	100-40	100-60	100-80	100-100
150	0.3992	0.2799	0.1795	0.1272	0.0798
250	0.4188	0.2837	0.1969	0.1356	0.0954
350	0.4447	0.3289	0.2265	0.1446	0.1057

shows different network structures have important effects on their performance. It has been known well that there are three typical structures in traffic networks: random graph (RG), small-world network (SW), and regular lattice (RL). We start by constructing networks according to regular ER [20] and WS [21] algorithms. It has been demonstrated that SW networks have both a small value of average shortest path, like RG, and a high clustering coefficient, like RL. Such a definition corresponds to networks having a high value of global efficiency and a high value of local efficiency, that is, to networks extremely efficient in exchanging information both at a global and at a local scale [22].

In this example, to investigate the effects of topologies on NDP, the network design decisions for different network structures with the same nodes and edges are given. Three typical network topologies are showed in Figure 3. These networks are generated according to the procedure in [21], where self-links and repeated links are forbidden. Particularly, the SW network is formed by a RL rewired with probability 0.1. All topologies are set to have the same average degree $\langle k' \rangle \approx 4$ and 20 nodes. Other parameters are equal to that in Example 1.

The number of the brackets represents the initial performance before the network design. Table 6 provides the optimization results for three typical networks. From the table, we can see that PER of SW network is the largest, but RL is the smallest one. The result can be explained easily. The SW network has the small average shortest path which decreases the distance between two arbitrary nodes. Therefore it will be

TABLE 6: The optimization results for different network topologies.

Budget	Topology		
	SW	ER	RL
200	0.5483	0.5228	0.4604
250	0.5486	0.5231	0.4650
300	0.5486	0.5234	0.4659
350	0.5487	0.5238	0.4666
400	0.5488	0.5241	0.4667

an efficient network to improve the performance of traffic network. But for the RL network, the larger average shortest path increases the distance between two arbitrary nodes, which reduces the network performance. Therefore, the result is in accordance with the prediction of “structure-determinative-function.”

5. Conclusion

The transportation network performance is an important indicator to evaluate its reliability. In order to improve the network vulnerability, the performance should be considered in network design decisions. This paper studies a new form of transportation network design problem by optimizing the network performance, and a bilevel programming model is proposed to describe this problem. In the model, the upper level is to optimize the system performance within limited budgets, while the lower level is user equilibrium assignment problem. Finally, numerical examples show that, by redesigning the network, the performance (the total system cost) will increase (reduce) greatly. In addition, the network structure has profound effects on its performance. Compared with the other two topologies, the form of small-world network is the best one.

In this paper, it is assumed that the OD matrix is fixed. In reality, OD demand would fluctuate every day, even every hour. Therefore it would be interesting to examine the dynamic OD demand in network design problem further.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper is partly supported by National Basic Research Program of China (2012CB725400), NSFC (71271024), FANEDD (201170), the Program for New Century Excellent Talents in University (NCET-12-0764), and the Fundamental Research Funds for the Central Universities (2012JBZ005).

References

- [1] H. Yang and M. G. H. Bell, “Models and algorithms for road network design: a review and some new developments,” *Transport Reviews*, vol. 18, no. 3, pp. 257–278, 1998.
- [2] Z. Gao, H. Sun, and H. Zhang, “A globally convergent algorithm for transportation continuous network design problem,” *Optimization and Engineering*, vol. 8, no. 3, pp. 241–257, 2007.
- [3] Z. Y. Gao and Y. F. Song, “A reserve capacity model of optimal signal control with user-equilibrium route choice,” *Transportation Research B*, vol. 36, no. 4, pp. 313–323, 2002.
- [4] M. Abdulaal and L. J. LeBlanc, “Continuous equilibrium network design models,” *Transportation Research B*, vol. 13, no. 1, pp. 19–32, 1979.
- [5] T. L. Friesz, G. Anandalingam, N. J. Mehta, K. Nam, S. J. Shah, and R. L. Tobin, “The multiobjective equilibrium network design problem revisited: a simulated annealing approach,” *European Journal of Operational Research*, vol. 65, no. 1, pp. 44–57, 1993.
- [6] P. Marcotte, “Network optimization with continuous control parameters,” *Transportation Science*, vol. 17, no. 2, pp. 181–197, 1983.
- [7] Q. Meng, H. Yang, and M. G. H. Bell, “An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem,” *Transportation Research B*, vol. 35, no. 1, pp. 83–105, 2001.
- [8] V. Latora and M. Marchiori, “Efficient behavior of small-world networks,” *Physical Review Letters*, vol. 87, no. 19, Article ID 198701, 2001.
- [9] V. Latora and M. Marchiori, “Economic small-world behavior in weighted networks,” *European Physical Journal B*, vol. 32, no. 2, pp. 249–263, 2003.
- [10] V. Latora and M. Marchiori, “How the science of complex networks can help developing strategies against terrorism,” *Chaos, Solitons and Fractals*, vol. 20, no. 1, pp. 69–75, 2004.
- [11] Q. Qiang and A. Nagurney, “A unified network performance measure with importance identification and the ranking of network components,” *Optimization Letters*, vol. 2, no. 1, pp. 127–142, 2008.
- [12] D. Zhu, Y. Hu, Y. Li, and B. Yu, *A New Measure For Airline Networks Performance Evaluation and Critical Cities Identification*, Fudan University, Shanghai, China, 2006.
- [13] E. Jenelius, T. Petersen, and L.-G. Mattsson, “Importance and exposure in road network vulnerability analysis,” *Transportation Research A*, vol. 40, no. 7, pp. 537–560, 2006.
- [14] P. M. Murray-Tuite and H. S. Mahmassani, “Methodology for determining vulnerable links in a transportation network,” *Transportation Research Record*, vol. 1882, pp. 88–96, 2004.
- [15] D. E. Boyce, “Urban transportation network equilibrium and design models: recent achievements and future prospectives,” *Environment and Planning*, vol. 16A, pp. 1445–1474, 1984.
- [16] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1985.
- [17] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, Australia, December 1995.
- [18] K. Premalatha and A. M. Natarajan, “A new approach for data clustering based on PSO with local search,” *Computer and Information Science*, vol. 1, no. 4, pp. 139–145, 2008.
- [19] P. C. Fourie and A. A. Groenwold, “The particle swarm optimization algorithm in size and shape optimization,” *Structural and Multidisciplinary Optimization*, vol. 23, no. 4, pp. 259–267, 2002.
- [20] P. Erdős and A. Rényi, “On the evolution of random graphs,” *Publications of the Mathematical Institute of the Hungarian Academy of Science*, vol. 5, pp. 17–61, 1960.

- [21] D. J. Watts and S. H. Strogatz, "Collective dynamics of "small-world" networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [22] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

Research Article

The Model of Severity Prediction of Traffic Crash on the Curve

Jian-feng Xi,^{1,2} Hai-zhu Liu,² Wei Cheng,³ Zhong-hao Zhao,² and Tong-qiang Ding²

¹ State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China

² College of Transportation, Jilin University, Changchun 130022, China

³ Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650224, China

Correspondence should be addressed to Wei Cheng; chengwei_ding@163.com

Received 5 September 2013; Revised 20 November 2013; Accepted 29 November 2013; Published 9 January 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Jian-feng Xi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the study of traffic crashes on curved road segments as the focus of research, a logistic regression based curve road crash severity prediction model was established based on a sample crash database of 20000 entries collected from 4 regions of China and 15 evaluation indicators involving driver, driving environment, and traffic environment factors. Maximum Likelihood Estimation and step-back technique were deployed for data analysis, the conclusion of which is that the three main contributory factors on curve road crash severity are weather, roadside protection facility, and pavement structure. Hosmer and Lemeshow tests were used to verify the reliability of the model, and the model variables were discussed to a certain degree as well.

1. Introduction

As a major component of road geometric design, curved road segments, due to their alignment characteristics, are most prone to traffic crashes among all road geometric elements. In 2010, crashes on curved segments accounted for 10.5% of the total number of traffic crashes in China. Correspondingly, the number of deaths accounted for 12.89% of the total number of deaths, as shown in Figure 1.

Many researches have been conducted with regard to the characteristics and causes of crash on curved segments, ranging from those based on macroeconomic statistics, where crash rates of straight and curved segments are compared [1, 2], and those based on vehicle dynamics, where the focuses are on improving vehicle's safety and through performance on curves [3, 4], to those that are based on the study of human factors, where the studies of traffic crash have been heavily focused on drivers and driver behavior [5]. Vehicle speed and geometric design are also topics of study when curve road crashes are concerned [6, 7]. Regarding crash severity, certain mathematical models have been established, such as the Bayesian Method [8], the Ordered Probit Model [9], and the Neuronal Network Approach [10]. However, among the aforementioned studies and researches, few had aimed

specifically at investigating crash severity on curved road segments.

Analysis and prediction models are based on crash data. In the past, data used by traffic crash studies on curve roads was mostly gathered from specific segments. Crash samples and crash-related factors had not been duly considered either. For these reasons, crash analysis would often fail to explain the real causes of crash, and the models were often questionable in terms of reliability and adaptability.

In this study, the crash data consists of 500 randomly chosen samples, according to the data gathering and management standards of China's "National Road Traffic Crash Information System," from a crash database of 20000 valid crash data entries that covers 4 regions of China from 2004 to 2008. The database comprises 60 items of crash information, which is comprehensive enough to recreate the process of a crash and provide an important basis for the analysis of causes. In terms of analysis method, the crash data is the foundation of analysis of the cause of crash and the structure and form of the crash determines the model for crash causation analysis. Since the attributed causes of each crash cannot be isolated from a single point of interpretation [11] and various internal and external factors are affecting drivers' reaction to road traffic safety, this paper made a

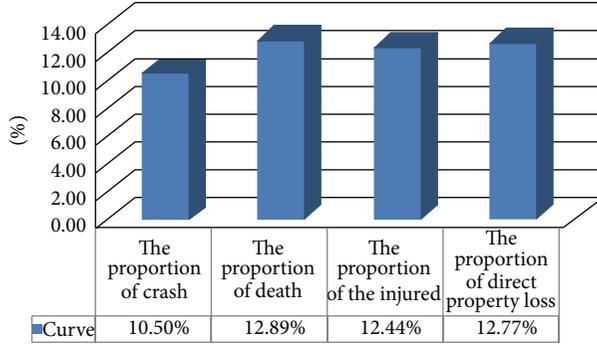


FIGURE 1: National crash loss on the curve in 2010.

comprehensive analysis of the severity of crash on curves from three aspects that include driver, vehicle type, and the road environment.

In recent years, due to its innate suitability, the logistic regression model has been widely used in practical problems, such as Banking [12], Genomics [13], and Psychopathology [14]. Many researchers used the logistic regression model to analyze the safety of roads, such as crashes of single-vehicle motorcycles [15], and crash prevention [16, 17]. Descriptive factors are defined as natural numbers that start from 0 and can be transformed into discrete variables. By this transformation, traffic crash data can meet the requirements of logistic regression model; therefore, the severity prediction model of traffic crash on curved segments may be built based on these data. The model [18] can overcome the deficiencies of analysis method and linear regression analysis of the traditional Mantel-Haenszel model, for it can contain multiple influencing factors that include an analysis of both discrete and continuous variables. The model can effectively analyze mixed influences and interactions from external variables and therefore also provides a methodological basis for the quantitative description relationship between multiple influencing factors and the prediction of severity on curves.

2. Establishing the Logistic Regression Model

2.1. Binary Logistic Regression Probability Formula. The severity of traffic crash on curved segments is regarded as dependent variable y , when the i th crash has any number of deaths, $y_i = 1$; otherwise, $y_i = 0$.

Assuming there are p influence factors that are related to dependent variable y and are denoted $X = (x_1, x_2, \dots, x_p)$, the probability of fatal crash under the impact of influencing factors is

$$P = \left\{ y_i = \frac{1}{X} \right\} = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}, \quad (1)$$

where x_i ($i = 1, 2, \dots, p$) are the influencing factors of crash severity; it can be a continuous variable, categorical variables, or dummy variables; β_i ($i = 1, 2, \dots, p$) are the regression coefficients.

2.2. Binary Logistic Regression Model Parameter Estimation. Considering the severity of the crash to be a dichotomous dependent variable, x_1, \dots, x_p are its corresponding independent variables. Let $x_i = (1, x_{i1}, \dots, x_{ip})'$, and $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ be its corresponding variable vector, $i = 1, \dots, n$. Thus, a logistic model may be established as [15]

$$\text{logit}(y) \triangleq \ln \left(\frac{p_i}{1 - p_i} \right) = x_i' \beta, \quad (2)$$

in which $p_i = P(y_i = 1/x_i) = e^{x_i' \beta} / (1 + e^{x_i' \beta})$.

The model used Maximum Likelihood Estimation method for variable estimation and the likelihood function of variable $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ can easily be derived from the binary logistic regression model, which is

$$l(\beta) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i}, \quad (3)$$

where $y_i = 0$ or 1. Therefore, the logarithmic likelihood function is

$$L(\beta) = \ln l(\beta) = \sum_{i=1}^n [y_i x_i' \beta - \ln(1 + x_i' \beta)]. \quad (4)$$

For $(\partial L(\beta)) / \partial \beta = 0$, it resulted in

$$\sum_{i=1}^n [x_i (y_i - p_i)] = 0. \quad (5)$$

In the logistic regression model, the Newton-Raphson method is used. Define $s_i = y_i - p_i$, $v_i = p_i(1 - p_i)$, and let $S = (s_1, \dots, s_n)'$, $V = \text{diag}(v_i)$; then

$$\begin{aligned} \frac{\partial L(\beta)}{\partial \beta} &= \sum_{i=1}^n x_i (y_i - p_i) = X' S, \\ -\frac{\partial^2 L(\beta)}{\partial \beta \partial \beta'} &= \frac{\partial}{\partial \beta'} \left(\sum_{i=1}^n x_i (y_i - p_i) \right) \\ &= -\frac{\partial}{\partial \beta'} \left(\sum_{i=1}^n x_i p_i \right) \\ &= \sum_{i=1}^n p_i (1 - p_i) x_i x_i' \\ &= X' V X, \end{aligned} \quad (6)$$

in which

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}. \quad (7)$$

If the number of iterations is k , then the maximum likelihood estimate of variable β becomes

$$\begin{aligned}\hat{\beta}^k &= \hat{\beta}^{k-1} - \left[\left(\frac{\partial^2 L(\beta)}{\partial \beta \partial \beta'} - i \frac{\partial L(\beta)}{\partial \beta} \right) \right]_{\beta=\hat{\beta}^{k-1}} \\ &= \hat{\beta}^{k-1} + \left[(X'VX)^{-1} (X'S) \right]_{\beta=\hat{\beta}^{k-1}}.\end{aligned}\quad (8)$$

3. Selection of Crash Impact Factor

Through a detailed analysis on the structure of the crash database, it is found that the database is a superdimensional structure with radial, multidimensional, and multilevel characteristics. Each crash record contains multiple data attributes and each value reflects a traffic crash's characteristics in one aspect. In addition, the setting of data attributes is based on five factors, which are the information of the crash, personnel information, vehicle information, road information, and environmental information. Therefore, logically the data attributes and five factors of people, vehicles, roads, climate environment and crash basic information formed a two-level formation, in which the five factors of people, vehicles, roads, environment, and crash basic information are in the upper layer, and the data attributes are in the lower layer. Also, different attributes and the attribute value formed a set complying with their specific logical relations and within each attribute there existed a specific hierarchy.

500 crash samples, randomly selected from a crash information database of 20000 entries, have been analyzed. The causes of crash are mostly related to bad subjective judgment and human errors, vehicle performance issues, change of external environment, and change of road conditions. However, vehicle performance is generally not considered in crash analysis and is neglected in this study.

Therefore, from three-level analysis system of drivers, driving environment, and road environment, the model selects 15 evaluation items as independent variables. The evaluation of independent variables is shown in Table 1.

4. Prediction Model of Traffic Crash Severity on Curved Segments

The model used stepwise regression method to analyze the independent variables and step-back technique to obtain the results. In step 1, 15 independent variables were all put into the model and the variables based on the probability of the likelihood ratio for the test were assumed. In step 13, the weather X_6 , roadside protective facility type X_{12} , the road pavement X_{14} , and constant were selected. Results are shown in Table 2.

Taking a significance level of 0.05 and using the reverse stepwise method and 13 times of screening for selection, the model obtained the correlation of crash severity on curves with weather, road-side protective facility type, and road pavement structure. From Hosmer and Lemeshow tests, the results in Table 3 showed that the significance level of 0.674 is greater than 0.05 and thus proved that the original assumption is valid. Also, the chi-square value from 7.856

of the first screening reduced to 0.79 in the 13th screening, which proved that the model is correct.

The probability of crash with death from the logistic regression is expressed by

$$\begin{aligned}P &= \left(y_i = \frac{1}{X} \right) \\ &= \frac{e^{-1.634+2.258X_6-1.321X_{12}-1.46X_{14}}}{1 + e^{-1.634+2.258X_6-1.321X_{12}-1.46X_{14}}}.\end{aligned}\quad (9)$$

5. Conclusion

(1) Not all of the 15 impact factors are selected to put into the curve road crash severity prediction model. Some were excluded partly due to weak correlations, but it does not mean that they have little impact on the severity of a crash. For example, visibility and road surface conditions are closely related to weather conditions and their effects may be indirectly reflected in the final model, if such factors are to be further considered.

(2) During bad weather conditions, road surface friction coefficient decreases and visibility of road condition reduced, and vehicles driving on curved segments will be prone to cross over the median, causing side scraping, rollover, or rear-end crashes. During rainy or foggy weathers, water film will form on the surface of the road reducing tire friction. Therefore, weather is one of the main factors that affect crash severity on curved segments, and the severity will be even greater especially under ice and snow conditions.

(3) Highway safety design often focuses on road alignment and profile, but lacks due considerations toward road-side facilities. However, in the real world, road-side environment in general is closely related to traffic crashes. On curve roads, off-road crashes frequently occur and it is mainly because of improper speed control and road-side environment, especially on mountainous curves where one side of the road is either deep grooves or cliffs. If curved segments have roadside protection facilities, it can greatly reduce the severity of the crash. Hence, better design and installment of road-side protection facilities should be one of the priorities for preventing severe crashes from happening.

(4) Compared to gravel pavement, asphalt roads generally have higher crash rates. Also, compared to the cement concrete pavement, asphalt roads have smoother surface and greater involuntary horizontal movement from vehicles. Therefore, drivers usually feel more comfortable driving on an asphalt-paved road surface, which often leads to high speed driving and speed related crashes. Asphalt pavement is more sensitive to the temperature; its structural strength decreases at high temperature and cracks at low temperature. However, drivers would often neglect the effect of temperature changes on the condition of road surface and their effects on safe driving, which is another very important factor in causing severe crashes.

(5) The existing road traffic crash database of China was originally designed to determine the responsibility of crash for legal purposes. With its 60 data items or attributes,

TABLE 1: Evaluation of independent variables.

Evaluation of the category	Variable symbol	Variable content	Variable assignment
Driver	X_1	Gender	Female = 0, male = 1
	X_2	Age classification	Age from 16 to 25 = 0, 26 to 35 = 1, 36 to 45 = 2, above 46 = 3
	X_3	Household register	Agriculture household = 0, nonagricultural household = 1
	X_4	Driving experience classification	1–5 years = 0, 6–10 years = 1, 11–15 years = 2, 16–20 years = 3, above 20 years = 4
	X_5	The accident responsibility	Minor responsibility = 0, equal responsibility = 1, main responsibility = 2, all responsibility = 3
Driving environment	X_6	Weather	Sunny = 0, not sunny (rain, snow, fog, cloudy, wind) = 1
	X_7	Terrain	Plain = 0, others = 1
	X_8	Visibility	Above 200 m = 0, 100–200 m = 1, 50–100 m = 2, below 50 = 3
	X_9	Road surface condition	Dry = 0, not dry (rain, ice) = 1
Road environment	X_{10}	Lighting condition	Day = 0, night with light = 1, night without light = 2
	X_{11}	Traffic signal type	With signal control = 0, without signal control = 1
	X_{12}	Roadside facilities protection type	With roadside protection = 0, without roadside protection = 1
	X_{13}	Road physical isolation	With isolated = 0, without isolated = 1
	X_{14}	Pavement structure	Asphalt pavement = 0, not asphalt pavement = 1
	X_{15}	Road conditions	Road surface in good condition = 0, pavement damage = 1

TABLE 2: Variable estimation.

Variable	B	S.E.	Wals	df	Sig.	Exp (B)
Weather	2.258	1.443	2.448	1	0.118	9.56
Roadside protective equipment	-1.321	0.853	2.398	1	0.122	0.267
Road pavement	1.46	0.531	7.546	1	0.006	4.304
Constant	-1.634	0.403	16.409	1	0.000	0.195

TABLE 3: Hosmer and Lemeshow test.

Step	Chi-square	Freedom	Significance level
1	7.856	8	0.448
13	0.79	2	0.674

although very comprehensive, it still needs future improvement in order to better describe the whole process of an actual traffic crash and to be used as the basis for crash prediction and prevention studies.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This project was supported by the NSFC (Grant no. 50808093).

References

- [1] T. Miller, D. Lestina, M. Galbraith, T. Schlaw, P. Mabery, and R. Deering, “United States passenger-vehicle crashes by crash geometry: direct costs and other losses,” *Accident Analysis and Prevention*, vol. 29, no. 3, pp. 343–352, 1997.
- [2] C. Zhang and J. N. Ivan, “Effects of geometric characteristics on head-on crash incidence on two-lane roads in connecticut,” *Journal of the Transportation Research Board*, vol. 1908, pp. 159–164, 2006.
- [3] S. K. Rao, “Unit curve for design of highway transitions,” *Journal of Transportation Engineering*, vol. 121, no. 2, pp. 169–175, 1995.
- [4] J. B. Beniley and W. E. Gallagher, “Review of highway curve design,” *Journal Highway*, vol. 18, no. 11, pp. 7–16, 2005.
- [5] S. Schick, “Accident related factors,” Trace Coordinator V3, 2009.
- [6] R. Yuanyuan, *Research on Driving Dangerous Area and Driving Behavior Model in Road Curved Section*, Traffic Environment and Safety Technology, Jilin University, Changchun, China, 2011.

- [7] W. Chaoshen, *Study on Road Traffic Safety Analysis and Countermeasures in Curve*, Transportation Planning and Management, Chang'an University, Xi'an, China, 2010.
- [8] J. Ma, K. M. Kockelman, and P. Damien, "A multivariate poisson-lognormal regression model for prediction of crash counts by severity, using bayesian methods," *Accident Analysis and Prevention*, vol. 40, no. 3, pp. 964–975, 2008.
- [9] B. Mediouny, Z. Anis, and H. S. Habib, "Motorway traffic crash prediction based on the network approach: application to the ringway of Paris," *Evaluation at Optimization System in the Production Services*, vol. 1, pp. 1–8, 2011.
- [10] S. M. Rifaat and H. C. Chin, "Accident severity analysis using ordered probit model," *Journal of Advanced Transportation*, vol. 41, no. 1, pp. 91–114, 2007.
- [11] R. Howard, *Statistical Modeling and Analysis of Injury Severity Sustained by Occupants of Passenger Vehicle Involved in Crashes with Large Trucks*, University of Nevada, Reno, Nev, USA, 2010.
- [12] S. Y. Sohn and H. S. Kim, "Random effects logistic regression model for default prediction of technology credit guarantee fund," *European Journal of Operational Research*, vol. 183, no. 1, pp. 472–478, 2007.
- [13] C. M. Zhang, H. D. Fu, Y. Jiang, and T. Yu, "High-dimensional pseudo-logistic regression and classification with applications to gene expression data," *Computational Statistics and Data Analysis*, vol. 52, no. 1, pp. 452–470, 2007.
- [14] C. R. García-Alonso, J. Guardiola, and C. Hervás-Martínez, "Logistic evolutionary product-unit neural networks: innovation capacity of poor guatemalan households," *European Journal of Operational Research*, vol. 195, no. 2, pp. 543–551, 2009.
- [15] V. Shankar and F. Mannering, "An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity," *Journal of Safety Research*, vol. 27, no. 3, pp. 183–194, 1996.
- [16] M. A. Abdel-Aty and H. T. Abdelwahab, "Predicting injury severity levels in traffic crashes: a modeling comparison," *Journal of Transportation Engineering*, vol. 130, no. 2, pp. 204–210, 2004.
- [17] R. Harb, E. Radwan, X. Yan, A. Pande, and M. Abdel-Aty, "Freeway work-zone crash analysis and risk identification using multiple and conditional logistic regression," *Journal of Transportation Engineering*, vol. 134, no. 5, pp. 203–214, 2008.
- [18] A. S. Al-Ghamdi, "Using logistic regression to estimate the influence of accident factors on accident severity," *Accident Analysis and Prevention*, vol. 34, no. 6, pp. 729–741, 2002.

Research Article

Modeling of Car-Following Required Safe Distance Based on Molecular Dynamics

Dayi Qu, Xiufeng Chen, Wansan Yang, and Xiaohua Bian

Automobile and Transportation College, Qingdao Technological University, Qingdao 266033, China

Correspondence should be addressed to Dayi Qu; dyqu@263.net

Received 1 October 2013; Revised 18 November 2013; Accepted 4 December 2013; Published 8 January 2014

Academic Editor: Wuhong Wang

Copyright © 2014 Dayi Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In car-following procedure, some distances are reserved between the vehicles, through which drivers can avoid collisions with vehicles before and after them in the same lane and keep a reasonable clearance with lateral vehicles. This paper investigates characters of vehicle operating safety in car following state based on required safe distance. To tackle this problem, we probe into required safe distance and car-following model using molecular dynamics, covering longitudinal and lateral safe distance. The model was developed and implemented to describe the relationship between longitudinal safe distance and lateral safe distance under the condition where the leader keeps uniform deceleration. The results obtained herein are deemed valuable for car-following theory and microscopic traffic simulation.

1. Introduction

Car-following models elaborate the situation that vehicles follow one another in the same lane and capture drivers' maneuvering decisions under different conditions. The models are the most fundamental part of microscopic traffic simulation [1]. From the research of Reuschel and Pipes using the operational research theory method on car-following, the models can be classified as stimulus-response models (Gazis et al., 1961; Newell, 1961), safe distance models (Gipps, 1981), psychophysical models (Wiedemann, 1974), and artificial intelligence models (Kikuchi and Chakroborty, 1992; Wu et al., 2000) [2–6]. Among them, the car-following model based on the safe distance has broadly been used in practice [7].

Parker put forward an expectation distance model for the first time through the research on car-following behavior on the fast road sections [8]. Peter Hides carried out further research on microscopic behavior of urban traffic flow and preliminarily established car-following models for urban traffic flow based on expectation distance [9]. Zhang et al. gave deep research on psychological and physical reaction mechanism of the driver and presented multiregime model based on the driver's psychological reaction [10]. At present, the car-following model based on safe distance has been

widely applied to microscopic simulation of road traffic, which is one of the hot research topics in the field of traffic engineering [11, 12]. Reaction time for a car-following maneuver responding to an unexpected hazard in the roadway has been conducted in several studies [13–18]. From their results, the mean reaction times identified are rarely greater than 1.50 seconds. The safe distances in the articles above all gave the longitudinal clearance from a preceding vehicle in following procedure, while there always exists lateral impact in car-following. Gunay carried out research taking into lateral clearances for vehicles change and established the model taking the lateral positions of vehicles into account [19].

This paper concentrates on the model of car-following required safe distance. Required safe distance is necessary for ensuring the driving safety in car following procedure, while the factors such as the longitudinal clearances between vehicles and the speed also have implications on the vehicle. Thus, we make analysis on the influencing factors and perform the simulation for the required safe distance and lateral safe distance under different speed conditions of the leader and the follower, which indicates that the model of required safe distance both benefits the operating safety and driving efficiency.

The structure of the paper is organized as follows. After the introduction, analysis of molecular dynamics is described

combined with the car following procedure in Section 2. Section 3 states the models of required safe distance including required safe distance and lateral safe distance. The influencing factors are illustrated from the perspective of longitudinal clearance and the speed in Section 4. Section 5 presents the test and validation to demonstrate the application of the model. The final section summarizes the findings and the conclusions of the paper.

2. Analysis of Molecular Dynamics

Among the molecules, attractive force and repulsive force simultaneously exist, both of which increase with the distance decreasing. But their changing rules are different that the attractive force declines more slowly. Molecular forces makes the molecules difficult to approach and be far away, as a phenomenon that the gas is difficult to be compressed and be expanded at a certain temperature. Vehicles in the car-following fleet show similar characteristics. From the aspect of safety, drivers try to keep a safe distance with the leader; at a certain speed, the phenomenon shows that the fleet is difficult to be compressed. So if the leader speed increases, for the sake of the efficiency, the follower will accelerate not to be fallen behind by the leader for long time. Thus, the clearance between adjacent vehicles will not be continually enlarged and the whole fleet will not be expended too much. Here, we define these characters as molecular car-following behavior, the theory of studying which is called molecular car-following theory.

In the car-following procedure, each vehicle is independent, but it also affects one another. Relative to the leader, every vehicle is the follower, while it is the leader that is relative to the rear vehicle. Similar to the dynamics relationship of molecules, there is an equilibrium distance called required safe distance in the fleet, just like the clearance between molecules where the resultant force is zero. As shown in Figure 1, on the assumption that the braking efficiency of three vehicles is identical, the braking distance is only influenced by the speed. And the required safe distance is proposed into front, lateral, and rear distance according to different directions. We call the furthest distance the follower keeps with the leader as required frontier and the furthest distance the follower keeps with the rear vehicle as required trailing edge. With respect to a vehicle, required frontier is actively obtained, while required trailing edge is passively obtained.

We define the state that the follower positions at the required trailing edge and drives at the same speed of the leader as equilibrium state. As illustrated in Figure 1, the front required safe distance is equal to the equilibrium distance in molecular dynamics with the same direction as vehicles' and also brings about influence on the followers. With speed as reference to the leader itself, the minimum safe distance the leader expects the follower to keep is called rear distance. Moreover, for the purpose to avoid confusing with the lateral safe distance, the front safe distance and the rear distance are together defined as required safe distance.

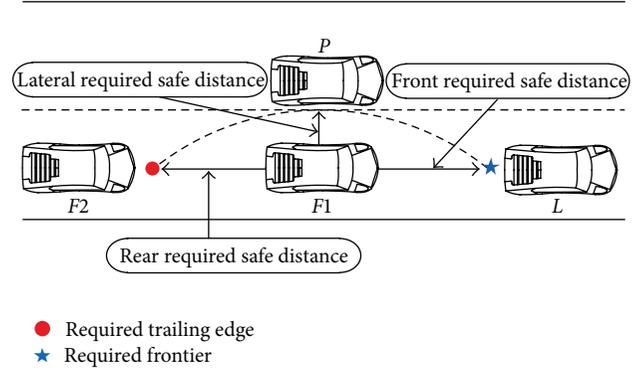


FIGURE 1: Vehicle required safe distance in car-following.

When the driver perceives the leader's travelling state being changed, the required safe distance is the minimum safe distance that the driver starts to act until stopping braking. The follower will not be fallen behind for long time considering the efficiency while will not be too near considering the safety in the following procedure.

3. Modeling Required Safe Distance

3.1. Required Safe Distance. When the leader maintains the braking state of uniform deceleration, supposing that v_L and v_F are respectively the original speeds of the leader and the follower and a_{Lm} and a_{Fm} are, respectively, the maximum decelerations, then the relative speed is $(v_F - v_L)$.

The finding shows that t_r is the sum of reaction time and braking coordination time ranging from 0.8 second to 1.0 second; t_i is the build-up time of deceleration ranging from 0.1 second to 0.2 second; then the braking distance can be get

$$X = v_0 \left(t_r + \frac{t_i}{2} \right) + \frac{v_0^2}{2a_m}, \quad (1)$$

where v_0 is the initial speed, a_m is the maximum deceleration.

If the leader operates with uniform motion or uniform acceleration, the follower can easily keep following in a safe state. After a period of time, the follower will drive at its original speed not to collide with the leader as long as the leader keeps a uniform motion. The follower always tries to achieve speed close to the leader's to improve travelling efficiency. If the leader conducts uniform deceleration, the required safe distance of the follower should be analyzed based on the changes of vehicles' speed before and after.

(1) $v_F > v_L$. Under the condition that the speed of the follower is faster than the leader and the leader does a uniform deceleration, the follower may collide with the leader without timely deceleration. At this moment, the leader will actively do uniform deceleration and the braking distance with no reaction time is

$$X_L = \frac{v_L^2 + v_L t_i a_{Lm}}{2a_{Lm}}. \quad (2)$$

If the follower decelerates after the time t_r , where t_r is the perceiving time after the leader changed, then the braking distance is

$$X_F = v_F t_r + \frac{v_F^2 + v_F t_i a_{Fm}}{2a_{Fm}}. \quad (3)$$

From the analysis above, we can get required safe distance of the follower in uniform deceleration condition as

$$X_R = v_F t_r + \frac{v_d t_i}{2} + \frac{v_F^2}{2a_{Fm}} - \frac{v_L^2}{2a_{Lm}} + d, \quad (4)$$

where d is the minimum clearance between adjacent vehicles after stopping, ranging from 2 meters to 5 meters.

(2) $v_F = v_L$. This situation is similar to the above. So the required safe distance is:

$$X_R = v_F t_r + \frac{v_L^2}{2} \left(\frac{1}{a_{Fm}} - \frac{1}{a_{Lm}} \right) + d. \quad (5)$$

(3) $v_F < v_L$. If there is a bigger clearance than required safe distance between adjacent vehicles, two vehicles can keep a safe state during a period of time. But if the follower continues driving at the initial speed, collisions may occur later on. So the follower should decelerate when the speed of the follower approaches the leader speed.

Supposing that the leader decelerates with maximum deceleration when the speed of follower is equal to the leader speed, then the time from decelerate till this moment is

$$t_1 = t_i + \frac{(v_L - (a_{Lm} t_i / 2) - v_F)}{a_{Lm}}. \quad (6)$$

During this time, the follower operates with the same speed, and the distance is

$$X_{F1} = v_F t_i + \frac{v_F}{a_{Lm}} \left(v_L - \frac{a_{Lm} t_i}{2} - v_F \right). \quad (7)$$

And when the follower must decelerate, the braking distance is:

$$X_{F2} = v_F t_r + \frac{v_F^2 + v_F t_i a_{Fm}}{2a_{Fm}}. \quad (8)$$

So the total distance of the follower is

$$X_F = v_F (t_i + t_r) - \frac{v_F v_d}{a_{Lm}} + \frac{v_F^2}{2a_{Lm}}. \quad (9)$$

And the leader keeps uniform deceleration and its driving distance is

$$X_L = \frac{v_L^2 + v_L t_i a_{Lm}}{2a_{Lm}}. \quad (10)$$

So required safe distance of the follower can be gotten as follows:

$$X_R = v_F (t_i + t_r) - \frac{2v_F v_d + v_L^2}{2a_{Lm}} + \frac{v_F^2}{2a_{Fm}} - \frac{v_L t_i}{2} + d. \quad (11)$$

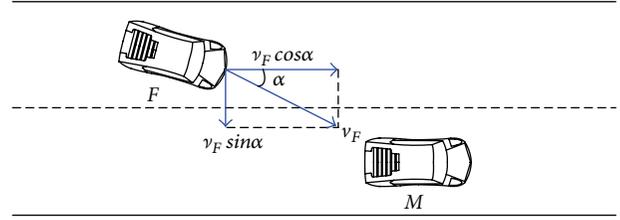


FIGURE 2: Speed decomposition diagram when the vehicle F shifts.

3.2. *Lateral Required Distance.* The lateral required distance is the minimum lateral clearance that the vehicle keeps not to collide with the nearest one on the adjacent lane. Considering drivers' physical and mental characters, even if the lateral clearance is less than the lateral required distance demanded at a certain time, it does not necessarily cause accidents.

As illustrated in Figure 2, α is the angle at which the follower deviates from the target lane. If a large longitudinal clearance exists between the adjacent vehicles, the vehicle F can adjust timely; then the vehicle M will not result in heavy impact on the vehicle F . Moreover, the faster the speed of the follower is, the further the vehicle deviates during the reaction time and the more probably two vehicles collide. Thus the lateral required distance is influenced by both the longitudinal clearance and the vehicle speed.

At a certain time, v_F represents the speed of the vehicle F and α is the maximum deviation angle. In Figure 2, decomposing v_F into lateral and longitudinal directions, $v_F \sin \alpha$ is the lateral velocity component and $v_F \cos \alpha$ is the longitudinal velocity component. At the end of the reaction time t , the lateral displacement component may be less than the lateral clearances, and it may lead to accidents. The lateral required distance can be written as

$$X_S = v_F \cdot t \cdot \sin \alpha, \quad (12)$$

where t is the reaction time, α is the maximum possible deviation angle.

The analysis above reveals that the lateral required distance is linear with the speed of the follower.

Putting formulas (4) and (12) together, when $v_F > v_L$, the relationship between X_R and X_S is shown as follows:

$$X_R = \frac{X_S}{t \cdot \sin \alpha} t_r + \frac{v_d t_i}{2} + \frac{X_S^2}{2a_{Fm} \cdot (t \cdot \sin \alpha)^2} - \frac{v_L^2}{2a_{Lm}} + d. \quad (13)$$

Since formulas (4) and (12) are both monotonically increasing functions of the speed v_F , the lateral required distance X_S increases with X_R increasing.

Similarly, for $v_F = v_L$, the relationship between X_R and X_S is expressed as

$$X_R = \frac{X_S}{t \cdot \sin \alpha} t_r + \frac{v_L^2}{2} \left(\frac{1}{a_{Fm}} - \frac{1}{a_{Lm}} \right) + d. \quad (14)$$

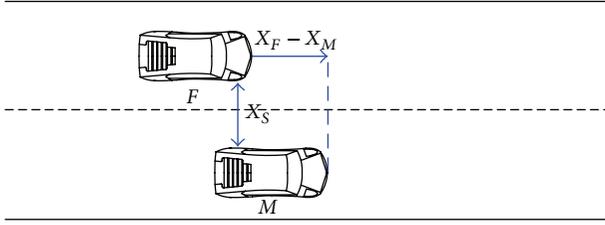


FIGURE 3: Positional relationship diagram between adjacent lane vehicles.

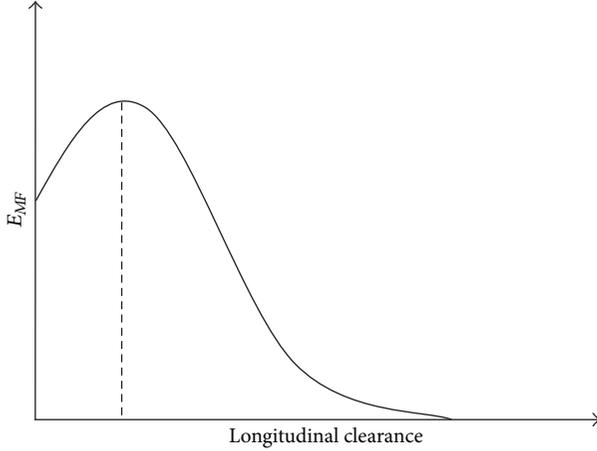


FIGURE 4: Influencing curve on follower of longitudinal clearance.

For $v_F < v_L$, the relationship between X_R and X_S is

$$X_R = \frac{X_S}{t \cdot \sin \alpha} (t_i + t_r) - \frac{X_S}{t \cdot \sin \alpha} \cdot \frac{v_d}{a_{Lm}} - \frac{v_L^2}{2a_{Lm}} + \left(\frac{X_S}{t \cdot \sin \alpha} \right)^2 \cdot \frac{1}{2a_{Fm}} - \frac{v_L t_i}{2} + d. \quad (15)$$

However, in the real-world situation, the lateral clearance between the follower and lateral vehicle should be a little greater than that defined above because the lateral clearance will be enlarged in some cases where the follower decelerates or avoids collisions laterally.

4. Influencing Analysis of Influential Factors

4.1. Influencing Analysis of Longitudinal Clearance. Figure 3 demonstrates that the longitudinal clearance $X_F - X_M$ between the two vehicles has influence on the follower in adjacent lane. The influencing curve is illustrated in Figure 4.

On the assumption that the longitudinal clearance is $(v_F \cdot \cos \alpha - v_M) \cdot t$, it is exactly zero at the end of the reaction time, which is the most probable point where two vehicles collide. The study shows that under the condition where the longitudinal clearance and its speeds are unchanged, the influence on the follower by $X_F - X_M$ follows normal

TABLE 1: Table of safe distance model formula.

Model	Formula
Safe distance model based on headway	$s_1 = v_0 t_d + d$
Safe distance model by braking process	$s_2 = v_0 t_d + \frac{v_0^2}{2a_m} + d$
The model of required safe distance	Models of required safe distance at three following states as above formulas (4), (5), and (11).

distribution. E_{MF} is the influence on the follower by the lateral vehicle then

$$E_{MF} = \frac{1}{\sqrt{2\pi}\sigma} e^{-[X - (v_F \cdot \cos \alpha - v_M) \cdot t]^2 / 2\sigma^2}, \quad 0 \leq X < +\infty, \quad (16)$$

where X is the longitudinal clearance, σ is a undetermined parameter. And at the point that the longitudinal clearance is $(v_F \cdot \cos \alpha - v_M) \cdot t$, the biggest influencing point is $1/\sqrt{2\pi}\sigma$.

4.2. Influencing Analysis of the Speed. In the driving procedure, the faster the speed is, the lighter steering wheel is for the driver, and the vehicle is more probably to produce the lateral deviation, which may directly put much mental pressure on drivers. In this case, the driver of the follower will drive away from the lateral vehicle or decelerate to avoid colliding.

As the lateral vehicle quickly approaches the follower, the following driver will concentrate on it. To a certain extent, the driver's pressure will increase. If the follower is within the range influenced by the lateral vehicle, the lateral clearance and the longitudinal clearance are unchanged, and the larger the speed difference between adjacent vehicles is, the stronger influence the lateral vehicle has on the follower; thus, drivers should decelerate to avoid dangers.

5. Testing and Validation of the Models

5.1. Simulation on the Required Safe Distance. Through the tests of three following states for $v_F > v_L$, $v_F = v_L$ and $v_F < v_L$ the model of required safe distance and two traditional models are compared to analyze the difference. One of the two traditional models is based on headway and another is based on braking process. Formulas of the three models are shown in Table 1.

In order to make a better simulation and analysis compared with the traditional models of safe distance, we assign values to parameters of the formulas, as shown in Table 2.

The relationship between the safe distance and v_F is, respectively, simulated as in Figure 5 ($v_F > v_L$) and Figure 6 ($v_F = v_L$) on the assumption of $v_L = 50 \text{ km} \cdot \text{h}^{-1}$, and Figure 7 describes the relationship between the safe distance and v_L supposing $v_F = 50 \text{ km} \cdot \text{h}^{-1}$. From Figure 5, the result shows that the safe distance based on braking process is too large to decrease the road capacity, and the safe distance based on headway is too small to cause collision in car following procedure. While the model of required safe distance takes the driving characters of vehicles before and

TABLE 2: Calibration table of formula parameters.

Parameter	d (m)	t_d (s)	a_m (m·s ⁻²)	t_i (s)	t_r (s)
Value	4	1.6	7	0.2	0.9

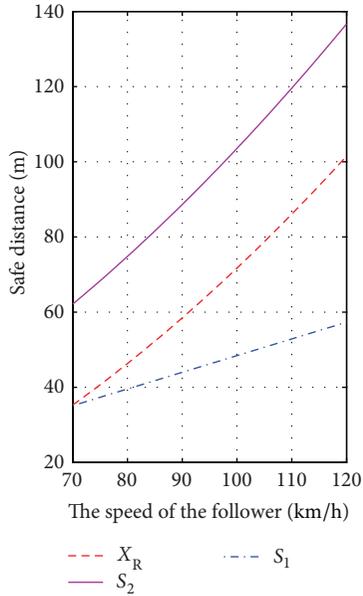


FIGURE 5: Simulation curve for $v_F > v_L$.

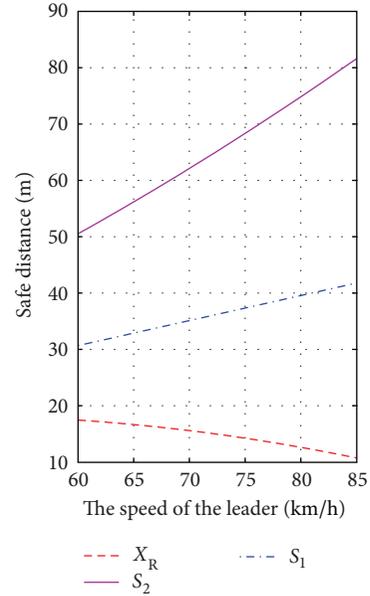


FIGURE 7: Simulation curve for $v_F < v_L$.

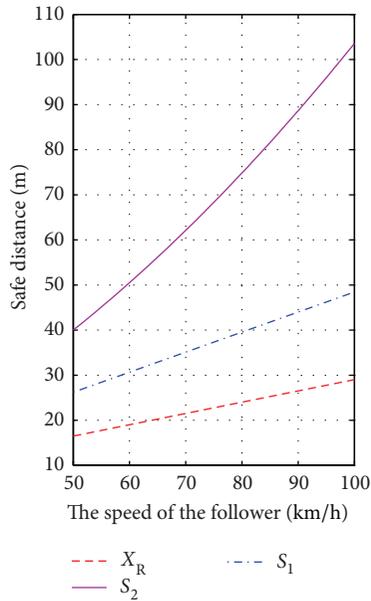


FIGURE 6: Simulation curve for $v_F = v_L$.

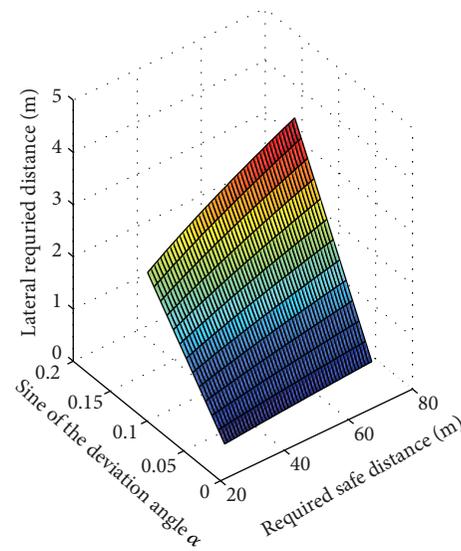


FIGURE 8: Simulation diagram for $v_F > v_L$.

after into consideration, so the follower need not adjust the clearance when keeping the state consistent with the leader, which consequently avoids the situation above. Figures 6 and 7 demonstrates that safe distances based on two traditional models are larger, which ensures two vehicles not to collide but reduces travel efficiency. However, the model of required

safe distance not only ensures safe operation of vehicles but also improves travel efficiency.

5.2. Simulation on the Lateral Safe Distance. The simulation curves of the lateral required distance for $v_F > v_L$, $v_F = v_L$, and $v_F < v_L$ are illustrated in Figures 8, 9, and 10. Figures 8 and 9 are under the condition that the leader's speed is

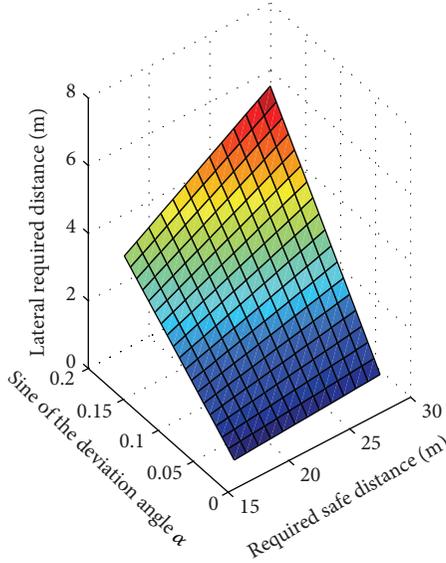


FIGURE 9: Simulation diagram for $v_F = v_L$.

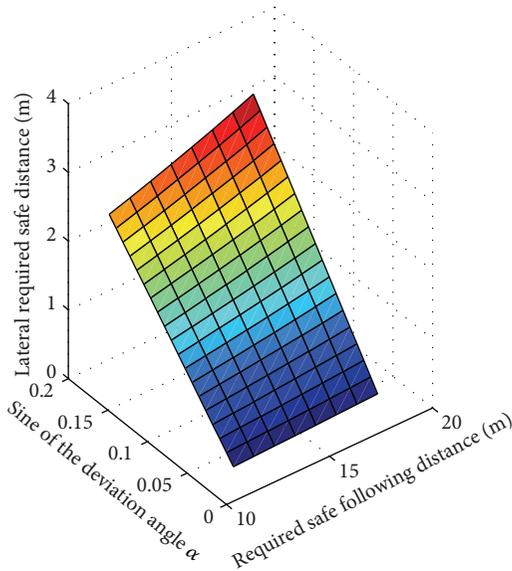


FIGURE 10: Simulation diagram for $v_F < v_L$.

$50 \text{ km}\cdot\text{h}^{-1}$, and Figure 10 is assuming that the leader's speed is $70 \text{ km}\cdot\text{h}^{-1}$. As shown in the figures, the lateral required distance increases with the deviation angle increasing, and with X_R increasing, the changing amplitude will increase; similarly, the required safe distance also increases with X_R increasing. But it almost has no impact on X_S when α is small. If the angle α is large, this needs a larger lateral required distance changing largely with v_F increase, and this is also consistent with the actual situation.

6. Conclusion

This paper presents the concept of car-following required safe distance, which is based on molecular dynamics. We develop

the models of required safe distance and lateral safe distance and apply mathematical reasoning method to get their expressions. And the influence on lateral required distance is analyzed from the aspects of longitudinal clearance and the speed. For better understanding, we make the comparison between the model of required safe distance and two traditional models, which reveals that the model of required safe distance shows the priority in driving safety and efficiency. Finally, upon the completion of the tests and validation using the actual survey data in crossings, it is pointed out that the model of car-following required safe distance based on molecular dynamics is significant for vehicles driving safely and efficiently, which can provide effective theoretical basis for the adaptive cruise control system.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgments

The paper is supported by the National Natural Science Foundation of China (51178231) and Provincial Natural Science Foundation of Shandong (ZR2012EEL128). The authors would like to thank the referees.

References

- [1] T. Toledo, "Driving behaviour: models and challenges," *Transport Reviews*, vol. 27, no. 1, pp. 65–84, 2007.
- [2] D. C. Gazis, R. Herman, and R. W. Rothery, "Nonlinear follow-the-leader models of traffic flow," *Operation Research*, vol. 9, no. 4, pp. 545–567, 1961.
- [3] G. F. Newell, "Nonlinear effects in the dynamics of car following," *Operation Research*, vol. 9, no. 2, pp. 209–229, 1961.
- [4] P. G. Gipps, "Behavioral car-following model for computer simulation," *Transportation Research B*, vol. 15, no. 2, pp. 105–111, 1981.
- [5] R. Wiedemann, *Simulation of Road Traffic in Traffic Flow*, University of Karlsruhe, Karlsruhe, Germany, 1974.
- [6] C. Kikuchi and P. Chakroborty, "Car following model based on a fuzzy inference system," *Transportation Research Record*, vol. 1194, pp. 82–91, 1992.
- [7] M. F. Aycin and R. F. Benekoha, "Comparison of car-following models for simulation," in *Proceedings of the 78th Annual Meeting of Transportation Research Board (TRB '99)*, 1999.
- [8] M. T. Parker, "The effect of heavy goods vehicles and following behaviour on capacity at motorway roadwork sites," *Traffic Engineering & Control*, vol. 37, no. 9, pp. 524–531, 1996.
- [9] H. Peter, "A car-following model for urban traffic simulation," *Traffic Engineering & Control*, vol. 39, no. 5, pp. 300–302, 1998.
- [10] Y. L. Zhang, J. E. Clark, and E. C. James, "A multivegime approach for microscopic traffic simulation," in *Proceedings of the 77th Annual Meeting of Transportation Research Board (TRB '98)*, Washington, DC, USA, 1998.
- [11] S. Druitt, "An introduction to microsimulation," *Traffic Engineering & Control*, vol. 39, no. 9, pp. 480–483, 1998.
- [12] E. R. Boer, "Car following from the driver's perspective," *Transportation Research F*, vol. 2, no. 4, pp. 201–206, 1999.

- [13] D. B. Fambro, R. J. Koppa, D. L. Picha, and K. Fitzpatrick, "Driver perception-brake response in stopping sight distance situations," *Journal of the Transportation Research Board*, vol. 1628, pp. 1–7, 1998.
- [14] K. I. Ahmed, *Modeling driver's acceleration and lane changing behaviors [Ph.D. thesis]*, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 1999.
- [15] G. S. Gurusinghe, T. Nakatsuji, Y. Tanaboriboon, K. Takahashi, and J. Suzuki, "A car-following model incorporating excess critical speed concept," *Journal of Eastern Asia Society for Transportation Studies*, vol. 4, no. 2, pp. 171–183, 2001.
- [16] P. Ranjitkar, T. Nakatsuji, G. Gurusinghe, and Y. Azuta, "Car-following experiments using RTK GPS and stability characteristics of followers in platoon," in *Proceedings of the 7th International Conference on Applications of Advanced Technology in Transportation*, pp. 608–615, Boston, Mass, USA, 2002.
- [17] Y. Uchiyama, K. Ebe, A. Kozato, T. Okada, and N. Sadato, "The neural substrates of driving at a safe distance: a functional MRI study," *Neuroscience Letters*, vol. 352, no. 3, pp. 199–202, 2003.
- [18] J. S. Kong, F. Guo, and X. P. Wang, "A vehicle rear-end anti-collision method base on safety distance model," *Automotive Electronics*, vol. 24, no. 11, pp. 251–253, 2008.
- [19] B. Gunay, "Car following theory with lateral discomfort," *Transportation Research B*, vol. 41, no. 7, pp. 722–735, 2007.

Research Article

Urban Traffic Signal System Control Structural Optimization Based on Network Analysis

Li Wang, Xiao-ming Liu, Zhi-jian Wang, and Zheng-xi Li

Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Beijing 100144, China

Correspondence should be addressed to Li Wang; li.wang@ncut.edu.cn

Received 14 August 2013; Revised 24 October 2013; Accepted 3 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Li Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Advanced urban traffic signal control systems such as SCOOT and SCATS normally coordinate traffic network using multilevel hierarchical control mechanism. In this mechanism, several key intersections will be selected from traffic signal network and the network will be divided into different control subareas. Traditionally, key intersection selection and control subareas division are executed according to dynamic traffic counts and link length between intersections, which largely rely on traffic engineers' experience. However, it omits important inherent characteristics of traffic network topology. In this paper, we will apply network analysis approach into these two aspects for traffic system control structure optimization. Firstly, the modified C-means clustering algorithm will be proposed to assess the importance of intersections in traffic network and furthermore determine the key intersections based on three indexes instead of merely on traffic counts in traditional methods. Secondly, the improved network community discovery method will be used to give more reasonable evidence in traffic control subarea division. Finally, to test the effectiveness of network analysis approach, a hardware-in-loop simulation environment composed of regional traffic control system, microsimulation software and signal controller hardware, will be built. Both traditional method and proposed approach will be implemented on simulation test bed to evaluate traffic operation performance indexes, for example, travel time, stop times, delay and average vehicle speed. Simulation results show that the proposed network analysis approach can improve the traffic control system operation performance effectively.

1. Introduction

Advanced urban traffic signal control systems normally use multilevel hierarchical control mechanism to simplify the network control process. In this mechanism, several key intersections of traffic network will be selected and the network will be divided into several control subareas in which the signal of intersections will be optimized according to the traffic states variation of key intersections. In 1971, Walinchus [1] firstly built the concept of "traffic control subarea." Stockfisch [2], Pinell et al. [3] and Kell and Fullerton [4] proposed the guideline for computer signal system selection based on intersection traffic state analysis, road segment length, vehicle arrival rate, and so forth. Yagoda et al. [5] and Chang [6] defined the traffic control index and threshold value of algorithm for traffic control subarea division. However, they have not considered the characteristics of dynamic traffic network topology, and typical advanced traffic signal

control systems like SCOOT [7] and SCATS [8] normally executed the "control subarea division" and "key intersection selection" based on traffic counts and link length between intersections. In practice, this configuration process relies on the experience of traffic engineers [9, 10]. It is difficult to assure the reasonability and effectiveness of the system control structure configuration because of lack of reliable theoretical support.

Generally traffic network can be represented as a weighted network graph, where an intersection corresponds to a node, a road segments to a link, and traffic flow parameters of the segments (like traffic flow, link length, travel time, etc.) to link weight. This type of weighted network has typical topology characteristics like nonhomogeneous and scale-free. However traditional traffic signal system structure configuration does not consider these characteristics. It is necessary to apply traffic network analysis method to reduce the uncertainty of traffic network control configuration and

ensure the process of key intersection selection and subarea division.

Recently, there are several papers and research results in this area. In aspect of key node assessment, several approaches can be referred, for example, the betweenness method, node deletion method, node contraction method, network nodes nearness and neighborhood key degrees assessment, and so forth [11, 12]. In aspect of network subarea division, it is found be similar with the network community discovery. Although there are several community discovery algorithms have been developed, to the best knowledge of authors, there is still lack of direct findings on how to realize the control structure optimization for traffic signal networks. For example, K-L algorithm [13] needs to know the size of the two communities; spectrum bisection method [14] can only divide the network into odd communities; G-N algorithm [15] cannot confirm the suitable iteration process easily while amount of the community structure is unknown; and W-H algorithm [16] is mainly used to dig the community structure that contains designated nodes.

In this paper, we will improve the network analysis approaches and apply them into traffic signal control field, especially using modified C-means clustering method for node assessment and improved community discovery algorithm for control subarea division. In Section 2, traffic network will be abstracted as network graph and typical indicators of network graph are introduced. In Section 3, traffic signal intersection importance assessment will be implemented by the C-means clustering approach. And then community discovery algorithm will be applied to the implementation of traffic signal network subarea division in Section 4. Finally, experiment conducted by SCOOT system and VISSIM simulation platform will be constructed to compare and verify the effectiveness of network analysis approach based on real traffic data of Beijing city.

2. Traffic Network Modeling

As mentioned above, urban traffic network can be abstracted as a weighted graph in which the intersection corresponds to a node, the road section to a link, and the link length to the link weight. Based on this type of weighted network graph, the following indicators will be used in this paper.

- (1) Node connectivity (i.e., node degree) $\langle k \rangle$: the number of links connecting to node k .
- (2) Node betweenness $C(i)$: the ratio of the number of the shortest path crossing the node i to all the shortest paths between the nodes in the network:

$$C(i) = \sum_{s \neq t \neq i} \left[\frac{g_{st,i}}{n_{st}} \right], \quad (1)$$

where $g_{st,i}$ indicates the number of the shortest paths through node i between node s and node t and n_{st} is the number of all the shortest paths between node s and node t .

- (3) Network module degree Q :

$$Q = \text{Tr}(E) - \|E^2\|, \quad (2)$$

where E is a symmetrical matrix with elements E_{ij} representing the ratio of the number of links connecting community i and j to the number of whole network links, and $\|E^2\| = \sum_i \sum_{j,k} E_{ij} E_{ik}$; $\text{Tr}(E)$ is the trace of matrix E .

3. Traffic Network Node Importance Assessment Based on Clustering

It is known that weighted network normally has typical non-homogeneous and scale-free characteristics, which means that the importance of each node in traffic network is different. Existing network nodes importance assessment methods are basically derived from graph theory and need to be improved for traffic control application. In traffic engineering fields, we normally describe the signal intersection as “key,” “important,” “general,” “unimportant,” and other types. Therefore, we select C-means clustering method [17] to classify the traffic intersections of network.

C-Means clustering method can be used to classify the nodes of network into different categories and describe the degree of node in each category. The basic ideas of the clustering process can be described as follows: assume the cluster centers and calculate distance from each node to the centers; adjust the cluster centers according to the sum of distance until the clustering process meets minimize conditions.

Specifically, suppose that there are n samples $X = \{x_1, x_2, \dots, x_n\}$, where $x_k = (x_{k1}, x_{k2}, \dots, x_{km})^T \in R^m$, $k = 1, \dots, n$ is the eigenvector of sample x_k ; c , $2 \leq c \leq n$ is the number of categories that all samples will be divided into; $P = (p_1, p_2, \dots, p_c) (\in R^{m \times c})$ constitutes a cluster prototype matrix, where $p_i = (p_{i1}, p_{i2}, \dots, p_{im})^T \in R^m$ is the i th clustering prototype; $U = (\mu_{ik})_{c \times n} \in R^{c \times n}$ is a matrix with elements $\mu_{ik} \in [0, 1]$ meaning the degree of sample x_k to p_i ; then the objective function of clustering can be expressed as

$$\min J_b(U, P) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})^b (d_{ik})^2, \quad (3)$$

$$\text{s.t.} \quad \sum_{i=1}^c \mu_{ik} = 1, \quad k = 1, 2, \dots, n,$$

where b is a flexible parameter in $[1.5, 2.5]$ and d_{ik} represents Euclidean distance between x_k and p_i .

It is obvious that the type of data samples is important for clustering. In traffic network, the data samples for assessments of intersection not only need to reflect the static topology characteristics, but also need to contain the information of dynamic traffic flow. In this paper we select three typical indicators to form the data samples of each network node.

- (1) Node connectivity $\langle k \rangle$ reflects the characteristics of the topology of the node.
- (2) Node betweenness $C(i)$ reflects the control ability of nodes among the network traffic flow dissemination.
- (3) Node traffic flow in rush hour reflects the dynamic traffic counts variation and capacity of the intersections.

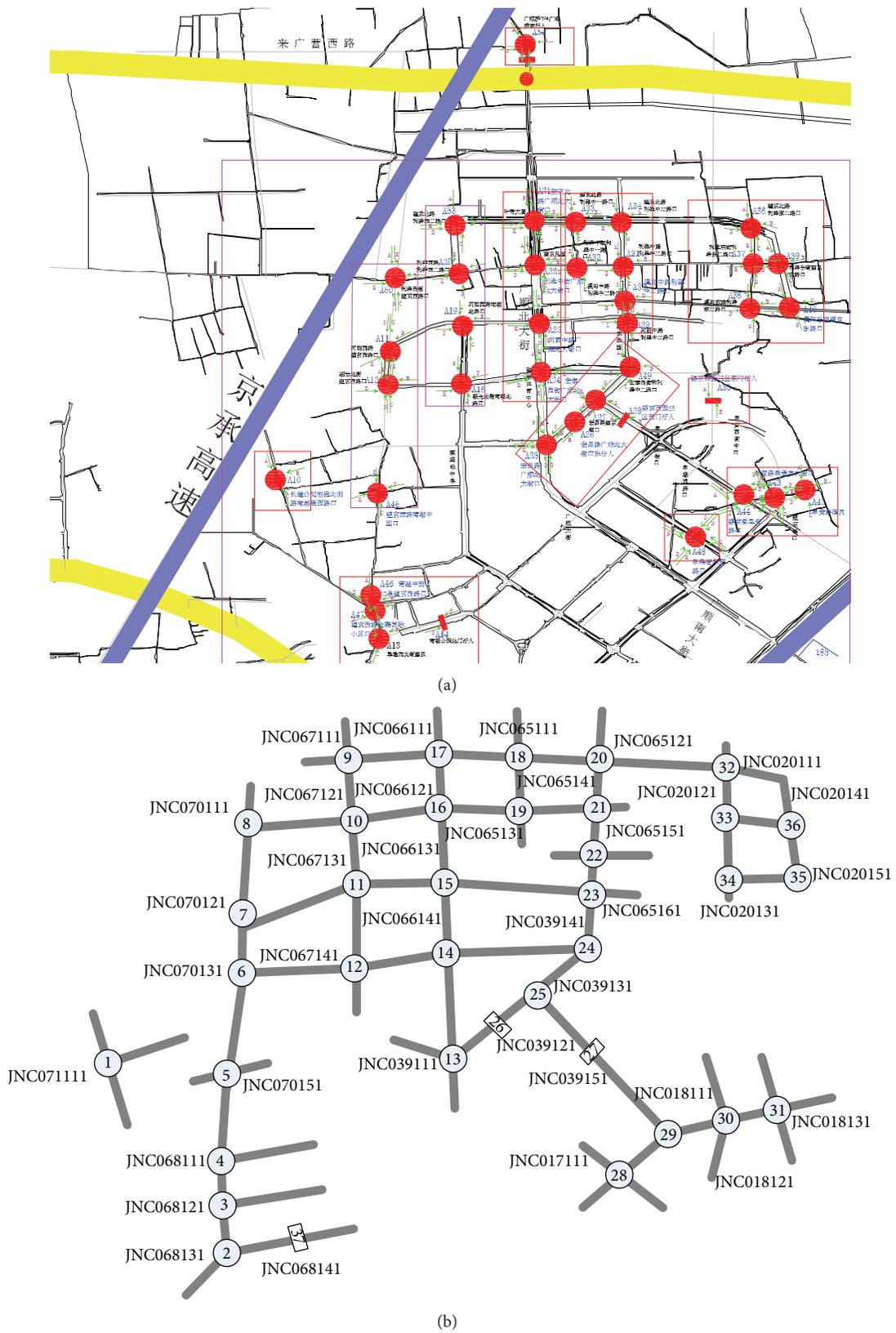


FIGURE 1: Traffic control network of wangjing area in Beijing. (a) Map of control network. (b) Topology graph of traffic signal network.

TABLE 1: Intersection importance evaluation index (18:00-19:00, October 17, 2011).

No. of node	Intersection name	Node degree	Node traffic flow	Node betweenness
1	Nanhuxiyuan South	2	1933	0.0023
2	Futongxidajie South	2	3123	0.0116
3	Wangjingxilu-Jinyuguoji	2	2971	0.0105
4	Nanhu-Zhongyuan Ertiao West	2	3234	0.0128
5	Wangjingxilu-Nanhuzhongyuan	3	2557	0.0302
6	Huguang Beijie West	3	3410	0.0372
7	Heyin West	3	1918	0.0418
8	Wangjing Xilu North	3	2249	0.0221
9	Lizexierlu	3	1311	0.0035
10	Lizexi Street	3	1676	0.0279
11	Nanhu Beilu North	3	3573	0.0232
12	Huguangbei Street	3	7067	0.0232
13	Huguangzhongjie East	3	5012	0.0256
14	Hongtaixijie West	4	6828	0.0383
15	Heyinzhonglu West	4	6642	0.0476
16	Lizezhongjie West	4	6325	0.101
17	Wangjing Beilu-Guangshunbeidajie	3	5034	0.0604
18	Lizezhongyilu	3	3197	0.0139
19	Lizezhongyilu South	3	2154	0.0116
20	Lizezhongerlu	3	3812	0.0186
21	Lizezhongjie	4	2632	0.0453
22	Xiyangzhonglu Mid	4	1173	0.0778
23	Heyinzhong	3	2213	0.1626
24	Hongtaixi Street	2	3384	0.0383
25	Hongchang Road	3	1588	0.0581
26	Hongchang Peds	2	1634	0.0767
27	Wangjingxiyuansanqu Peds	2	1856	0.072
28	Furongjie-Fuanlu	1	1188	0.0023
29	Wangjingjie-Fuan	3	3249	0.1463
30	Fuanlu-Futongxidajie	2	2855	0.0778
31	Fuanlu-Guofengbeijing	1	813	0.0023
32	Lizedongerlu	3	1587	0.0232
33	Lizedong Street	4	1070	0.1498
34	Lizedongerlu South	3	1645	0.0581
35	Xiyang East	2	1177	0.0035
36	Wangjing Beilu East	3	842	0.0302

Then, the node assessment algorithm can be described by the following steps.

Step 1 (parameter settings). The data sample x_k includes three items: node degree $\langle k \rangle$, node betweenness $C(i)$, and rush hour traffic counts. Let category factor be $c = 3, 4$ corresponding to different number of node categories. To realize the advantages of clustering, let index $b = 2$, iteration termination threshold $\varepsilon = 0.001$, initial iteration counter $t = 1$ and then produce a prototype model of clustering $P^{(t)}$.

Step 2. Calculate the partition matrix $U^{(t)} = \{\mu_{jk}^{(t)}\}$. For all i, k , if $\exists d_{ik}^{(t)} > 0$, there are $\mu_{ik}^{(t)} = \{\sum_{j=1}^c [(d_{ik}^{(t)}/d_{jk}^{(t)})^{2/(b-1)}]\}^{-1}$; if $\exists i, r$ such that $d_{ir}^{(t)} = 0$, there is $\mu_{ir}^{(t)} = 1$; and when $j \neq r$, there is $\mu_{ij}^{(t)} = 0$.

Step 3. Update cluster prototype matrix $P^{(t+1)} = \{p_i^{(t+1)}\}$, $p_i^{(t+1)} = (\sum_{k=1}^n (\mu_{ik}^{(t)})^b \cdot x_k) / (\sum_{k=1}^n (\mu_{ik}^{(t)})^b)$.

Step 4. If $\|P^{(t)} - P^{(t+1)}\| < \varepsilon$, the algorithm stops and outputs partition matrix U and cluster prototype P ; otherwise let $t = t + 1$ and turn to (2). Normally $\|\cdot\|$ can use F -norm; that is,

$$\|P^{(t)} - P^{(t+1)}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^c (p_{ij}^{(t)} - p_{ij}^{(t+1)})^2}.$$

4. Traffic Control Subarea Division Based on Community Discovery

In this paper, using the network module degree in (2) as assessment indicator, we modify Newman cohesion community discovery [18] approach for traffic control subarea division. The division algorithm can be described as follows.

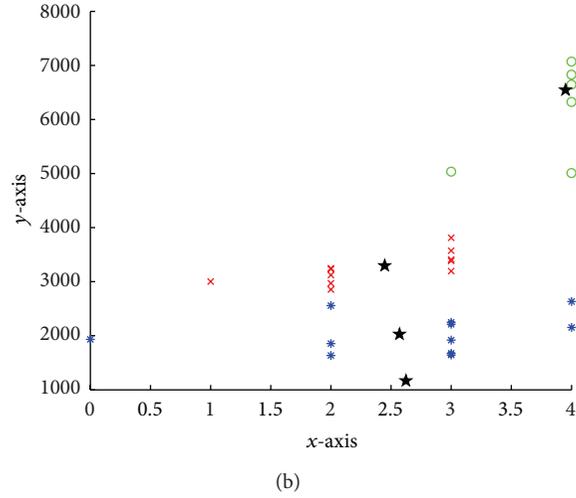
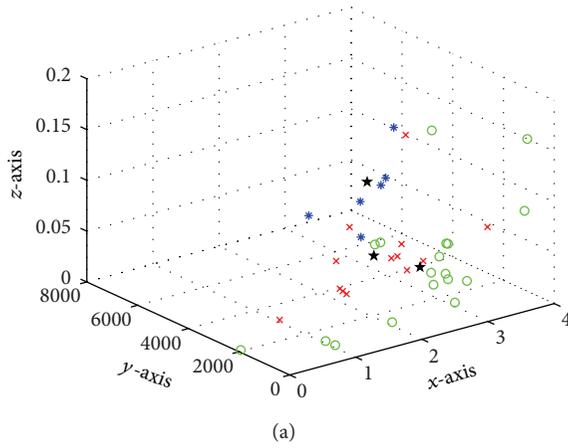


FIGURE 2: C-means result with $c = 3$. (a) Node cluster distribution, (b) Plane projection of betweenness and traffic flow indicators.

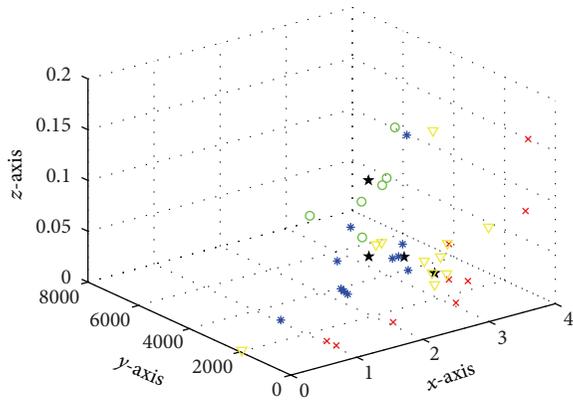


FIGURE 3: C-means result with $c = 4$.

Step 1. Divide the traffic network diagram into N communities which means that each node represents one community. The initial E_{ij} and the sum a_i of its rows of matrix E satisfy

$$a_i = \frac{k_i}{2m},$$

$$E_{ij} = \begin{cases} \frac{1}{2m}, & \text{if there exists link between } i \text{ and } j \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where k_i is the degree of node i and m is the total number of links in the network.

Step 2. Merge the communities that contain links and record the increment of the network module degree $\Delta Q = 2(E_{ij} - a_i a_j)$. Merging should be executed along the direction that can make ΔQ toward maximum values. Then, renew E_{ij} and add the ranks and rows related to the community i and j .

Step 3. Repeat Step 2 until the whole network merges into one community.

Step 4. Consider the dendrogram of community structure and select the max module degree. Then, determine the optimal division of network community.

5. Experiment

5.1. Traffic Network Modeling. Choose Wangjing area in Beijing illustrated in Figure 1(a) as test bed to evaluate the effectiveness of above proposed traffic control network structure optimization methods. The abstracted weighted network graph can be referred to in Figure 1(b).

5.2. Experiment Traffic Data Collection. On the basis of traffic raw data in October 17, 2011 from Beijing Traffic Management Bureau, we calculate the indexes of node degree, nodes betweenness, and node rush hour (18:00-19:00) traffic flow as shown in Table 1. The calculation process follows steps as follows.

- (1) Calculate the node degree of each node according to the weighted network topology.
- (2) Calculate the weight of sides of each segment according to the link length, the distance matrix D between the nodes, and the shortest distance R . Then acquire the number of each node in the shortest distance matrix to calculate the betweenness of each node.
- (3) Count the vehicle flow data in rush hours and calculate the flow rate for each intersection.

5.3. Intersection Importance Assessment. Apply C-means clustering algorithm described in Section 3 to get network nodes importance clustering. Figures 2 and 3 show the cluster distribution as $c = 3$ and $c = 4$, respectively, in which x -axis

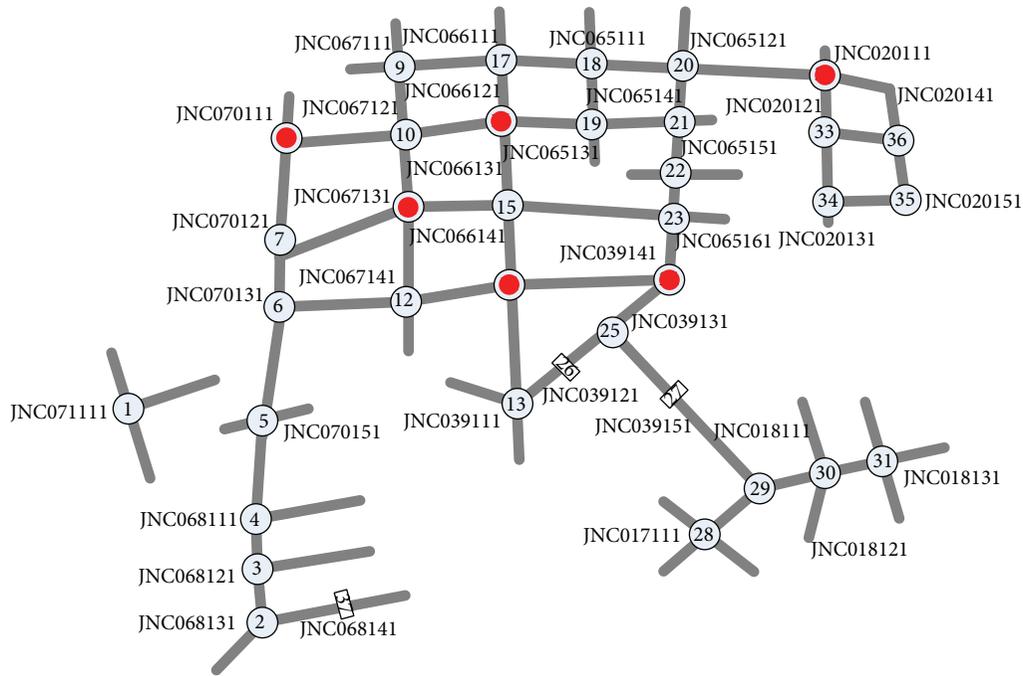


FIGURE 4: Key nodes of wangjing area traffic network.

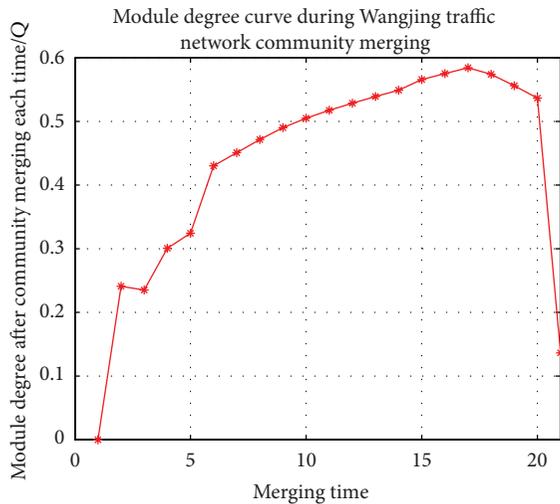


FIGURE 5: Module degree curve during Wangjing traffic network community emerging.

means the node degree, y -axis means node rush hour traffic flow (veh/hour), and z -axis means the node betweenness. Figure 2(b) shows the plane projection of node betweenness and traffic flow indicators when $c = 3$.

Practically traffic signal control systems generally take traffic nodes as three categories: key node, important node, and general node. For convenience, take 3 categories to cluster the network intersections. Figures 2(a) and 2(b) show that:

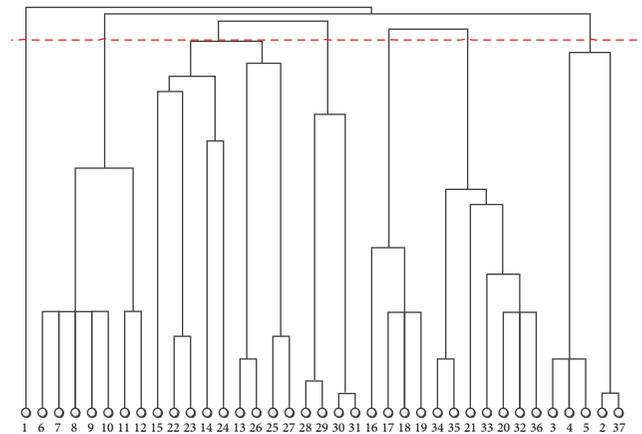


FIGURE 6: Community division of Wangjing traffic network.

- (1) the node degree of cluster centers is almost 2.5 or 3.5 which means that the key node should have node degree of 3 or 4;
- (2) node betweenness of cluster center nodes are 0.018, 0.023, and 0.042, respectively;
- (3) rush hour traffic flow of cluster center nodes is 1600 veh/hour, 3100 veh/hour, and 6500 veh/hour respectively.

It is very close to the key node selection criteria in traffic engineering practice. We can choose node 8, node 11, node 14,

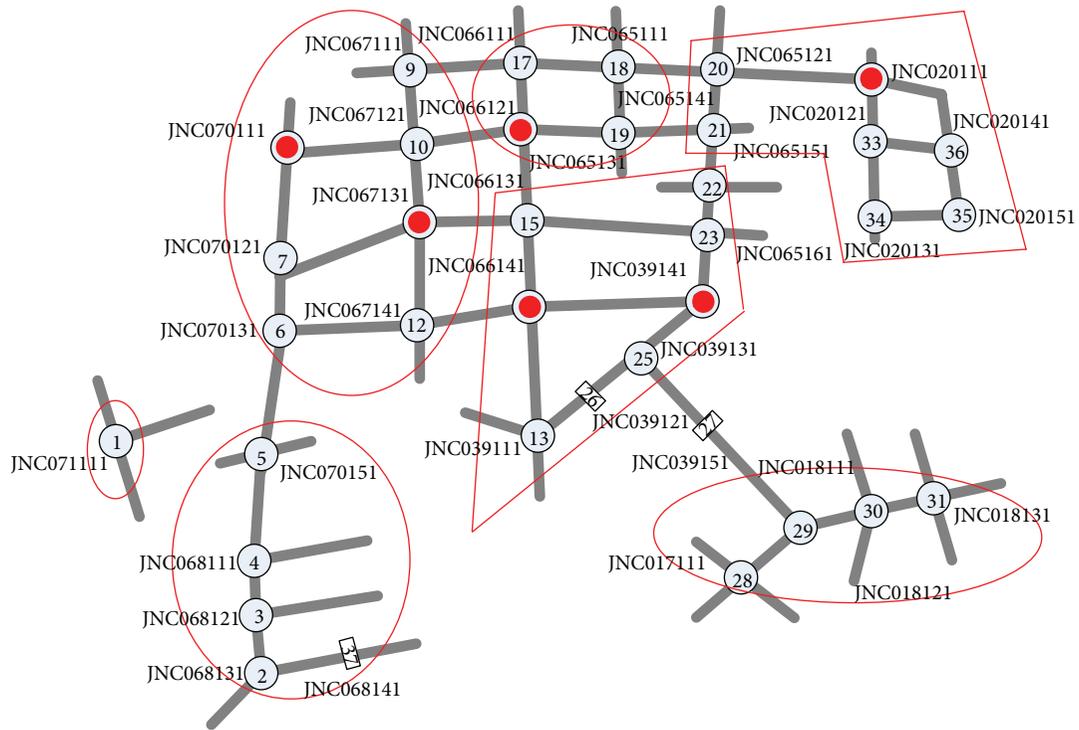


FIGURE 7: Control subarea division of Wangjing network based on community discovery.

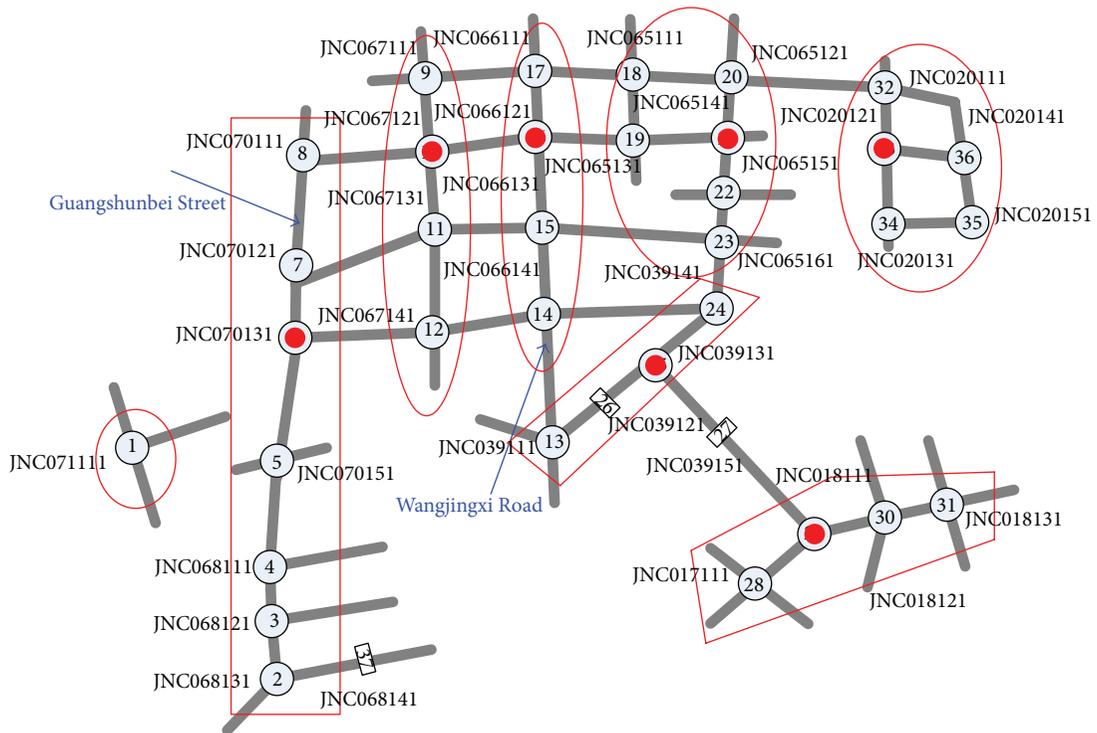


FIGURE 8: Control subarea division of Wangjing network based on distance and traffic flow.

node 16, node 24, and node 32 as the key nodes as big points shown in Figure 4 and bold in Table 1.

5.4. Traffic Control Subarea Division. Complete the community merging process mentioned in Section 4 and record the value of module degree Q at each step. The variation curve is shown in Figure 5 in which the horizontal axis represents the combination times and the vertical axis represents the module degree.

From the module degree change curve in Figure 5, we can see that the 17th calculation corresponds to the maximum value of network module degree $Q = 0.5843$, which means that when $n = 17$, we can get the best community discovery division of network. The result of proposed community discovery is shown as Figure 6 and the corresponding division of traffic control subarea of Wangjing network is shown as in Figure 7. It is necessary to mention that there are two key traffic intersections selected in Section 5.3 locating in the same control subarea. For this situation, we need to decide the right one according to the requirement of practice.

5.5. Traffic Simulation. To test the effectiveness of network analysis approach, a hardware-in-loop simulation environment has been built as the following steps: (1) construct the traffic network simulation of Wangjing area based on VISSIM simulation platform and (2) implement online real-time control on road network with PC-SCOOT system. To avoid the influence of signal control algorithm to different control structure configuration, the same traffic signal control algorithm has been applied in simulation with traditional configuration based on distance between intersections and node traffic flow shown as Figure 8 and proposed configuration by this paper as Figure 7.

Configure the simulated road network using PC-SCOOT system as the following steps: (1) set the SCOOT system structure by traditional method and proposed method separately; (2) assign the control subareas and the intersections Modern and IP address utilizing the DBAS commands; (3) make configuration to each intersection in detail; (4) set the simulated intersection model with IP address, phase information, and detector information of signal controller utilizing VB software based on VISSIM COM interface; (5) set traffic detectors at Wangjingxi Road and Guangshunbei Street (as shown in Figure 8) in VISSIM model and input traffic data. Then, three different traffic demand inputs: 300 veh/hour/lane, 600 veh/hour/lane, and 1000 veh/hour/lane, have been loaded separately into the network to simulate the traffic state in low hour, flat hour and rush hour of one day.

Tables 2 and 3 show the traffic operation performance comparison of SCOOT system based on traditional approach and proposed network approach. In Table 2, the SCOOT system based on proposed method achieves better operation performance during low hour, flat hour, and rush hour. Specially, the travel time from North to South of Wangjingxi Road average decreased by 3.1% and from South to North of Wangjingxi Road average decreased by 7.3%; the travel time

TABLE 2: Average travel time of Wangjingxi Road and Guangshunbei Street (s).

	Wangjingxi Road		Guangshunbei Street	
	N to S	S to N	S to N	N to S
Low hour				
Traditional method	293	213	254	204
Proposed method	292	191	189	210
Decrease	-0.34%	-10.32%	-25.59%	2.94%
Flat hour				
Traditional method	254.76	227.64	322.92	191.50
Proposed method	254.72	212.16	221.04	198.29
Decrease	-0.0157%	-6.8%	-31.54%	3.55%
Rush hour				
Traditional method	292.46	216.57	339.24	210.51
Proposed method	267.43	205.96	218.31	205.96
Decrease	-8.56%	-4.9%	-35.65%	-2.16%
Average decrease	-3.1%	-7.3%	-31.41%	-1.35%

from South to North of Guangshunbei Street decreased by 31.41% and from North to South decreased 1.35%.

In Table 3, traffic network performance has been improved based on proposed configuration method, in which the network average stop delay decreased by 12.5%, the network average stops decreased by 19.61% and the network average speed increased by 11.74%. It is obviously verifying the availability and feasibility of the proposed approach in this paper.

6. Conclusion and Future Research

Traditional key intersection selection and control subarea division method for advanced traffic signal control system do not consider the inherent characteristic of topology structure of the traffic network. In this paper, we use three different indexes based on network analysis to integrate the network topology structure indicators with traffic flow conditions which reflect the characteristic of the network from both static and dynamic aspects. It is more reasonable in theoretical analysis and traffic engineering in practice. Moreover, the advanced community discovery methods are applied to establish or adjust the traffic network control structure before the traffic signal control system come into operation where the proposed approach does not change their original system control strategy and signal optimization algorithm. In our traffic simulation research, compared to traditional method, the proposed control structure optimization approach will obviously decrease the travel time and stops of the test area. It really provides strong support for the application of this network control structural optimization in traffic signal network control.

However, the following issues still need to be aware of in future research: how to achieve more reasonable concordance between key node assessment and control subarea division in practice; whether the dynamic division of subcontrol area

TABLE 3: Wangjing traffic network simulation performances.

	Average stop delay (s)	Average stops	Average delay (s)	Average speed (km/h)
Low hour				
Traditional method	78.115	2.192	113.105	34.221
Proposed method	74.518	2.203	109.349	34.663
Decrease	-4.6%	-0.50%	-3.32%	1.29%
Flat hour				
Traditional method	99.04	4.574	185.926	27.222
Proposed method	89.91	4.056	144.688	30.878
Decrease	-24.76%	-51.84%	-41.19%	27.33%
Rush hour				
Traditional method	117.599	7.454	278.869	21.689
Proposed method	108.882	7.025	246.971	23.565
Decrease	-7.41%	-5.76%	-11.44%	8.65%
Average decrease	-12.5%	-19.61%	-19.42%	11.74%

is feasible and how to actualize this process; whether the stability of the network traffic control system will be affected after network analysis approach is applied.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper was supported by China National Natural Science Foundation (51308005), Scientific Research Project of Beijing Education Committee (PXM2013_014212_000029, 000031, 000028, KM20111009011), the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institution (CIT&TCD201304002), and China National 863 High Technology Plan (2012AA112401).

References

- [1] R. J. Walinchus, "Real-time network decomposition and sub network interfacing," *Highway Research Record*, pp. 20–28, 1971.
- [2] C. R. Stockfisch, "Guidelines for computer signal system selection in urban areas," *Traffic Engineering*, vol. 43, no. 3, pp. 30–63, 1972.
- [3] C. Pinnell, J. J. DeShazo Jr., and R. L. Wilshire, "Areawide traffic control systems," *Traffic Engineering & Control*, vol. 45, no. 4, pp. 16–20, 1975.
- [4] J. H. Kell and I. J. Fullerton, *Manual of Traffic Signal Design*, Prentice Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1991.
- [5] H. N. Yagoda, E. H. Principe, C. E. Vick, and B. G. Leonard, "Subdivision of signal systems into control areas," *Traffic Engineering*, vol. 43, no. 12, pp. 42–46, 1973.
- [6] E. C.-P. Chang, "Evaluation of interconnected arterial traffic signals," *Transportation Planning Journal*, vol. 15, no. 1, pp. 137–156, 1986.
- [7] P. B. Hunt, D. I. Robertson, and R. I. Winton, "SCOOT—a traffic responsive method of coordinating signals," TRL Laboratory Report 1014, 1981.
- [8] ATS Technology (HK) Limited, "SCATS Adaptive Control Introduction," pp. 17–21, 1992.
- [9] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.
- [10] D. Bretherton, M. Bodger, N. Baber, and S. Controls, "SCOOT—the future [urban traffic control]," in *Proceeding of 12th IEEE International Conference on Road Transport Information and Control*, pp. 301–306, April 2004.
- [11] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [12] A.-M. Kermarrec, E. Le Merrer, B. Sericola, and G. Trédan, "Second order centrality: distributed assessment of nodes criticality in complex networks," *Computer Communications*, vol. 34, no. 5, pp. 619–628, 2011.
- [13] B. W. Kernighan and S. Lin, "A efficient heuristic procedure for partitioning graphs," *The Bell Systems Technical Journal*, vol. 49, pp. 291–307, 1970.
- [14] A. Pothen, H. D. Simon, and K.-P. Liou, "Partitioning sparse matrices with eigenvectors of graphs," *SIAM Journal on Matrix Analysis and Applications*, vol. 11, no. 3, pp. 430–452, 1990.
- [15] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [16] F. Wu and B. A. Huberman, "Finding communities in linear time: a physics approach," *European Physical Journal B*, vol. 38, no. 2, pp. 331–338, 2004.
- [17] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 21, no. 3, pp. 193–199, 2002.
- [18] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, vol. 69, no. 6, Article ID 066133, 5 pages, 2004.

Research Article

Multi-Objective Optimization of Traffic Signal Timing for Oversaturated Intersection

Yan Li, Lijie Yu, Siran Tao, and Kuanmin Chen

School of Highway, Chang'an University, The Middle Section of the Second Ring Road (South Part), Xi'an 710064, China

Correspondence should be addressed to Yan Li; lyan@chd.edu.cn

Received 31 August 2013; Revised 17 November 2013; Accepted 2 December 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Yan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For the purpose of improving the efficiency of traffic signal control for isolate intersection under oversaturated conditions, a multi-objective optimization algorithm for traffic signal control is proposed. Throughput maximum and average queue ratio minimum are selected as the optimization objectives of the traffic signal control under oversaturated condition. A simulation environment using VISSIM SCAPI was utilized to evaluate the convergence and the optimization results under various settings and traffic conditions. It is written by C++/CRL to connect the simulation software VISSIM and the proposed algorithm. The simulation results indicated that the signal timing plan generated by the proposed algorithm has good efficiency in managing the traffic flow at oversaturated intersection than the commonly utilized signal timing optimization software Synchro. The update frequency applied in the simulation environment was 120 s, and it can meet the requirements of signal timing plan update in real filed. Thus, the proposed algorithm has the capability of searching Pareto front of the multi-objective problem domain under both normal condition and over-saturated condition.

1. Introduction

Current traffic signal control technologies usually have lower efficiency when the saturation degree is high. Many methods were proposed to improve the efficiency of traffic signal under traffic conditions with high saturation degree [1–3]. However, very little of those methods can be widely utilized for the reason of the requirements or limitations of those methods [4]. In order to establish one specific traffic control method with the capability of dealing with oversaturated condition, the nature of traffic signal control should be discovered.

Traffic signal control methods try to establish the connection between observed traffic parameters, like counts, delay and queue length, with traffic signal parameters, such as phase sequence, cycle length, and split. In this way, various traffic signal optimization algorithms can seek for the values of traffic signal parameters to obtain the optimal values of one or several traffic parameters by considering traffic signal parameters as independent variables under specific traffic condition. This principle is followed by almost all the commonly used traffic control optimization methods, which include TRRL (Transport and Road Research Laboratory) method [5], HCM (Highway Capacity Manual) method [6]

and adaptive traffic control software like SCATS (Sydney Coordinated Adaptive Traffic System) [7] and SCOOT (Split Cycle Offset Optimizing Technique) [8]. Then, it can be indicated that the descriptions of traffic flow characteristics are the most critical factor in traffic signal timing optimization. However, the traffic flow becomes unstable when the traffic demand approaches or exceeds the capacity [9]. The detrimental effects, such as spillback, residual queue, or De-facto red, make it hard to describe traffic flow characteristics accurately. Thus, traffic signal optimization methods established by traffic flow formulas under normal traffic condition are no longer suitable for oversaturated conditions. The data-driven based heuristic algorithms could be an ideal method to obtain optimized traffic signal timing plan under oversaturated conditions. The heuristic algorithms do not rely on the traffic flow formulas but to seek for optimization scenarios based on real time traffic data. For the reason that various factors can affect the effects of traffic signal control under oversaturated condition, it will be better to consider more impact factors in the process of traffic signal optimization.

The focus of this paper is to present a multi-objective optimization method to obtain a relatively better signal

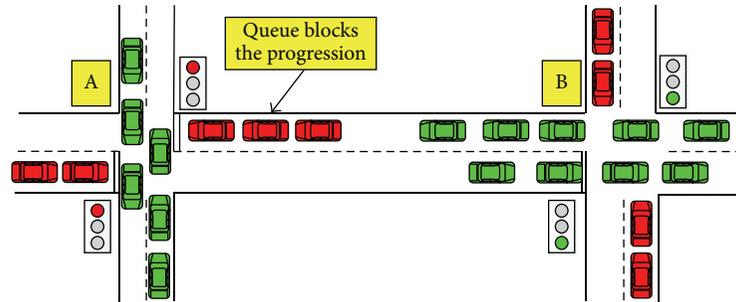


FIGURE 1: Residual queues at intersection.

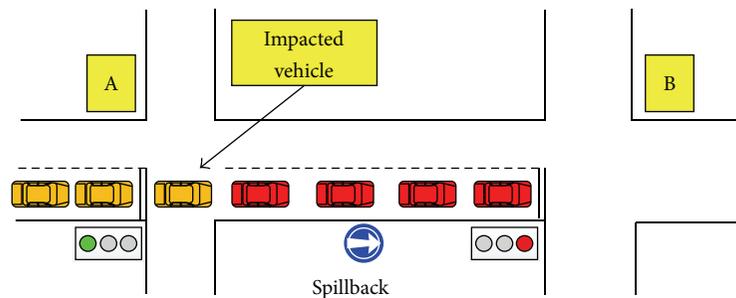


FIGURE 2: Spillback at intersection.

timing plan for oversaturated intersection. Before we design the algorithm, the characteristics of oversaturated traffic flow are analyzed to acquire the optimized objectives of the algorithm. Then, we present the details of the algorithm, which include coding scheme, optimization objectives, and algorithm selection. At last the convergence and simulation results of the algorithm under different conditions are summarized and analyzed.

2. Characteristics of Oversaturated Flow

The traffic status can be determined by traffic intensity, that is, the V/C ratio. At signalized intersection, the over-saturation can be defined as the condition of an approach with residual queue [10], which is illustrated as Figure 1. As it is difficult to measure the residual queue directly, it can be estimated by loop detector [11] or mobile sensors [12] information using shockwave theory.

The traffic flow characteristics will be different when traffic flow is under over-saturation conditions, or approaching saturation conditions [13]. When the traffic condition is approaching saturation, the traffic flow will become unstable. A small fluctuation from any vehicle in a platoon may cause adverse consequences and reduce the efficiency of traffic system sharply [14]. When the traffic status is under oversaturated condition, the deterministic queuing model can be adopted to estimate the queue length. For the reason of low stability of saturated traffic flow, the parameters estimated by the models listed above can hardly be utilized in the process of signal timing.

Generally, oversaturated condition will firstly appear at one or several isolated intersection with relatively high

saturation degree. Then, the congestion begins to spread to adjacent intersections with detrimental effect like “spillback,” which is shown in Figure 2. Intersections along a route with larger traffic volume may spread the congestion much faster than other routes [15]. Finally, the whole road network will be congested, even in a “lock-out” status [16]. According to process of traffic congestion generation, the traffic control strategies under oversaturated condition should be utilized at isolated intersections to avoid detrimental effects like spillback or residual queue.

3. Algorithm Selection and Design

3.1. Selection of the Algorithm. Throughput maximum and queue ratio maintenance are two conflicting objectives, which is a typical multi-objective optimization problem (MOP) [17]. Many intelligent algorithms, such as Evolutionary Algorithm (EA) [18], Particle Swarm Optimization (PSO) [19], Simulated Annealing Algorithm (SA) [20], and Ant Colony Optimization (ACO) [21], can be utilized to obtain the Pareto front of the MOP. However, of all the algorithms, only the Genetic Algorithms (GA), one category of the Evolutionary Algorithm, was successfully applied in the traffic signal control system for commercial productions [22]. A Non-dominated Sorting Genetic Algorithm II (NSGA-II) [23] has better performance than other multi-objective evolutionary algorithms [24], and thus it is selected as the traffic signal control optimization method in this paper.

The NSGA-II algorithm improved the NSGA by adding elitist strategy, density estimation strategy, and fast non-dominated sorting strategy. The time complexity of NSGA-II drops to $O(rN^2)$ from $O(rN^3)$ of NSGA, which is not

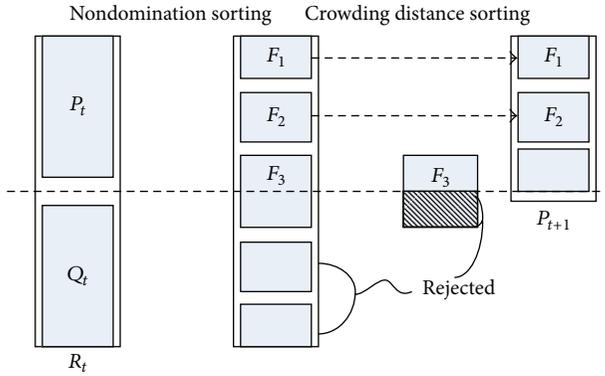


FIGURE 3: Major procedure of NSGA-II algorithm.

larger than other MOEA. The non-dominated sorting and the elitist strategy can improve the performance and keep the better solution. The density estimation strategy avoids the share parameter, which is hard to determine. These strategies make the NSGA-II algorithm have better performance than other MOEA. However, the convergence and diversity of the NSGA-II may be worse for the reason of density estimation strategy.

3.2. NSGA-II Algorithm. The major procedure of NSGA-II algorithm is shown in Figure 3. The main loop of the NSGA-II algorithm is listed as the following steps.

Step 1. Initially, a random parent population P_0 is created. The population is sorted based on the nondomination. Each solution is assigned a fitness (or rank) equal to its nondomination level. The usual binary tournament selection, recombination, and mutation operators are used to create an offspring population Q_0 of size N . Let $t = 0$.

Step 2. A combined population $R_t = P_t \cup Q_t$ is formed. The population R_t is of size $2N$. Then, the population R_t is sorted according to nondomination. The best non-dominated set F_i is formed.

Step 3. The crowded-comparison operator \prec_n is chosen to sort the non-dominated set F_i in descending order. The best N members of the set are chosen for the new population P_{t+1} .

Step 4. The new population P_{t+1} is now used for selection, crossover and mutation to create a new population Q_{t+1} .

Step 5. When the termination condition meets, the loop stops; otherwise, $t = t + 1$, and turn to Step 2.

3.3. Coding Scheme. The coding scheme has evidently effects on genetic manipulation, especially for crossover. The binary encoding, real number encoding and structural encoding are three mostly used coding scheme in genetic algorithms. In the traffic signal control optimization problem, the coding scheme should get the capability to describe the signal timing plan. If the binary encoding is applied, additional constraints

are needed in the process of crossover and mutation. Besides, the logical constraints between genes should also be considered (like the cycle length should be the summation of all the green times and lost times). All these constraints increase the complexity of the calculation. Hence, the real number encoding was selected as the coding scheme in traffic signal control optimization problem. The green time of each phase is the executable genetic operator of gene.

3.4. Selection of Optimization Objectives. Most practitioners tend to make simple changes to splits or phase sequence to minimize delay before moving to more complex approaches. Considering the traffic control strategies for oversaturated conditions are immature, it is better to keep the current control strategy instead of applying a new one. It is also important to implement strategies (where possible) to prevent over-saturation from occurring in the first place, rather than reacting to the issues after the fact. Once the traffic demand goes back to the normal level, the frequently used traffic control strategy will make the practitioners manage the traffic flow easily. In this way, the throughput maximum should be selected as one of the optimization objectives.

The queue problem is the major detrimental effects under oversaturated condition. Signal timing optimization algorithm should get the capability of maintaining the queue ratio of each approach in an acceptable range. In this control strategy, the green times should be adjusted to balance the queues at intersection's approaches during oversaturated conditions. The objective is to minimize the number of blocked intersections by critical intersection's queues.

Based on the discussion listed above, two optimization objectives, "throughput maximum" and "queue ratio maintenance," are selected in the proposed traffic signal timing optimization algorithm.

3.5. Optimization Objective Functions. In order to achieve the optimization objective of throughput maximum, the algorithm tends to extend cycle length to avoid the lost time brought by phase transition process. However, the queue of the conflicted intersection approaches will form very quickly under oversaturated condition and get extremely high probability to be "spillback" if the long cycle length is selected. At this time, the proposed algorithm should have the capabilities of maintaining the queue ratio at each approach to prevent the detrimental effects. It is assumed that high resolution traffic data are available for signal timing plan optimization. The output of the algorithm will be the green times of all the phases. The phase sequence is predefined.

Under oversaturated condition, traffic system is erratic, especially at the critical route of subnetwork or critical movement of the intersection. Hence, the primary target of traffic control under oversaturated condition is to maintain the traffic flow in a stable state and to discharge as many vehicles as possible, especially for the critical movement. For the purpose of maximizing the throughput vehicles of the critical movement, a weight coefficient should be multiplied by the actual throughput numbers in the corresponding optimization objective function. Thus, the throughput maximum

optimization function in the algorithm can be obtained as follows:

$$O_c = \max \sum_{i=1} (\omega_i \cdot n_i(g_i, C, t)), \quad (1)$$

where O_c is the weighted maximum throughput vehicle number of the intersection, ω_i is the weight coefficient of the i th movement, and $n_i(g_i, C, t)$ is the actual throughput vehicles of the i th movement, which is a function of the green time of movement g_i , cycle length C , and time; this parameter can be directly obtained from the VISSIM simulation environment in this research.

Many traffic optimization measures can be classified as the queue maintenance strategy, such as average queue length, maximum queue length, average maximum queue length and dissipated queue length. The queue occupancy ratio is utilized as the optimization objective by considering the impact of acceptable queue space. By introducing this conception in the optimization objective function, the detrimental effect “spillback” can be eased effectively. Meantime, the queue length of each phase can also be optimized. The queue ratio maintenance optimization function in the algorithm can be obtained as follows:

$$O_l = \min \frac{1}{p} \sum_{p=1}^P \frac{\text{len}_p}{L_p}, \quad (2)$$

where p is the number of the phases, $\text{len}_p(t)$ is the queue length of the p th phase, and L_p is the maximum acceptable queue of the corresponding phase.

4. Evaluation Setup

4.1. Simulation Environment. The proposed algorithm was deployed and evaluated in a prevailing microscopic traffic simulation environment, VISSIM. The advantages of VISSIM over other simulation packages include the following

- (1) VISSIM provides the largest flexibility for users to calibrate driving behaviors and traffic conditions.
- (2) VISSIM was developed under .NET framework, which brings flexibility for add-on program development.
- (3) VISSIM provides the best tools for the development of signal control strategies, such as the NEMA controller emulator, Vehicle Actuated Programming (VAP) language, signal control Application Programming Interfaces (SCAPI), and so forth.

VISSIM SCAPI method was used to develop the signal control emulator in this research. SCAPIs were written in C++/CLR language [25] and the original version of SCAPI controller requires signal control algorithms to be embedded into a single dynamic link library (DLL) file. To facilitate the development, a middleware was developed, which can synchronously collect all the real-time detectors/phases states from the VISSIM network to the controller emulator then

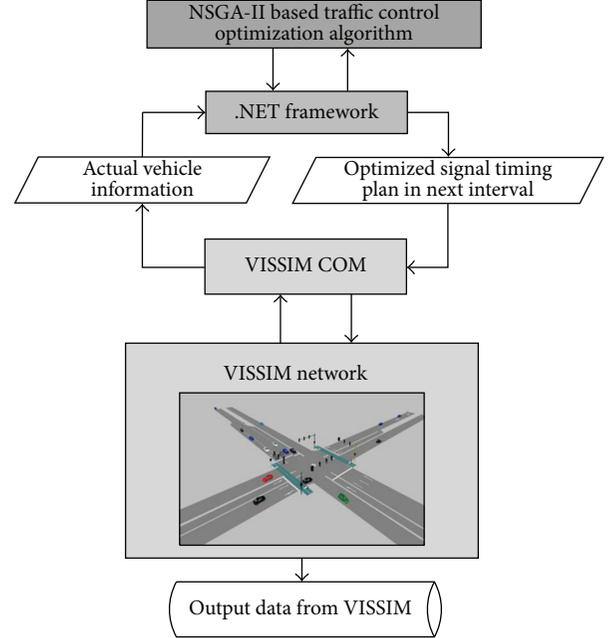


FIGURE 4: Simulation environment of VISSIM.

return the new desired phase states back to the VISSIM network. At each time step, the controller runs the algorithm, makes decisions according to the current state, and then returns the new desired phase states to the VISSIM network. The concept of this simulation environment is illustrated as in Figure 4.

The algorithm was tested under a PC environment with the configuration of Intel Core i7-3770 3.4 GHz processor and 8 GB DDR3 RAM. It takes less than 1 s to optimize the signal timing plan of single intersection on average. The VISSIM software has the capability of simulating the network ten times faster than the real time at minimum. Thus, a 120 s interval, which is larger than most cycle length applied in the real filed, is selected as the update frequency. This parameter should be adjusted based on the hardware configuration. During the period of updating the signal control parameters, the simulation software can provide the simulated traffic flow parameters in the next 20 minutes, which can become the input of the algorithm. The signal timing plan will be updated synchronously in the simulation environment when the current cycle is finished.

4.2. Experimental Design. A simple four way intersection is selected to evaluate the proposed algorithm. The geometric characteristics of the intersection are shown in Figure 5. The experimental configurations are listed in Table 1, and experimental traffic volume is listed in Table 2. Within the signal timing plan, the NEMA's (National Electrical Manufacturers Association) ring structure (see Figure 7) [26] is utilized to achieve the purpose of adjusting the green flexibly based on traffic volume.

TABLE 1: Experimental configurations of the proposed algorithm.

Parameters	Desired value
The length of intersection approach (m)	550
Desired speed (km/h)	60
Number of genes	4 for signal ring 5 for signal ring and phase sequence 8 for dual-rings
Population size	200
Crossover probability	0.8
Mutation probability	0.1
Evolution generation	100

TABLE 2: Experimental traffic volume (pcu).

Movement	Eastbound		Westbound		Northbound		Southbound	
	Low	High	Low	High	Low	High	Low	High
Left turn	65	130	30	60	30	60	40	80
Throughput	620	1240	700	1400	370	740	510	1020
Right turn	35	70	20	60	20	40	50	100

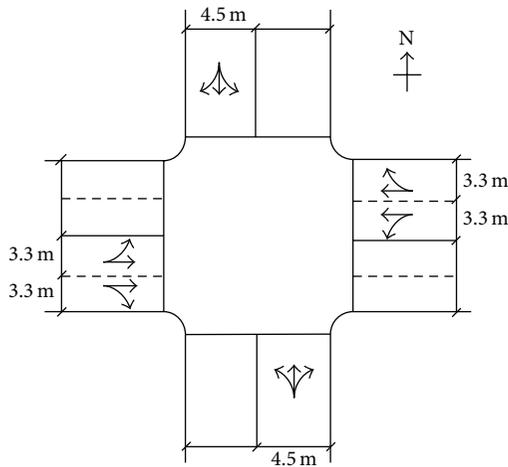


FIGURE 5: Intersection geometry.

5. Results and Discussions

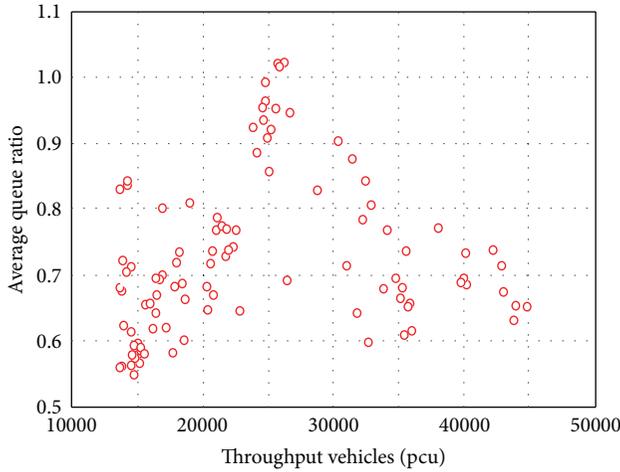
The convergence and results of proposed algorithm at various gene numbers and traffic volumes are analyzed. In order to evaluate the effect of the algorithm, the widely used Webster model and Synchro software are utilized as a reference algorithm to the proposed algorithm.

5.1. Convergence of the Algorithm. In order to test the convergence of the algorithm, the optimization results of the proposed algorithm after 50 generations and 100 generations are compared. As illustrated in Figure 6, the converged speed of the proposed algorithm is relatively fast. After 50 generations, the optimization results have already got a relatively good

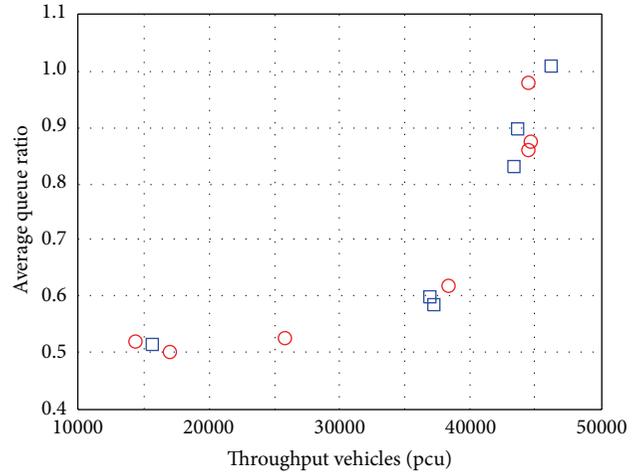
convergence, which is similar to the optimization results after 100 generations. In this way, the evaluation generation can be reduced when it is necessary to consider the calculation time. Although the accuracy of the solution may be lower, the accuracy can still be acceptable based on the experimental results.

5.2. Optimization Results with Different Gene Numbers. As the NEMA's ring structure is utilized in the proposed algorithm, the number of genes can stand for different signal timing scenarios. In this paper, a comparison of scenarios with 4, 5, and 8 genes is selected. The scenario with 4 genes stands for a signal ring timing plan with fixed four phases (left lagging). While the scenario with 5 genes means a signal ring with flexible phase sequence (lead/lagging). And the 8 genes scenario is a standard NAME dual-ring traffic signal timing plan. Figure 7 shows the single ring structure with left lag phase sequence and the standard dual ring structure. Figure 8 illustrates the optimization results of the three scenarios listed above. It can be indicated that the feasible optimal solutions of the 4 and 5 genes are much more than the 8 genes scenario. The optimization results will be better if the distribution of optimal solution can cover the whole period. The distribution of the optimization solutions of 5 genes are more balanced than other scenarios. With the increase of the gene's number, the evaluation generations and population size will rise significantly. Meantime, the complexity of the algorithm will increase too.

5.3. Comparison of the Optimization Results under Different Traffic Volume. As illustrated in Figure 9, the proposed algorithm has the capability of optimizing traffic signal control under different traffic conditions. In the high volume scenario, the traffic volume of southbound exceeds its



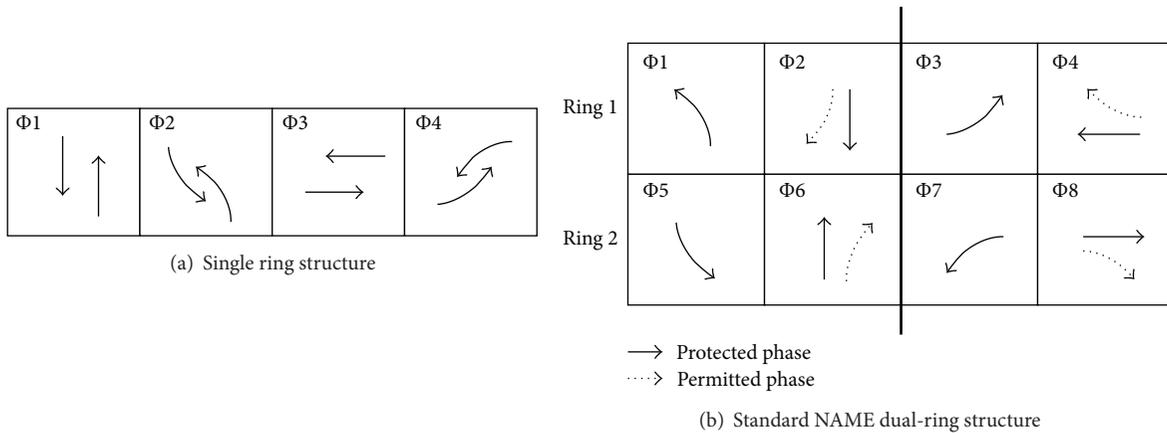
(a) Initial distribution



○ After 50 generations
 □ After 100 generations

(b) Distribution of optimal solutions after 50 and 100 generations

FIGURE 6: Distribution of optimal solutions after various generations.



(a) Single ring structure

(b) Standard NAME dual-ring structure

FIGURE 7: Schematic diagrams of the ring structure.

capacity. The selected intersection is under an oversaturated condition. The optimization results of the Webster method and Synchro are shown in Table 3, and the comparisons of the selected parameters from the methods are shown in Table 4. The Webster model cannot calculate the cycle length with a saturation degree greater than 1; thus the Akcelik’s model is applied to obtain the signal timing plan under oversaturated condition. Based on the optimal results, the average queue ratio of the intersection can be maintained less than 1 by applying the proposed algorithm.

Compared with the optimal results with Webster model and Synchro software, the proposed algorithm has similar performance with other method under normal condition. However, under oversaturated condition, queues build up quickly under the signal timing plan by common used models. For the reason of applying specific optimization objectives, the proposed algorithm has the capability of

searching and converging to Pareto front of the multi-objective problem domain under oversaturated conditions.

It should be noticed that the proposed algorithm has no obvious advantages under the normal conditions. In this way, it is suggested to apply the mature signal timing optimization methods or software, such as Synchro, to determine the signal timing plan under the normal conditions instead of applying the new one. It is also important to implement the proposed algorithm to prevent over-saturation from occurring in the first place, rather than reacting to the issues after the fact. Once the traffic demand goes back to the normal level, the frequently used traffic control method will make the practitioners manage the traffic flow easily.

6. Conclusions

A multi-objective optimization algorithm is present to optimize traffic signal timing at oversaturated intersection. In

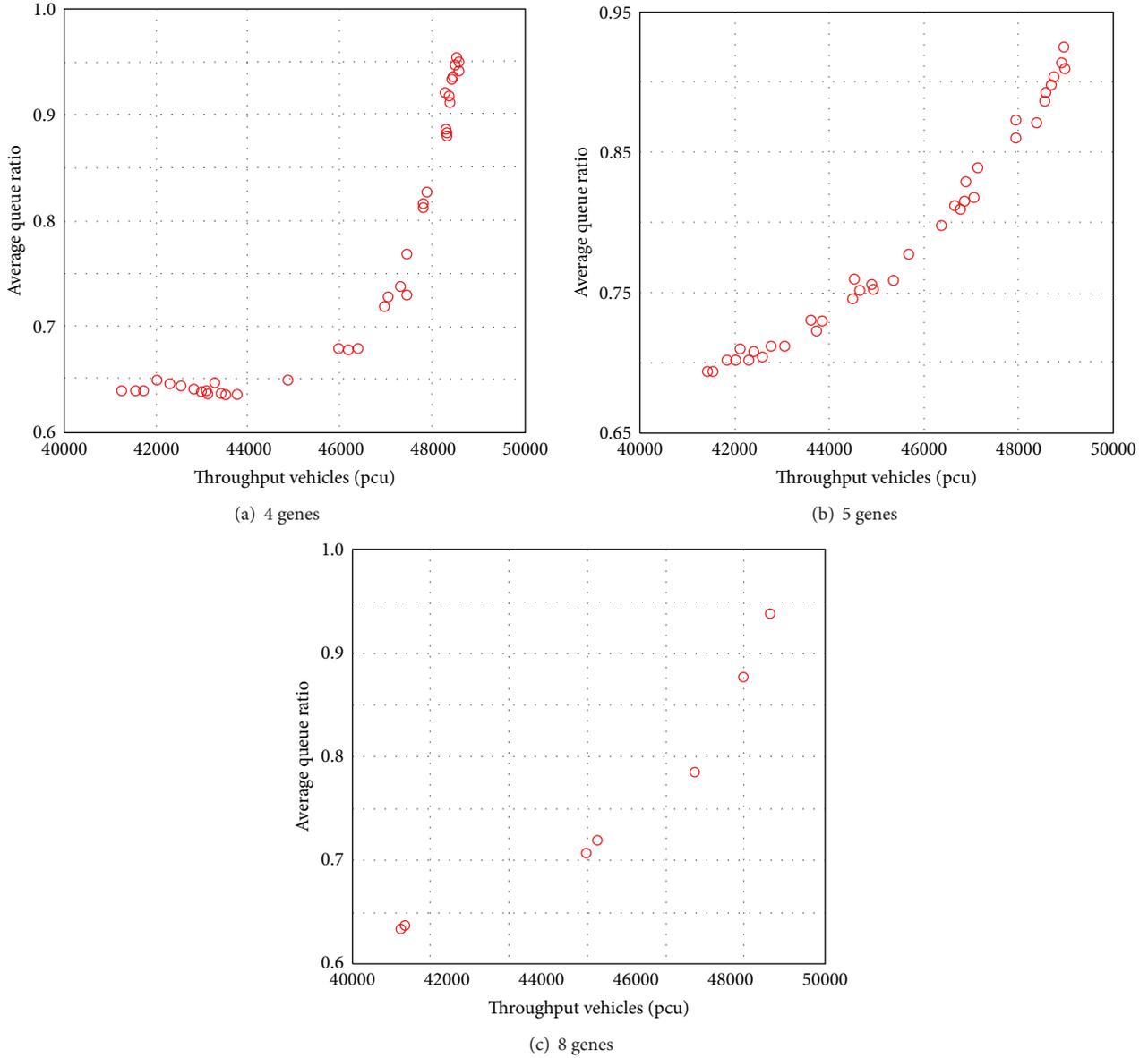


FIGURE 8: Distribution of the optimal solutions with various gene numbers.

TABLE 3: Optimized signal timing plan of Webster model and Synchro software.

Algorithm	Cycle length (s)	Phase 1 (s)	Phase 2 (s)	Yellow (s)
Webster model under normal condition	66	35	23	4
Webster model under over-saturated condition	132	74	50	4
Synchro model under normal condition	70	36	26	4
Synchro model under over-saturated condition	90	45	37	4

TABLE 4: Comparison of the evaluation results between proposed method and other models.

Algorithm	Throughput vehicles (pcu)		Average queue ratio	
	Normal	Over-saturated	Normal	Over-saturated
NSGA II (4 genes)	23417	44492	0.46	0.94
NSGA II (5 genes)	24582	46706	0.44	0.94
NSGA II (8 genes)	24814	48628	0.45	0.90
Webster	23981	29598	0.48	>1
Synchro	24173	35916	0.45	>1

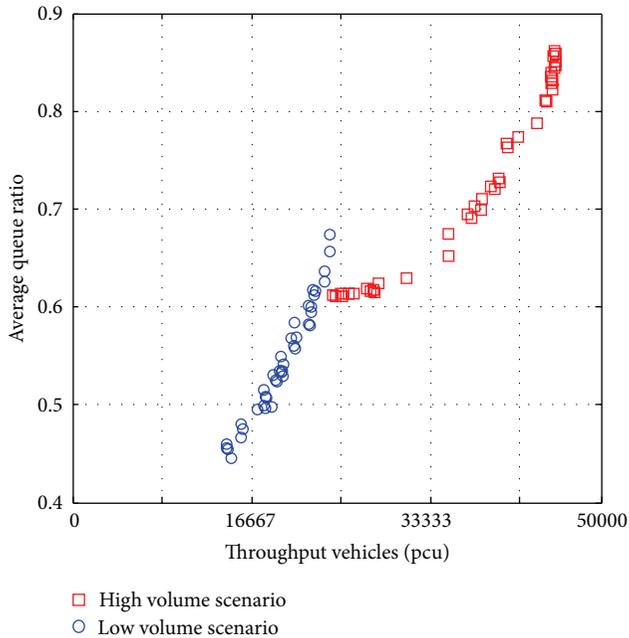


FIGURE 9: Distribution of the optimal solutions under different traffic volumes.

the proposed algorithm, throughput maximum and average queue ratio minimum are selected as the optimization objectives to meet the characteristics of oversaturated traffic flow. The proposed algorithm was tested under various kinds of phases and traffic conditions. The gene number can reflect the phase sequence and traffic control structure of the signal timing. The more genes are selected, the more accurate the algorithm is. Meantime, the proposed algorithm will be difficult to get feasible solution if the large gene number is selected because of the computational complexity. Thus, it is suggested that the gene number should be no more than 8. The real number encoding is utilized in the proposed algorithm, and the executable genetic operator of gene is the green time of the corresponding phase. The update frequency applied in the simulation environment was 120 s, and it can meet the requirements of signal timing plan update in real filed. The experimental results from VISSIM SCAPI simulation environment indicate that the proposed algorithm has the ability of obtaining traffic signal timing plan with better performance than the commonly utilized signal timing optimization software under oversaturated conditions. The performance of proposed algorithm under normal condition is not better than the commonly used traffic signal control optimization methods for the reason of the discrepancies between the optimization objectives and the traffic flow characteristics. The proposed algorithm has the capability of searching Pareto front of the multi-objective problem domain. Further jobs should be concerned on the signal timing optimization method for oversaturated coordinated intersections or small scale road network and real field applications with the traffic signal controller.

Acknowledgments

This research was supported by the National Nature Science Foundation of China (no. 51208054) and the Fundamental Research Funds for the Central Universities (no. 2013G1211005).

References

- [1] M. M. Abbas, Z. M. Adam, and D. Gettman, "Development and evaluation of optimal arterial control strategies for oversaturated conditions," *Transportation Research Record*, no. 2259, pp. 242–252, 2011.
- [2] G. Abu-Lebdeh and R. F. Benekohal, "Design and evaluation of dynamic traffic management strategies for congested conditions," *Transportation Research A*, vol. 37, no. 2, pp. 109–127, 2003.
- [3] G. Abu-Lebdeh and R. F. Benekohal, "Signal coordination and arterial capacity in oversaturated conditions," *Transportation Research Record*, no. 1727, pp. 68–76, 2000.
- [4] Y. Li, X. Guo, J. Yang, Y. Liu, and S. He, "Mechanism analysis and implementation framework for traffic signal control of oversaturated intersection group," *Journal of Transportation Systems Engineering and Information Technology*, vol. 11, no. 4, pp. 28–34, 2011.
- [5] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proceedings of the IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.
- [6] Transportation Research Board, *Highway Capacity Manual*, National Academy of Sciences, Washington, DC, USA, 5th edition, 2010.
- [7] A. G. Sim and K. W. Dobinson, "The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 2, pp. 130–137, 1980.
- [8] R. D. Bretherton and G. T. Bowen, "Recent enhancements to SCOOT-SCOOT version 2.4," in *Proceedings of the 3rd International Conference on Road Traffic Control*, pp. 95–98, London, UK, May 1990.
- [9] F. Dion, H. Rakha, and Y.-S. Kang, "Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections," *Transportation Research B*, vol. 38, no. 2, pp. 99–122, 2004.
- [10] R. W. Denney, L. Head, and K. Spencer, *Signal Timing under Saturated Conditions*, Federal Highway Administration, Washington, DC, USA, 2008.
- [11] X. Wu, H. X. Liu, and D. Gettman, "Identification of oversaturated intersections using high-resolution traffic signal data," *Transportation Research C*, vol. 18, no. 4, pp. 626–638, 2010.
- [12] X. J. Ban, P. Hao, and Z. Sun, "Real time queue length estimation for signalized intersections using travel times from mobile sensors," *Transportation Research C*, vol. 19, no. 6, pp. 1133–1156, 2011.
- [13] B. B. Park and C. Li, "An analytical approach for estimating the highway capacity manual signalized intersection delay variability," *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 4, pp. 327–333, 2011.
- [14] F. Zheng and H. Van Zuylen, "Uncertainty and predictability of urban link travel time: delay distribution-based analysis," *Transportation Research Record*, no. 2192, pp. 136–146, 2010.

- [15] A. Chen, H. Yang, H. K. Lo, and W. H. Tang, "Capacity reliability of a road network: an assessment methodology and numerical results," *Transportation Research B*, vol. 36, no. 3, pp. 225–252, 2002.
- [16] E. J. Smaglik, D. M. Bullock, D. Gettman, C. M. Day, and H. Premachandra, "Comparison of alternative real-time performance measures for measuring signal phase utilization and identifying oversaturation," *Transportation Research Record*, no. 2259, pp. 123–131, 2011.
- [17] K. Ghoseiri, F. Szidarovszky, and M. J. Asgharpour, "A multi-objective train scheduling model and solution," *Transportation Research B*, vol. 38, no. 10, pp. 927–952, 2004.
- [18] J.-H. Kim, J.-H. Han, Y.-H. Kim, S.-H. Choi, and E.-S. Kim, "Preference-based solution selection algorithm for evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 20–34, 2012.
- [19] J. Jie, J. Zeng, C. Han, and Q. Wang, "Knowledge-based cooperative particle swarm optimization," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 861–873, 2008.
- [20] W.-C. Hong, "Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm," *Neurocomputing*, vol. 74, no. 12–13, pp. 2096–2107, 2011.
- [21] R. Putha, L. Quadrifoglio, and E. Zechman, "Comparing ant colony optimization and genetic algorithm approaches for solving traffic signal coordination under oversaturation conditions," *Computer-Aided Civil and Infrastructure Engineering*, vol. 27, no. 1, pp. 14–28, 2012.
- [22] B. Park, C. Messer, and T. Urbanik II, "Traffic signal optimization program for oversaturated conditions: genetic algorithm approach," *Transportation Research Record*, no. 1683, pp. 133–142, 1999.
- [23] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [24] D. Sun, R. F. Benekohal, and S. T. Waller, "Multiobjective traffic signal timing optimization using non-dominated sorting genetic algorithm," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 198–203, Columbus, Ohio, USA, 2003.
- [25] D. Lacks and T. Kocak, "Developing reusable simulation core code for networking: the grid resource discovery example," *Journal of Systems and Software*, vol. 82, no. 1, pp. 89–100, 2009.
- [26] National Electrical Manufacturers Association, *National Transportation Communication for ITS Protocol Object Definitions for Actuated Traffic Controller Units*, NEMA, Washington, DC, USA, 1996, no. TS3.5, 1996.

Research Article

Sustainable Transport Data Collection and Application: China Urban Transport Database

Tian Jiang, Zhongyi Wu, Yu Song, Xianglong Liu, Haode Liu, and Haozhi Zhang

China Urban Sustainable Transport Research Center, China Academy of Transportation Sciences, Ministry of Transport, Beijing 100029, China

Correspondence should be addressed to Tian Jiang; peter_zealot@hotmail.com

Received 14 July 2013; Revised 13 October 2013; Accepted 21 November 2013

Academic Editor: Geert Wets

Copyright © 2013 Tian Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Transport policy making process of national and local governments should be supported by a comprehensive database to ensure a sustainable and healthy development of urban transport. China Urban Transport Database (CUTD) has been built to play such a role. This paper is to make an introduction of CUTD framework including user management, data warehouse, and application modules. Considering the urban transport development features of Chinese cities, sustainable urban transport development indicators are proposed to evaluate the public transport service level in Chinese cities. International urban transport knowledge base is developed as well. CUTD has been applied in urban transport data processing, urban transport management, and urban transport performance evaluation in national and local transport research agencies, operators, and governments in China, and it will be applied to a broader range of fields.

1. Introduction

With the economy growing rapidly in recent years, the total number of vehicles increases dramatically and leads to urban traffic congestion, environmental pollution, and energy issues. To enhance the level of service for travelling in urban areas, both central and local governments in China have applied suitable strategies and policies to reduce the traffic congestion and air pollution. Public transport has been considered as the most promising way to solve these problems and is given the priority in urban transport development. The State Council of China, at the beginning of 2013, issued *The Guideline on Further Implementing the Prioritized Development of Urban Public Transport*, which ensures the leading role of public transit in urban transport.

There are several successful cases of database to provide urban transport data to support policy making, transport planning, and evaluation. Several institutions around the world are working on data collection and sharing. We list online resources of their databases as follows:

(i) <http://www.regiolab-delft.nl/> developed by Region Laboratory Delft, Netherlands,

- (ii) http://safety.fhwa.dot.gov/tools/data_tools/fhwasa09002/ developed by Federal Highway Administration (FHWA), USA,
- (iii) National Household Travel Survey: <http://nhts.ornl.gov/>,
- (iv) <http://www.fhwa.dot.gov/policy/> developed by Federal Highway Administration (FHWA), USA,
- (v) Online resource as of 2010: <http://www.fmcsa.dot.gov/facts-research/facts-figures/analysis-statistics/dashome.htm>,
- (vi) Highway Performance Monitoring System: <http://www.nhtsa.gov/Data/>,
- (vii) <http://www.ntdprogram.gov/ntdprogram/> developed by Federal Motor Carrier Safety Administration (FMCSA), USA,
- (viii) <http://mobility.tamu.edu/> developed by National Highway Traffic Safety Administration (NHTSA), USA,
- (ix) <http://www.nra.ie/> developed by Federal Transit Administration (FTA), USA,
- (x) <http://www.astra.admin.ch/>,

- (xi) National Transit Database: <http://www.uitp.org/publications/>.

Problems exist in the reuse of traffic data, with various data formats, different aggregations, and different densities of meta-information (when existing). Miska et al. (2007) developed an International Traffic Database (ITDb) to deal with these problems by applying data name matching or translation to form a comprehensive standardized data pool, which can improve efficiency of the database [1, 2].

ITDb is a database that focuses on collecting and providing urban road traffic data (vehicle speed, traffic volume occupancy, etc.) in cities all over the world to provide help to academic researches or other applications [3, 4]. Another US database, the National Transit Database (NTD), which focuses on data collection and application of public transport data, is different from ITDb [5]. The NTD is developed to satisfy the data requirements of all levels of governments and the public. The NTD can evaluate the performance of nation's public transport system and can be used to calculate the amount of public transport system supporting funds. Public transport service quality evaluation cannot be made since daily service data is not collected.

China Transport Statistical Information Network and China Public Transport Network are two nationwide transport databases developed by Chinese companies and transport management agencies. However, the two databases have many inadequacies which cannot meet the needs for urban transport study or other practical application. The following are the current shortcomings of Chinese transport databases.

- (i) No comprehensive urban transport database (with data types including urban and suburban, domestic and international, public and private transport, GIS and urban transport, etc.).
- (ii) No data uploading and downloading mechanisms.
- (iii) No comparative analysis of urban transport data.
- (iv) No integrated decision making system.

Since the existing transport databases are either not suitable or comprehensive enough to monitor the development of public transit, evaluate policy effects on urban transport improvements, and support decision making, an integrated database needs to be established specially for sustainable urban transport development [6].

China Urban Transport Database (CUTD) is the first nationwide comprehensive urban transport database, with various types of data such as urban transport network, infrastructure, individual travel, public transport, and pedestrian data. The data sources of CUTD include GIS urban network map, traffic control feedback, traffic monitoring system, public transport operation, and so forth. With abundant and wide-covering urban transport data, CUTD supports decision-making, planning, management, and operation process of urban transport, helping the development of safe, convenient, efficient, economical, equitable, and sustainable urban transport systems for Chinese cities. As a project funded by Volvo Research and Educational Foundations (VREF), CUTD is accessible to partners and colleagues from all VREF Centers of Excellence (CoE).

2. The Framework of CUTD

The framework of CUTD includes modules of user management, data warehouse, and application [7, 8]. Indicators based on Chinese urban transport features are proposed to support transport development policy making and the assessment of development strategies' effects.

The three modules of CUTD: user management, data warehouse, and application (see Figure 1) are designed to realize the database's functions of storing transport data (ranging from raw data to statistical analysis results), running statistical analysis, and making evaluations. Each of the modules is designed and developed independently to focus on both technical aspects and the satisfaction of user demands [9]. The development of CUTD is based on its final operation. To meet the requirement of high performance, CUTD data has to be highly transfer-efficient and be of high quality. Other existing designs of databases, which are simply based on a single set of data types, are different from CUTD.

2.1. User Management. CUTD users include data providers and decision makers. Highway companies, bus and metro operators, and statistic departments of governments are data providers who are required to register the database and upload data regularly. All levels of governments and their supporting research agencies are decision makers who can make use of the data analysis and the evaluation and simulation results to develop transport development strategies. During the data uploading and user processes, the database administrator implements data quality management to ensure good data quality for the proper work of algorithms and tools. All datasets are strictly checked according to the quality management rules, which require a minimum set of meta-information.

Other potential users of CUTD data include domestic and international academic organizations (NGOs, universities, specialists, etc.) and individuals. Different levels right to the database are delivered to different groups of users. The Ministry of Transport of the People's Republic of China establishes a specific network for national and international data exchange and cooperation.

2.2. Data Warehouse. The data warehouse of CUTD includes three layers: geographic layer, transport layer, and data collection layer, which composite a stable and compact structure, improving the robustness of CUTD.

2.2.1. Geographic Layer. Based on GIS maps, the geographic layer provides map information that includes aspects of urban transport network and railway network, as shown by Figure 2. Transport infrastructures are divided into two groups, the link and the node, which can ensure a minimum set of meta-information for the users to search and get access to.

2.2.2. Transport Layer. Unlike the static urban transport structure information stored in the geographic layer, the transport layer of data warehouse consists of comprehensive data of all moving objects in urban transport system,

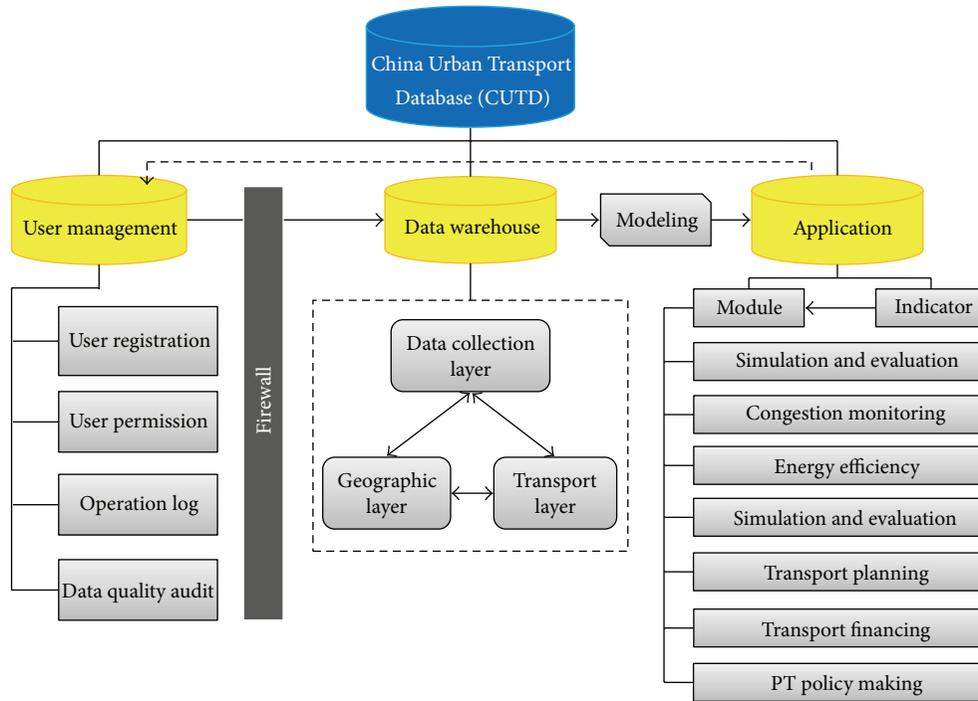


FIGURE 1: The framework of China Urban Transport Database.

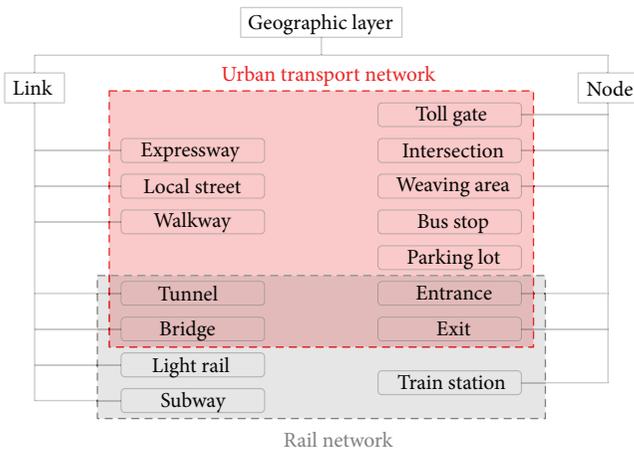


FIGURE 2: Structure of the geographic layer of CUTD.

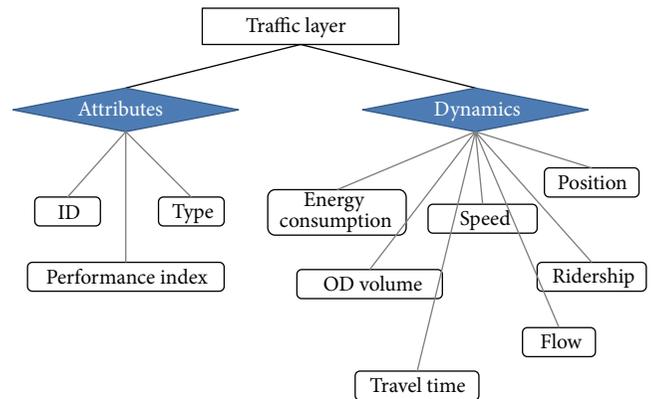


FIGURE 3: Structure of the traffic layer of CUTD.

including cars, buses, trains, motorcycles, bicycles, pedestrians, passengers, and so forth. Information including the original attributes and dynamics describing the movement of objectives are all stored in this layer, as shown by Figure 3.

2.2.3. *Data Collection Layer.* The data-collection layer consists of real-time traffic data collected by road traffic detection systems, which include the sensor type and location, the penetration rate of probe vehicle, and mobile device information. The collection of this type of data needs the support of roadside detecting equipment, GSM equipment, or vehicle-infrastructure communication equipment. Real-time traffic

data can be helpful to online traffic management, intelligent dispatching of public transport modes or taxis.

2.3. *Application.* The following (including but not limited) are function modules which can be realized based on the data warehouse of CUTD. These modules can support the fulfillment of the urban transport development objectives of CUTD.

- (i) Urban traffic simulation and evaluation [10, 11]. This module can provide simulation of urban traffic conditions (road, urban rail, or pedestrian, etc.) and evaluate traffic conditions.

- (ii) Traffic congestion monitoring and management. This module can dynamically monitor urban road traffic condition and identify congestion links or nodes.
- (iii) Energy consumption and emission monitoring and calculation.
- (iv) Urban transport planning.
- (v) Urban transport investment cost-benefit analysis and public transport enterprises' operation cost-audit and subsidy apportionment.
- (vi) Public transport policy making support.
- (vii) Travel service information.

3. China Sustainable Urban Transport Evaluation Indicator System

3.1. Sustainable Urban Transport Evaluation Indicators. For a certain city, CUTD data can be applied in the assessment of urban transport sustainable level and the evaluation of its public transport performance; therefore, the database plays a significant supporting role in the urban transport policy making process. The application of CUTD data can provide help to all levels of governments to improve their urban transport development strategies and urban transport management approaches.

Sustainable Urban Transportation System (SUTS) is a crucial part of an energy-saving, environment-friendly, and people-oriented society. To make an urban transport development towards SUTS, the evaluation of the sustainability of transport system should be carried out first. Establishing an indicator system for the SUTS is a suitable work to start with [12, 13]. The evaluation indicator system is preliminarily composed of 6 aspects and 26 indicators as shown in Table 1.

3.2. Public Transport Performance Indicators. Giving priority to urban public transport development has been regarded as a right strategic choice for urban transport development in Chinese cities. The Chinese government has made a plan to build a number of "Transit Metropolis" in the *Twelfth Five-Year Transport Development Plan*; therefore, the importance of the establishment of a national public transport performance indicator system has appeared. The indicator system can help evaluate and guide the public transport development in Chinese cities.

The following 8 aspects are selected to set indicators for the development of public transport performance indicator systems [14].

- (1) Public transport infrastructure performance (e.g., the ratio of bus parking spaces in the station).
- (2) Public transport service quality [15] (e.g., the availability indicators, the convenience indicators, the cost indicators, the comfortable indicators, and the safety and security indicators).
- (3) Public transport industry economy level (e.g., investment cost-benefit analysis).

TABLE 1: Sustainable urban transport evaluation indicators.

Aspect	Indicator
Transportation function	Road network density (km/km ²)
	Resident average travel time (min)
	Resident average transfer time (min)
	Commute time by public transport (min)
	Average speed at arterial roads (km/h)
	300 m-radius coverage ratio of public transit station (%)
	Public transit network density (km/km ²)
Economic and financial	Berths supply—demand ratio (%)
	Infrastructure invest/GDP (%)
	Public transit invest sharing (%)
	Financial subsidies sharing (%)
	Average annual growth rate of urban transport investment (%)
	Public transport affordability (%)
	Household travel cost (%)
Equity and safety	Pedestrian path area per capita (m ² /person)
	Public transport and nonmotor share (%)
	Percentage of villages with PT services (%)
	Accidental fatalities among 10,000 vehicles (person/10,000 vehicles)
	Annual growth rate of major accidents (%)
Energy consumption	Accidental economic loss (Yuan/vehicle)
	Fuel consumption per vehicle (liter/vehicle)
	Occupier of land (%)
Environmental influences	Traveler volume/road area (person/m ²)
	Pollutant emission per vehicle (gram/vehicle)
Management capacity	Traffic pollution sharing (%)
	Urban transportation management ability

- (4) Public transport operational level (e.g., transportation efficiency, vehicle-employee ratio, and public transport vehicle number per one million people).
- (5) Public transport information service level (e.g., bus intelligent dispatching, electronic bus stop ratio).
- (6) Safety and security level (e.g., fatalities per 10,000 vehicles, annual growth rate of major accidents).
- (7) Energy saving and emission reduction level (e.g., percentage of bus using clean-energy, fuel consumption per passenger/vehicle).
- (8) Public transport industry standardization level (e.g., public transport operation/service standardization, public transport terminology, and glossary standardization).

4. Data Collection Methods

CUTD data can be obtained through the following three channels.

- (1) Ministry of Transport (MOT) and Ministry of Finance (MOF) of China issued circular “Urban and Rural Passenger Transport Fuel Subsidy” in the end of 2009. The central government will establish the “Public Transport Development Foundation” in the near future. Referring to the experience of the US National Transit Database (NTD), all fund receivers or all fund applicants might be required to report energy consumption data and public transport development related data (such as public transport investment data, operation data, and maintenance data, etc.) to the nationwide urban transport database.
- (2) MOT carries out urban passenger transport statistics every year; therefore, the second data source can be developed by means of urban passenger transport statistics mechanism.
- (3) As a tentative plan, cooperation with Chinese cities such as Beijing, Shanghai, Chengdu, Zhengzhou, Jinan, Xi’an, and so forth can provide opportunities to obtain dynamic traffic data and develop the preliminary dynamic data report system. Also, the number of cooperation cities can be increased gradually in the future with the dynamic traffic data report system extended.

5. CUTD Application Examples

The application of CUTD has been pushed forward with the building of Urban Transport Planning, Control and Evaluation Lab, within which five systems including Urban Transport Data Collection and Analysis System, Urban Public Transport Intelligent Control and Information Management System, Urban Transport Data Management System, Urban Transport Planning Decision-Making Supporting System, Urban Transport Simulation and Evaluation System, and Urban Public Transport Monitoring Information Platform have been designed.

5.1. Urban Transport Data Collection and Analysis System. This system has two main functions of data collection and data analysis, which serve both data providers and decision makers. As shown by Figure 4, urban transport data is collected from data providers through extraction, cleansing, classification, and loading and imported into CUTD and IUTD (International Urban Transport Database). By analysis and comparison of the data from the two databases, query, statistical analysis, data mining, and information sharing can be realized. Various types of analysis results can then be shown to decision makers (or other data users).

For data providers, this system offers functions of on-page data reporting, spreadsheet uploading, and data approving to improve data input convenience and ensure high data quality. Providers can log into the system and type in data directly on pages as shown by Figure 5 or fill data into standard formatted excel spreadsheets and upload the files onto the system. Uploaded data can be reviewed, modified, or deleted. Data approving function examines the data, stores qualified

data for the use by analysis system and monitoring system, and returns unqualified data for modification.

5.2. Urban Public Transport Intelligent Control and Information Management System. This system includes two parts: taxi dispatching and information collection and intelligent bus dispatching and control.

The taxi dispatching and information collection module has functions of monitoring, data converging, and reporting generating. Taxi status monitoring, vehicle positioning, vehicle alarming, and vehicle control can be realized by collecting and managing taxi GPS information, driver information, and vehicle information through this module. These types of information can be integrated into a GIS program, which can display vehicle position and driving route on maps. Information of time, speed, vehicle, driver, and even vehicle monitoring video can also be transmitted and displayed on different interfaces. Taxi dispatching can be compiled by applying various telecommunication measures considering taxi information and the actual operational needs. Figure 6 shows an application of this module in Beijing’s taxi dispatching.

The intelligent bus dispatching and control module collects CAN bus information and GPS information and displays the position, speed, schedule following status, and other information of buses on real-time interfaces. This module has been applied in Dalian Development Area. There are three view types of bus operation status: simulation view as shown by Figure 7, track view, and schedule view as shown by Figure 8. Dispatching and managing tasks including arrival, departure, operation, stop, return, and fault can be carried out by reviewing the real-time information and issuing instructions through communication tools integrated with other functions.

At the same time, all types of bus operation data are stored into historical database as spreadsheets or operation log. The data can be reviewed by database users or analyzed by statistical programs.

5.3. Urban Transport Data Management System. Within the information construction process, various operation systems have been developed by different departments according to their own business demands. All those systems are separated from each other without any connection or communication. The Urban Transport Data Management System is built to unify all those databases. The following achievements are reached.

5.3.1. Data Management Specifications. Data storing specification ensures the reliability and integrity of data transfer and realizes the pooling and sharing of data.

Data coding specification provides unified conversion interface and unified storage of multisource data.

Data exchange specification provides standard data coding rules and realizes the data exchange between heterogeneous databases.

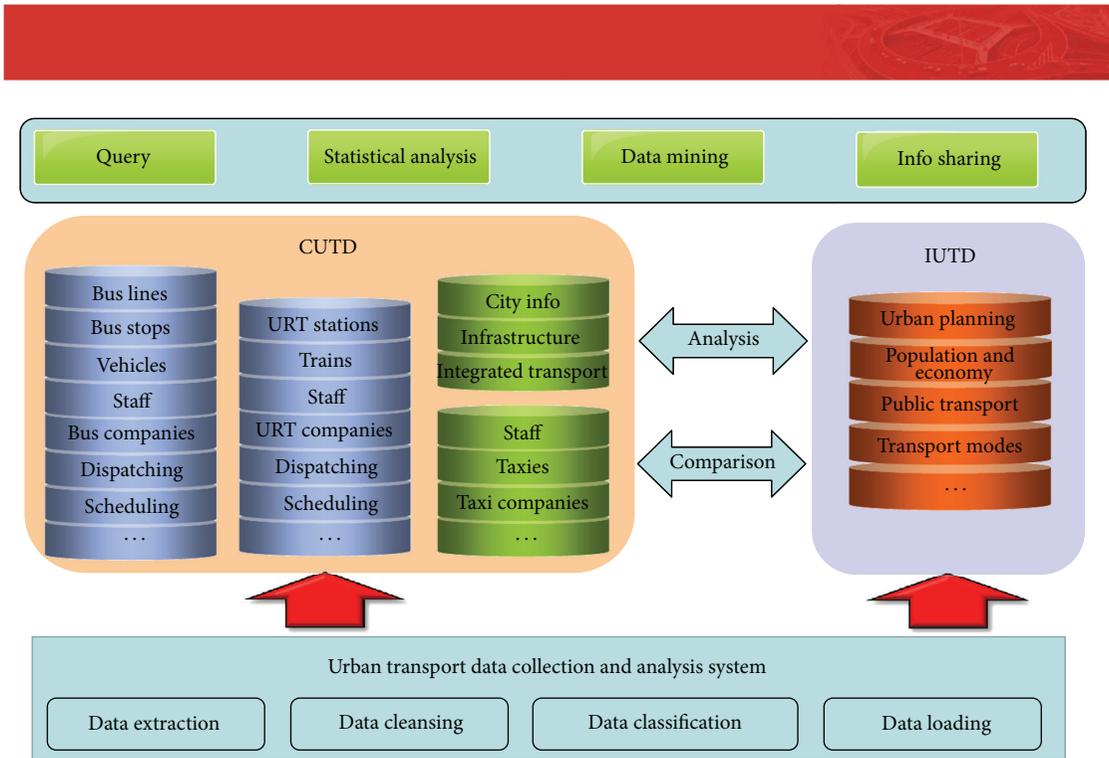


FIGURE 4: The software framework of Urban Transport Data Collection and Analysis System.



FIGURE 5: On-page data reporting.

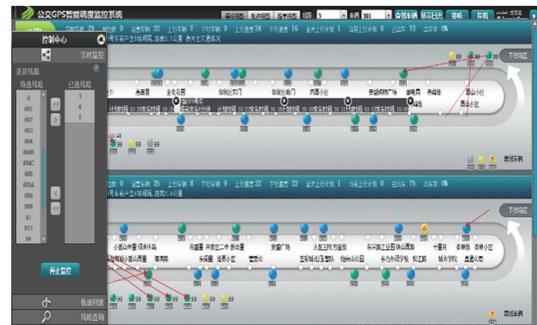


FIGURE 7: Bus operation monitoring and dispatching.



FIGURE 6: Taxi position monitoring based on GIS.

5.3.2. Collection of Domestic and International Urban Transport Data. CUTD has included 58 spreadsheets with over 63.22 million pieces of data.

IUTD has included 51 spreadsheets with over 0.80 million pieces of data.

Also, 0.974 million pieces of abnormal data have been cleaned and filtered.

CUTD data collections include the following.

- (i) Dynamic collection of real-time bus and taxi dispatching data.
- (ii) Exchange of data from domestic statistical information platform.

线路	上车时间	下车时间								
0700	07:00	07:20	07:40	08:22	08:24	08:34	08:54	09:09	09:24	09:34
0710	07:10	07:14	07:52	07:55	08:28	08:30	09:01	09:15	09:53	09:56
0720	07:20	07:24	08:02	08:05	08:38	08:40	09:11	09:25	10:03	10:06
0730	07:30	07:34	08:12	08:15	08:48	08:50	09:21	09:35	10:13	10:16
0740	07:40	07:44	08:22	08:25	08:58	09:00	09:31	09:45	10:23	10:26
0750	07:50	07:54	08:32	08:35	09:08	09:10	09:41	09:55	10:33	10:36
0800	08:00	08:04	08:42	08:45	09:18	09:20	09:51	10:05	10:43	10:46
0810	08:10	08:14	08:52	08:55	09:28	09:30	10:01	10:15	10:53	10:56
0820	08:20	08:24	09:02	09:05	09:38	09:40	10:11	10:25	11:03	11:06
0830	08:30	08:34	09:12	09:15	09:48	09:50	10:21	10:35	11:13	11:16
0840	08:40	08:44	09:22	09:25	09:58	10:00	10:31	10:45	11:23	11:26
0850	08:50	08:54	09:32	09:35	10:08	10:10	10:41	10:55	11:33	11:36
0900	09:00	09:04	09:42	09:45	10:18	10:20	10:51	11:05	11:43	11:46
0910	09:10	09:14	09:52	09:55	10:28	10:30	11:01	11:15	11:53	11:56
0920	09:20	09:24	10:02	10:05	10:38	10:40	11:11	11:25	12:03	12:06
0930	09:30	09:34	10:12	10:15	10:48	10:50	11:21	11:35	12:13	12:16
0940	09:40	09:44	10:22	10:25	10:58	11:00	11:31	11:45	12:23	12:26
0950	09:50	09:54	10:32	10:35	11:08	11:10	11:41	11:55	12:33	12:36
1000	10:00	10:04	10:42	10:45	11:18	11:20	11:51	12:05	12:43	12:46
1010	10:10	10:14	10:52	10:55	11:28	11:30	12:01	12:15	12:53	12:56
1020	10:20	10:24	11:02	11:05	11:38	11:40	12:11	12:25	13:03	13:06
1030	10:30	10:34	11:12	11:15	11:48	11:50	12:21	12:35	13:13	13:16
1040	10:40	10:44	11:22	11:25	11:58	12:00	12:31	12:45	13:23	13:26
1050	10:50	10:54	11:32	11:35	12:08	12:10	12:41	12:55	13:33	13:36
1100	11:00	11:04	11:42	11:45	12:18	12:20	12:51	13:05	13:43	13:46
1110	11:10	11:14	11:52	11:55	12:28	12:30	13:01	13:15	13:53	13:56
1120	11:20	11:24	12:02	12:05	12:38	12:40	13:11	13:25	14:03	14:06
1130	11:30	11:34	12:12	12:15	12:48	12:50	13:21	13:35	14:13	14:16
1140	11:40	11:44	12:22	12:25	12:58	13:00	13:31	13:45	14:23	14:26
1150	11:50	11:54	12:32	12:35	13:08	13:10	13:41	13:55	14:33	14:36
1200	12:00	12:04	12:42	12:45	13:18	13:20	13:51	14:05	14:43	14:46
1210	12:10	12:14	12:52	12:55	13:28	13:30	14:01	14:15	14:53	14:56
1220	12:20	12:24	13:02	13:05	13:38	13:40	14:11	14:25	15:03	15:06
1230	12:30	12:34	13:12	13:15	13:48	13:50	14:21	14:35	15:13	15:16
1240	12:40	12:44	13:22	13:25	13:58	14:00	14:31	14:45	15:23	15:26
1250	12:50	12:54	13:32	13:35	14:08	14:10	14:41	14:55	15:33	15:36
1300	13:00	13:04	13:42	13:45	14:18	14:20	14:51	15:05	15:43	15:46
1310	13:10	13:14	13:52	13:55	14:28	14:30	15:01	15:15	15:53	15:56
1320	13:20	13:24	14:02	14:05	14:38	14:40	15:11	15:25	16:03	16:06
1330	13:30	13:34	14:12	14:15	14:48	14:50	15:21	15:35	16:13	16:16
1340	13:40	13:44	14:22	14:25	14:58	15:00	15:31	15:45	16:23	16:26
1350	13:50	13:54	14:32	14:35	15:08	15:10	15:41	15:55	16:33	16:36
1400	14:00	14:04	14:42	14:45	15:18	15:20	15:51	16:05	16:43	16:46
1410	14:10	14:14	14:52	14:55	15:28	15:30	16:01	16:15	16:53	16:56
1420	14:20	14:24	15:02	15:05	15:38	15:40	16:11	16:25	17:03	17:06
1430	14:30	14:34	15:12	15:15	15:48	15:50	16:21	16:35	17:13	17:16
1440	14:40	14:44	15:22	15:25	15:58	16:00	16:31	16:45	17:23	17:26
1450	14:50	14:54	15:32	15:35	16:08	16:10	16:41	16:55	17:33	17:36
1500	15:00	15:04	15:42	15:45	16:18	16:20	16:51	17:05	17:43	17:46
1510	15:10	15:14	15:52	15:55	16:28	16:30	17:01	17:15	17:53	17:56
1520	15:20	15:24	16:02	16:05	16:38	16:40	17:11	17:25	18:03	18:06
1530	15:30	15:34	16:12	16:15	16:48	16:50	17:21	17:35	18:13	18:16
1540	15:40	15:44	16:22	16:25	16:58	17:00	17:31	17:45	18:23	18:26
1550	15:50	15:54	16:32	16:35	17:08	17:10	17:41	17:55	18:33	18:36

FIGURE 8: Bus schedule following status.



FIGURE 9: The interface of Urban Transport Data Management System.

- (iii) Collection of data from mastered basic researches and original surveys.
- (iv) Collection of relevant data from surveys in typical cities.
- (v) Collection of relevant data from surveys in transport enterprises.
- (vi) Collection of relevant result data from special surveys in cities by transport management departments.
- (vii) Collection of historical data (from urban transport yearbooks or transport yearbooks).
- (viii) Collection of fuel consumption data of urban-rural passenger transport and rural water passenger transport.

Classifications of IUTD data include the following.

- (i) Collection of data from basic researches.
- (ii) Collection of relevant result data from surveys in cities.

5.3.3. *A Unified Data Interface for Planning, Simulation, Monitoring, and Evaluation.* By developing the functions of data sharing and exchanging, a unified data interface for planning, simulation, monitoring, and evaluation is provided [16]. As shown in Figure 9, the Urban Transport Data Management System can realize four main functions including basic data element management, information resource menu management, data exchange management, and authorization management.

5.4. *Urban Transport Planning Decision-Making Supporting System.* This system has four models providing evaluation data for urban development indicators, public transport service quality, and future planning to support transport planning decision-making.

5.4.1. *Urban Public Transport Planning Model.* By analyzing basic travel information (collected from daily travel survey and bus passenger survey) and geospatial information (including bus geographic data and socioeconomic data) from urban transport database, this model can achieve travel plan analysis, bus corridor classification, bus line planning, public transport optimization program design, and public transport program review to support decision-making by public transport managers.



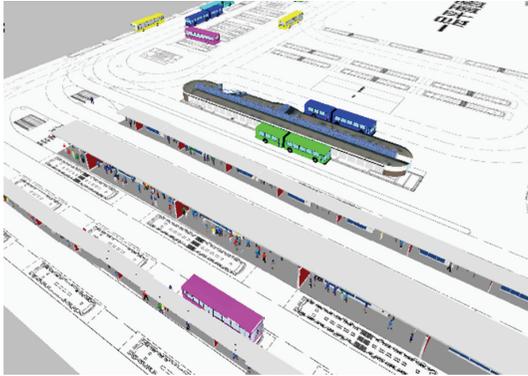


FIGURE 11: A simulation of a public transport hub.

- (iii) Evaluation of transport development impacts on future urban environment.
- (iv) Evaluation of urban transport impacts on health.
- (v) Estimation of vehicle and operation line energy consumption.
- (vi) Emission hot spot analysis.
- (vii) Evaluation of energy saving and emission reduction effects.

5.5. Urban Transport Simulation and Evaluation System. This system includes urban road network simulation and evaluation model and public transport service quality evaluation model.

5.5.1. Urban Road Network Simulation and Evaluation Model. This model is based on microsimulation software, VISSIM, which can realize the simulation of traffic status at different types of road sections and junctions under different control types, different signal timing, different speed limits, or different deceleration controls. The model can get relevant data (infrastructure and transport user data) from urban transport database and apply the data in simulation model construction and model calculation. Outputs of this model are values of evaluation indicators and 3D simulation video which can show the simulated traffic condition vividly and intuitively.

One of the applications of this model is simulation and evaluation of public transport hubs, as shown by Figure 11. Both the move of public transport vehicles and the move of passengers are simulated. By inputting different groups of vehicle and passenger data, different usage statuses can be demonstrated and different indicator values can be obtained to do evaluations. Also, infrastructure data can be modified to meet different operation and travel demands, as a simulation of infrastructure design or transformation. Evaluations and comparisons of different designs can then be carried out.

5.5.2. Public Transport Service Quality Evaluation Model. This model can dynamically correlate index system and bus operation database to make trend analysis of various indicators and comprehensively evaluate urban public transport

选择企业: 郑州市公共交通总公司 查询

选择: 安全性 便捷性 舒适性 可靠性 满意度 节能环保

指标名称	值	评分	排名	指标说明
安全操作合格率(%)	0.9615	88.30	--	被检查安全操作
驾驶员违规率	0.00576	100.00	--	统计期内, 面包车
驾驶员交通违章率	0.002438	92.60	--	统计期内, 面包车
车辆安全设施合格率(%)	0.85	68.30	--	根据国家标准
行车责任事故频率(次/百万公里)	0.4359	100.00	--	统计期内, 行车
安全运行里程(百万公里/次)	2.38	100.00	--	公共汽车企业
行车责任死亡事故频率(次/百万公里)	0.041	64.50	--	统计期内, 行车
重伤水平(人/百万公里)	0.43	81.75	--	统计期内, 公共
死亡水平(人/百万公里)	0.041	64.50	--	统计期内, 公共

FIGURE 12: Operation safety indicators scoring of Zhengzhou Bus Company.

development level by weighting the indexes. Figure 12 shows the operation safety indicators scoring of Zhengzhou Bus Company. Except for the scoring of safety indicators, from the interface it can be seen that there are also scoring of convenience indicators, comfort indicators, reliability indicators, satisfaction indicators, and energy saving (environmental) indicators.

5.6. Urban Public Transport Monitoring Information Platform.

In response to the national public transport development priority strategy and a series of regulatory strategies, as well as guidelines and guidance of the Ministry of Transport, surveys and expert discussions were carried out in Chengdu, Xi'an, Zhengzhou, and Shanghai. The Urban Public Transport Monitoring Information Platform was designed and developed. Monitoring of urban public transport infrastructures, passenger traffic, operation smoothness, and security and emergency responding is realized by collecting and analyzing bus and urban rail transport data. Evaluation of urban public transport development is made according to the Urban Public Transport Development Indicator System. Strong support is provided to promote the national "Public Transport City" demonstration construction projects.

5.6.1. Urban Public Transport Development Evaluation.

By pooling static data and dynamic data, obtaining data with statistical measures, and making comparisons with the data a year earlier, macrojudgment of urban transport situations can be made. This function can assist the evaluation of traffic management program implementation effects and provide data supports to the exploration of the interaction law between different departments.

5.6.2. Bus

Infrastructure Monitoring. By monitoring infrastructure, the development of public transport lines, stations, and vehicles can be mastered and support can be provided to the monitoring and evaluation of urban public transport development, and the strategic development planning decisions. Also, policy support can be provided to further strengthen infrastructure construction.

In terms of vehicles, vehicle type distribution (e.g., new-energy vehicle percentage, and vehicle age distribution) and indicators such as bus unit number per 10 thousand people

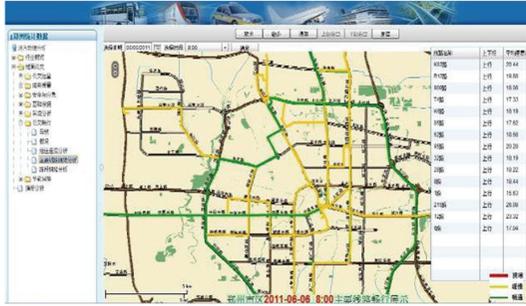


FIGURE 13: Congestion level of a major bus line network in Zhengzhou.

are counted to support the acquisition and maintenance of vehicles by bus companies.

Passenger Traffic Analysis. The passenger traffic analysis module analyzes the spatial and temporal distribution of urban passengers and is used to guide bus line and bus stop planning and the dispatching coordination of bus operators.

- (i) As the most basic indicators, passenger density and distribution are analyzed to provide basic data for hot spot analysis, congestion analysis, and bus corridor building.
- (ii) Main stop load factor reflects the distribution of passengers waiting at bus stops.
- (iii) Key bus line passenger distribution provides data support to transport planning and bus dispatching coordination.

Public Transport Monitoring. Making use of the research achievement of congestion model, the real-time matching of bus GPS data with nearly 80,000 lines' data and nearly 12,000,000 pieces of IC card data, the analysis of key bus line operation can be realized based on GIS.

By monitoring bus passenger flows during different periods in morning peaks and evening peaks, the system can analyze the spatial and temporal distribution of bus passengers between key areas and find out big passenger flows between key areas. The system can also evaluate and describe congestion level of road network and bus line network based on road grade and bus speed, as what was applied in Zhengzhou shown by Figure 13.

Energy Saving and Emission Reduction. Bus energy consumption monitoring system analyzes vehicle energy consumption, CO₂ emission, intensity, fuel consumption structure, and vehicle type structure according to emission standards. It plays an important role in helping understand the energy consumption, operation, and energy saving of transport system. Effective monitoring and statistics of vehicle emission is the basic measure to reduce pollution, save energy, and reduce emission. It can also lay the foundation for future

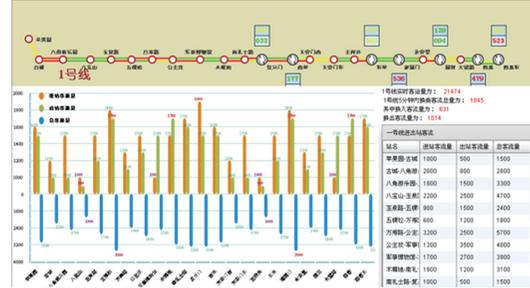


FIGURE 14: Dynamic subway passenger data demonstration.

research on urban transport energy and greenhouse gas emission reduction.

Security and Emergency Responding. The monitoring of security and emergency responding focuses on enterprise-level indicators including the number of over speed alarm, abnormal brake alarm, abnormal door switch alarm, and accidents.

5.6.3. Urban Rail Transport. In line level, by real-time monitoring and comparing passenger entering and exiting stations, passenger flow direction and moving trend can be dynamically demonstrated in forms of vector charts, stacked charts, and detailed statistics. Figure 14 shows the passenger data of Beijing Subway Line 1. The entering, exiting, and total passenger volumes are demonstrated in the form of a histogram. Passenger flow levels are displayed by different colors on the line map.

Based on urban rail passenger sorting model, key indicators (e.g., line load factor) are analyzed, by fine-grained monitoring of passenger traffic in each direction and each station.

6. Conclusion and Future Work

In this paper we have introduced China Urban Transport Database, a user-oriented platform for policy makers. This study has proposed an indicator system of Sustainable Urban Transportation System tailored for the current situation of urban development in China. The results of this study provide the evaluation basis for sustainable urban transportation in China. CUTD will be the primary source of information and statistics on the urban transportation systems of China, serving to improve China's sustainable urban transport development.

As a three-year project, the platform has been established in 2011 based on the framework proposed. Functions of simulation and evaluation, congestion monitoring, energy efficiency, transport planning, and public transport policy making have been implemented by corresponding systems. However, transport financing function is still waiting to be realized; databases still need to be expanded and improved; communications between different systems need to be strengthened. It is also important to improve the applicability of the various indicators, the scientific and operability of the

weight of each indicator, so greater adaptability is needed for the next stage of the study.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

China Urban Transport Database is funded by the Volvo Research and Educational Foundations (VREF).

References

- [1] M. P. Miska, M. Kuwahara, and S. Tanaka, "Online platform for sustainable traffic data storage," in *Proceedings of the 12th World Conference on Transport Research*, Lisbon, Portugal, July 2010.
- [2] M. P. Miska, A. Torday, H. Warita, and M. Kuwahara, "International traffic database—gathering traffic data fast and intuitive," in *Traffic Data Collection and Its Standardization*, J. Barceló and M. Kuwahara, Eds., vol. 144 of *International Series in Operations Research & Management Science*, Springer, Berlin, Germany, 2010.
- [3] M. P. Miska, A. Torday, H. Warita, and M. Kuwahara, "The international traffic database project," in *Proceedings of the ITS World Congress*, Beijing, China, 2007.
- [4] J. Barceló, M. P. Kuwahara, and M. Miska, "Traffic data availability and its standardization," in *Traffic Data Collection and Its Standardization*, J. Barceló and M. Kuwahara, Eds., vol. 144 of *International Series in Operations Research & Management Science*, Springer, Berlin, Germany, 2010.
- [5] U. S. Department of Transportation, *Traffic Detector Handbook*, vol. 2 of *Federal Highway Administration Report No: FHWA-HRT-06-139*, U.S. Department of Transportation, 3rd edition, 2006.
- [6] H. Guo, W. Wang, W. Guo, X. Jiang, and H. Bubb, "Reliability analysis of pedestrian safety crossing in urban traffic environment," *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.
- [7] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [8] J. A. Black, A. Paez, and P. A. Suthanaya, "Sustainable urban transportation: performance indicators and some analytical approaches," *Journal of Urban Planning and Development*, vol. 128, no. 4, pp. 184–209, 2002.
- [9] V. Pina and L. Torres, "Analysis of the efficiency of local government services delivery. An application to urban public transport," *Transportation Research A*, vol. 35, no. 10, pp. 929–944, 2001.
- [10] M. G. Karlaftis and P. S. McCarthy, "Subsidy and public transit performance: a factor analytic approach," *Transportation*, vol. 24, no. 3, pp. 253–270, 1997.
- [11] B. C. Richardson, "Sustainable transport: analysis frameworks," *Journal of Transport Geography*, vol. 13, no. 1, pp. 29–39, 2005.
- [12] H. Gudmundsson, "Making concepts matter: sustainable mobility and indicator systems in transport policy," *International Social Science Journal*, vol. 55, no. 176, pp. 199–217, 2004.
- [13] Y. Tyrinopoulos and C. Antoniou, "Public transit user satisfaction: variability and policy implications," *Transport Policy*, vol. 15, no. 4, pp. 260–272, 2008.
- [14] H. Castillo and D. E. Pitfield, "ELASTIC—a methodological framework for identifying and selecting sustainable transport indicators," *Transportation Research D*, vol. 15, no. 4, pp. 179–188, 2010.
- [15] A. A. Rassafi and M. Vaziri, "Sustainable transport indicators: definition and integration," *International Journal of Environmental Science and Technology*, vol. 2, no. 1, pp. 83–96, 2005.
- [16] C. Kennedy, E. Miller, A. Shalaby, H. MacLean, and J. Coleman, "The four pillars of sustainable urban transportation," *Transport Reviews*, vol. 25, no. 4, pp. 393–414, 2005.
- [17] K. J. Dueker and J. A. Butler, "A geographic information system framework for transportation data sharing," *Transportation Research C*, vol. 8, no. 1–6, pp. 13–36, 2000.

Research Article

Operational Impacts of Using Restricted Passenger Flow Assignment in High-Speed Train Stop Scheduling Problem

Huiling Fu,¹ Benjamin R. Sperry,² and Lei Nie¹

¹ School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

² Ohio University, Athens, OH 45701, USA

Correspondence should be addressed to Huiling Fu; fuhuiling00@163.com

Received 30 July 2013; Revised 1 October 2013; Accepted 29 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Huiling Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One key decision basis to the train stop scheduling process is the passenger flow assignment, that is, the estimated passengers' travel path choices from origins to destinations. Many existing assignment approaches are stochastic in nature, which causes unbalanced problems such as low efficiency in train capacity occupancy or an irrational distribution of transfer passengers among stations. The purpose of this paper is to propose a train stop scheduling approach. It combines a passenger flow assignment procedure that routes passenger travel paths freely within a train network and is particularly capable of incorporating additional restrictions on generating travel paths that better resemble the rail planner's purpose of utilizing capacity resources by introducing four criteria to define the feasibility of travel path used by a traveler. Our approach also aims at ensuring connectivity and rapidity, the two essential characteristics of train service increasingly required by modern high-speed rails. The effectiveness of our approach is tested using the Chinese high-speed rail network as a real-world example. It works well in finding a train stop schedule of good quality whose operational indicators dominate those of an existing stochastic approach. The paper concludes with a comprehensive operational impact analysis, further demonstrating the value of our proposed approach.

1. Introduction

As high-speed rail (HSR) networks are being constructed and expanded worldwide, rail operators face continuing problems of efficiently planning their high-speed train services. One such problem is the train stop scheduling problem (TSSP), which involves specifying a set or subset of stations within a HSR network where individual trains will stop. An efficient train stop schedule is crucial to improve train service connectivity, particularly for passengers whose origins and destinations (ODs) are not situated at limited terminals. On the other hand, the goal of scheduling many stops by a train on its route to ensure good connectivity conflicts with the goal of maintaining the rapidity of that train. From a practical perspective, a rail operator will always seek to maintain the most efficient schedule to provide the optimal balance between stop locations and frequencies.

Understanding passenger behavior requires a fundamental analysis to be conducted when designing services for any transportation modes [1]. For rail, a passenger flow

assignment procedure is a key decision basis to develop a train stop schedule within a train network. Most of the work from the literature that deals with the TSSP concentrated on selectively allocating train stops at stations along a rail line or in a network, where train stop pattern combinations of the so-called non-stop (i.e., direct express), skip-stop, and all-stop are employed. Decisions regarding the number of stops made by a train and at which stations in the network those stops are to take place should be rationally built on a simulation of passengers' travel paths choice behavior. In principle, scheduling train stops (or any other train service plans) should be consistent with the specific passenger demand. Two optimization models with the objectives of covering more passenger demand with fewer train stops as well as saving more travel time for passengers were developed by Hamacher et al. A genetic algorithm was introduced and tested on a partial rail network in southern Germany [2]. In the context of The Netherlands' rail network, three types of stations and train lines are usually referred to as regional (R) or stop trains for type 1, interregional (IR) for type 2,

and Intercity (IC) for type 3. Train lines of type 1 halt at all stations they pass. Lines of type 2 skip the small stations of type 1 and so forth. Goossens et al. established integer models combined with a multicommodity flow problem for multitype line planning problems to minimize operator's operating cost [3]. Using a 46 km long, six-station transit line in the northeastern US as the background, Ulusoy et al. optimized all-stop, short-turn, and express transit services by a cost-efficient operation model, and a logit-based model was used to estimate the ridership of all the seven pre-given train stop patterns [4]. In the setting of Taiwan's HSR line, Chang et al. formulated a multiobjective model with the TSSP embedded in that model to yield a train operation plan. The objectives included minimizing the operator's operating cost and the passengers' travel time loss. The model was solved by a fuzzy mathematical programming approach [5]. A bilevel programming model which was combined with a network equilibrium analysis of passenger flow assignment on trains in a lower-level problem was proposed by Lee and Hsieh. A numerical case study of the final train stop schedule included seven selectable train stop patterns among five stations along the line [6]. For the Beijing-Shanghai Chinese HSR line, Zhang et al. proposed a multiobjective 0-1 programming model to solve the TSSP. The objectives were minimizing train dwell time for travelers, improving the load of every train, and minimizing unsatisfied passenger demand for seven stations along the line [7]. Over a partial rail network mixed of both HSR and conventional rail lines, a bilevel programming model was applied by Deng et al. who exploited the user equilibrium theory to generate train stops together with a process of passenger flow assignment on trains [8].

From the aforementioned studies, we further examine three aspects of the TSSP, including two essential characteristics, connectivity and rapidity of a good train stop schedule, as well as the passenger flow assignment as follows. And we herein illustrate the motivation for this study.

- (1) Typically, a binary decision variable defining a stop being added at a station on a train or not is adopted. In doing this, as the number of stations becomes larger, it turns out to be hard to ensure a complete connectivity, meaning that travelers cannot always reach their destinations using direct or transfer connections. To the best of our knowledge, the first approach in the direction of guaranteeing connectivity is the use of a mixed integer linear programming (MILP), in which a binary decision variable indicating whether stop(s) is (are) added on a train for a passenger OD is considered in the problem formulation, working well in finding satisfactory solutions (see [9]).
- (2) Frequent train stops result in negative impacts including reducing train's rapidity, increasing passengers' total travel time as well as train operating cost. According to the literature, a common way to avoid the negative impacts of adding stops on trains is to define certain classes of trains and then restrict the total number of stops trains of each classification could make.

- (3) In the planning phase, each rail line generally tends to be organized towards certain passenger ODs, and a few of the major stations in a network are designed with good transfer capabilities. Accordingly, an ideal train stop schedule would intentionally organize passenger traffic that better resembles planner's purpose (i.e., passengers may be appropriately guided towards corresponding train services through using tool, e.g., a seat reservation system). However, previous research has not incorporated this idea in their stochastic passenger flow assignment procedures in the TSSP.

The main goal of the present work is to investigate a train stop scheduling approach that retains good connectivity and rapidity while also being able to cope with intentionally organizing passenger traffic in the passenger flow assignment procedure. For the purpose of retaining connectivity and rapidity, we decided to build on the recent approach of Fu et al. [9], using the same way of defining the binary decision variable. As an improvement, a restricted passenger flow assignment will be used in this paper to replace their normal formulation.

The paper is organized as follows. In Section 2, we formally describe the nominal version of our TSSP and illustrate the associated MILP formulation. In Section 3, we modify the MILP formulation by introducing a new passenger flow assignment constraint to deal with rail planner's purpose as mentioned previously. Numerical examples on real-world instances from the Chinese HSR network are presented in Section 4. Conclusions and future research are discussed in Section 5.

2. The Nominal TSSP

In this section, we describe the nominal TSSP that we consider in this paper. We restrict ourselves to the HSR network case used by Fu et al. [9] to be able to compare with their method and briefly recall the assumptions prior to illustrating the MILP formulation. Generally, in a rail network, trains running within a single rail line are known as in-line trains, and trains running across at least two linked up rail lines are known as cross-line trains. Both in-line and cross-line train OD patterns and potential train operating frequencies are predetermined as inputs for the model. Also, the track capacity is sufficient to meet the passenger demand by operating trains with rational load factors.

2.1. Notation. In the nominal TSSP, the aim is to effectively utilize capacity resources and to minimize passengers' generalized cost using a combination of train stop patterns. Our underlying approach partly borrows from Fu et al. [9] the notations of parameters and variables which are outlined below:

\mathcal{L} : set of trains among the given train OD patterns, in which ℓ_k represent trains indexed by k (or v as shown in the underlying formulation).

$\varepsilon(\ell_k)$: the number of train stop patterns that could be generated for train ℓ_k in set \mathcal{L} .

$f(\ell_k)$: estimated operating frequency of train ℓ_k .

V : set of stations in network, in which v_i represent stations (that can also be indexed by j , p , or q).

$v_i(\ell_k)$: stations (may not be stops) on the route of train ℓ_k .

E : set of tracks in network, in which e_l represent tracks indexed by l .

D : set of passenger ODs, in which $d(v_i, v_j)$ represent passenger demand between station v_i and v_j .

$D'(\ell_k)$: the collection of possible passenger ODs for which stops can be added on train ℓ_k , and $D'(\ell_k) = \{(v_i, v_j) \in D \mid \ell_k \in (v_i, v_j)\}$.

$\mathcal{L}(v_i, v_j)$: the collection of possible trains on which stops can be added for a passenger OD (v_i, v_j) , and $\mathcal{L}(v_i, v_j) = \{\ell_k \in L \mid (v_i, v_j) \in \ell_k\}$.

$\mathcal{L}(e_l)$: the collection of trains in set \mathcal{L} with their routes covering track e_l .

$A(v_i, v_j)$: set of feasible travel sections for travelers of passenger OD (v_i, v_j) within the given train OD patterns.

$a_n(\ell_k)$: passenger travel sections in set $A(v_i, v_j)$ indexed by n (or m); each travel section is uniquely on one train ℓ_k .

$\kappa(\ell_k)$: seating capacity of train ℓ_k .

$h(v_i, v_j)$: route length between stations v_i and v_j .

$N(\ell_k)$: the maximum number of stops that can be added on train ℓ_k .

$\eta(v_i, \ell_k)$: count parameter of whether a stop being added at station v_i on train ℓ_k .

$Y((v_i, v_{i+1}), \ell_k)$: accumulative passenger flow assigned on train ℓ_k between two adjacent stations v_i and v_{i+1} on the train route.

$T(a_n(\ell_k))$: passengers' generalized cost on travel sections $a_n(\ell_k)$.

$\tau^{it}(a_n(\ell_k))$: in-vehicle time on travel sections $a_n(\ell_k)$.

$\tau^{vt}(a_n(\ell_k))$: time converted from the ticket fares by time value on travel sections $a_n(\ell_k)$.

$\tau^{rt}(a_n(\ell_k))$: wait time on travel sections $a_n(\ell_k)$.

$\delta^{it}, \delta^{vt}, \delta^{rt}$: weights of generalized cost components.

$x((v_i, v_j), a_n(\ell_k))$: binary variable, it is 1 only if for passenger OD (v_i, v_j) , stops are added on train ℓ_k ; else it equals 0.

$y((v_i, v_j), a_n(\ell_k))$: variable of the passenger flow of OD (v_i, v_j) assigned on train ℓ_k with stops added on it.

2.2. Problem Formulation. The nominal MILP formulation for the TSSP is as follows:

$$\min \sum_{\ell_k \in \mathcal{L}} \sum_{v_i(\ell_k) \in V} (\kappa(\ell_k) \cdot f(\ell_k) - Y((v_i, v_{i+1}), \ell_k)) \cdot h(v_i, v_{i+1}), \quad (1)$$

$$\min \sum_{(v_i, v_j) \in D} \sum_{a_n(\ell_k) \in A(v_i, v_j)} T(a_n(\ell_k)) \cdot y((v_i, v_j), a_n(\ell_k)) \quad (2)$$

$$\text{s.t.} \quad \sum_{\ell_k \in \mathcal{L}(v_i, v_j)} y((v_i, v_j), a_n(\ell_k)) = d(v_i, v_j), \quad (3)$$

$$\forall (v_i, v_j) \in D,$$

$$\sum_{\ell_k \in \mathcal{L}(v_i, v_j)} x((v_i, v_j), a_n(\ell_k)) \cdot \kappa(\ell_k) \cdot f(\ell_k) \geq d(v_i, v_j), \quad \forall (v_i, v_j) \in D, \quad (4)$$

$$\sum_{(v_i, v_j) \in D'(\ell_k); a_n(\ell_k) \geq e_l} y((v_i, v_j), a_n(\ell_k)) \leq \kappa(\ell_k) \cdot f(\ell_k), \quad \forall e_l \in E, \ell_k \in \mathcal{L}(e_l) \quad (5)$$

$$\sum_{v_i \in V} \eta(v_i, \ell_k) \cdot x((v_i, v_j), a_n(\ell_k)) \leq N(\ell_k), \quad (6)$$

$$\forall \ell_k \in \mathcal{L},$$

$$y((v_i, v_j), a_n(\ell_k)) \leq M \cdot x((v_i, v_j), a_n(\ell_k)), \quad (7)$$

$$\forall \ell_k \in \mathcal{L}, (v_i, v_j) \in D,$$

$$x((v_i, v_j), a_n(\ell_k)) \in \{0, 1\}, \quad \forall \ell_k \in \mathcal{L}, (v_i, v_j) \in D, \quad (8)$$

$$y((v_i, v_j), a_n(\ell_k)) \in \mathbb{R}_+, \quad \forall \ell_k \in \mathcal{L}, (v_i, v_j) \in D. \quad (9)$$

Objective function (1) minimizes total trains' deadhead kilometers, where

$$Y((v_i, v_{i+1}), \ell_k) = \sum_{v_j \in V} \sum_{a_n(\ell_k) \in A(v_i, v_j)} y((v_i, v_j), a_n(\ell_k)), \quad (10)$$

$$\forall v_i(\ell_k) \in V.$$

Passengers' generalized cost in objective function (2) consists of three parts: in-vehicle time, consuming time converted from the ticket fares by time value, and wait time. Thus, $T(a_n(\ell_k))$ extends as

$$T(a_n(\ell_k)) = \delta^{it} \cdot \tau^{it}(a_n(\ell_k)) + \delta^{vt} \cdot \tau^{vt}(a_n(\ell_k)) + \delta^{rt} \cdot \tau^{rt}(a_n(\ell_k)). \quad (11)$$

Because train sets running on the same section do not necessarily operate at a uniform speed, different passengers for a given OD may have different in-vehicle time. Similarly, the ticket fares pricing would adopt a differentiation strategy in terms of either train speed classifications or being based on in-line trains and cross-line trains. Wait time depends on train operating frequencies.

Constraint (3) imposes passenger flow conservation in the assignment process. Demand-supply constraints are illustrated in (4) and (5). Constraint (4) ensures the total train

stop frequencies at a given station are adequate to meet the passenger demand requirements. Constraint (5) denotes that the flow of different passenger ODs assigned on a given train ℓ_k should not exceed that train's seating capacity. The condition of " $a_n(\ell_k) \geq e_l$ " means that only passenger OD(s) using travel sections $a_n(\ell_k)$ which pass through track e_l is (are) taken into account. Constraint (6) limits the maximum number of stops on train ℓ_k . The count parameter $\eta(v_i, \ell_k)$ equals 0 if a stop will not be added at station v_i ; and equals 1 if station v_i is to be added as a stop for more than one passenger OD. Constraint (7) ensures that if stop(s) is (are) not added on train ℓ_k for passenger OD (v_i, v_j) , its flow assigned on train ℓ_k equals 0; M is a very large positive number. Finally, the two types of decision variables are restricted in constraints (8) and (9).

3. Our TSSP with Restricted Passenger Flow Assignment

Passengers may have a large set of travel paths from which they choose without any restrictions. A main drawback of the nominal TSSP illustrated in the previous section is that it assigns passengers onto trains in a stochastic way, not taking into account guiding passengers towards certain train services on planner's capacity resources allocation strategy. Stochastic assignment can strongly affect the quality of train capacity occupancy, causing unbalanced problems; for example, short-distance travelers may preempt seats of long-distance travelers and be served by a long-distance train, while long-distance travelers are expelled out of that service. Additionally, the formulation ignores and is not easy to describe transfer behavior of travelers. Therefore, train service connectivity for travelers that mostly obtained from scheduling additional train stops would instead depend on the quality of the given train OD patterns.

To overcome these problems, and stimulated by applications of routing passenger travel paths freely in a public transit network discussed by, for example, Goossens et al. [3] and Borndörfer et al. [10], we convert the nominal TSSP into a TSSP embedded with a multicommodity flow problem (MCFP) by modifying constraint (3). Subsequently, we are able to assign passengers onto trains in a restricted way, via incorporating additional restrictions on generating passenger travel paths. Our modified version of the MILP model (1)–(9) includes objective functions (1) and (2), constraints (4)–(9) but uses the following new constraint:

$$\begin{aligned}
 & \sum_{v_q \in V, a_n(\ell_k):=(v_p, v_q)} y((v_i, v_j), a_n(\ell_k)) \\
 & - \sum_{v_q \in V, a_n(\ell_k):=(v_q, v_p)} y((v_i, v_j), a_n(\ell_k)) \\
 & = \begin{cases} d(v_i, v_j) & \text{if } v_p = v_i \\ 0 & \text{if } v_i \neq v_p \neq v_j \\ -d(v_i, v_j) & \text{if } v_p = v_j, \end{cases} \quad (12) \\
 & \forall (v_i, v_j) \in D, v_p \in V.
 \end{aligned}$$

In both versions of the MILP models, the passenger flow is assigned on a section (i.e., arc) basis (the reader interested in the MCFP applied for passenger flow assignment on a path basis is referred to, e.g., [11]). A travel section in the nominal version equates to a travel path. Comparatively, since transfer can be considered in constraint (12), a travel path now becomes splittable. That is to say, a travel path consists of at least one travel section, two travel sections compose a travel path if one transfer occurs during one single trip, and so forth. Accordingly in the modified version, we attach transfer time on travel sections $a_n(\ell_k)$ in (11) denoted by $\tau^{tt}(a_n(\ell_k))$ and assign a weight δ^{tt} to them; (11) thus reads

$$\begin{aligned}
 T(a_n(\ell_k)) &= \delta^{it} \cdot \tau^{it}(a_n(\ell_k)) + \delta^{vt} \cdot \tau^{vt}(a_n(\ell_k)) \\
 &+ \delta^{rt} \cdot \tau^{rt}(a_n(\ell_k)) + \delta^{tt} \cdot \tau^{tt}(a_n(\ell_k)). \quad (13)
 \end{aligned}$$

Apparently, $\tau^{tt}(a_n(\ell_k))$ is not equal to zero only if the section is a part of one travel path with transfer(s).

As the foundation of producing restrictions on generating passenger travel paths, we first recall the rail planner's purpose as already mentioned in the introduction section and make some further interpretations. A passenger rail line as part of a rail network is built, generally, tending to be organized towards certain passenger ODs along the corridor of priority. For the sake of efficient train capacity occupancy, in-line travelers should be organized onto in-line trains as much as possible, which is a practical representation of the important principle of that passenger travel distance should match train trip distance. Second, to adapt to passenger distributing capacity, stations with eligible facilities are recommended as main transfer hubs spread all over the entire rail network. From a functional perspective, transfer hubs have good performance if one transfer occurs within the same platform, and trains have convenient as well as fast connections with other high-speed or conventional trains or urban public transport.

We next illustrate our adopted restrictions in consideration of rail planner's objectives when generating passenger travel paths in the MCFP. A general rule is to limit a feasible travel paths set for passengers of each OD, using two primary criteria as follows.

- (1) A travel path, or a travel section being part of a path with transfer(s) possible for passengers of a certain OD is feasible, only if the ratio between the length itself and the trip distance of a train onto which passengers are probably assigned is greater than a given baseline value (notated as μ).
- (2) A travel path with transfer(s) is feasible if the transfer hub(s) is (are) selected from recommended stations.

From the viewpoint of providing better quality train services, alternative restrictions include the following:

- (3) a controlled percentage of nonstop trains for travelers are provided;
- (4) a single passenger is not required to transfer more than once during a single trip.

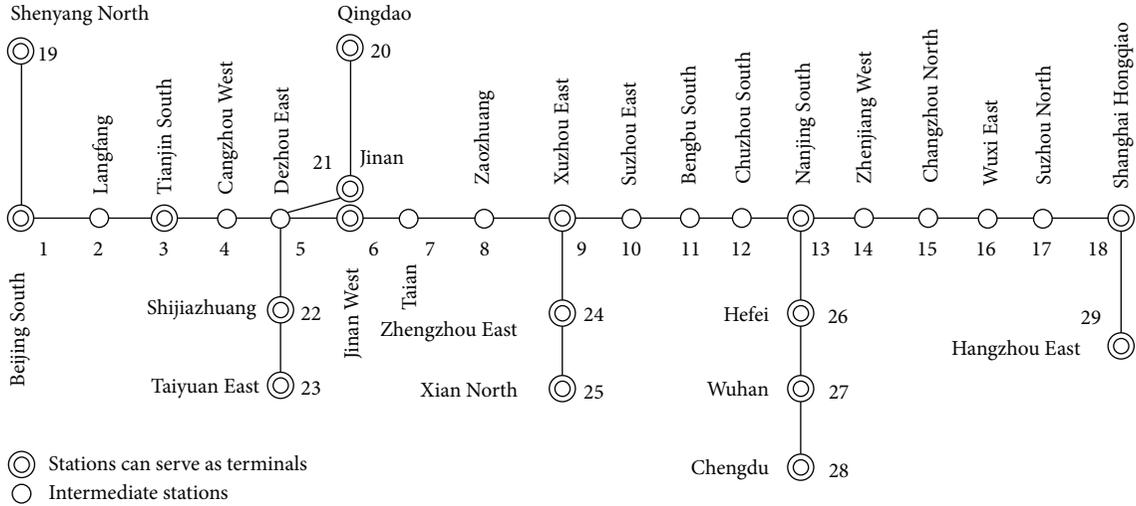


FIGURE 1: Topology of study HSR network.

Of course, many other restrictions on paths generation are allowed in practice; a rail operator could adopt many other criteria to define the feasibility of a path and choose a desirable set of paths, depending on the specific requirements. For example, even if criterion (1) is met, a rail operator could additionally delete travel paths on a train if that train precedes other trains not serving passengers among smaller stations. Certain travel paths could be likewise specified to passengers, for whom, in extreme case, no travel paths can be found by following the given criteria.

4. Numerical Examples

4.1. Data Set. This section presents a numerical test of the proposed approach for the TSSP using data from a real-world example on the Chinese railways. We focus our attention on the example adopted by [9] to allow for comparisons. For the sake of clarity, we stretch their real-world HSR network to be a topology with each station numbered as shown in Figure 1. Additional information reflecting each line or section’s operating condition is shown in Table 1. For the 2015 forecast year, our numerical example includes a total of 303 passenger ODs and total 453,205 passenger traffic per day associated with the Jinghu HSR (the center of the network, we abbreviate it as JHSR in the following). Due to data sizes, passenger flow OD matrix is not shown but is available from the authors. Table 2 shows the train OD patterns as given from the operational plan, which includes 7 in-line trains and 29 cross-line trains for the JHSR network. Other input parameters of our models are specified in Table 3.

4.2. Discussion of Results. Progressing to our appropriate code based on the original one (used in [9]), both models in two versions were applied to the network and solved on an Intel i5 2.4 GHz with 2 GB RAM in the environment of Microsoft Windows XP using the Lingo 10.0 optimizer. Note that in the modified version, we incorporated all the four restrictions into predetermined feasible travel paths set. In

addition to the already mentioned criteria, we provided that the terminal stations in Figure 1 served as transfer hubs, and 70 percent of travelers of one OD were organized towards nonstop trains with load factor not less than 0.85. During the solving process, a layered sequence method was employed to convert the multiobjective problem to a single-objective one. The models were first solved under single-objective (1); its optimization results were put into constraints as

$$\sum_{\ell_k \in \mathcal{L}} \sum_{v_i, v_{i+1} \in V} (\kappa(\ell_k) \cdot f(\ell_k) - Y((v_i, v_{i+1}), \ell_k)) \cdot h(v_i, v_{i+1}) = M^* \quad (14)$$

with M^* being the optimum solution; afterwards the problems were optimized under single-objective (2).

Using a branch-and-bound algorithm, our restricted TSSP took 57 minutes of computing time to find an optimum solution compared to 36 minutes for the nominal TSSP. The final train stop schedule obtained for the restricted TSSP is displayed in Table 4 and a comparison of our method with the method for the nominal TSSP is shown in Table 5. Due to the increase in complexity, as expected, the efficiency gotten for our solution is lower than that of the nominal one, but the gain in most operational indicators (to be discussed later) is quite significant and confirms the effectiveness of the proposed approach. Because the TSSP has to be solved in the planning phase, the computing time is acceptable.

According to Table 5, although it is a straightforward modification of an existing approach for the nominal case, our approach turns out to be effective in scheduling train stops of good quality, as indicated by operational metrics which clearly dominate those of the nominal TSSP. In spite of the differences, the two schedules have some similarities. First, both methods result in a flexible combination of train stop patterns, that is, nonstop, skip-stop, and all-stop. However, an “all-stop” train stop pattern is only recommended for short- or medium-distance trains (e.g., the train from station 13 to 18). Another similarity between the two methods is

TABLE 1: Line or section information of study HSR network.

HSR line or section	The year in operation	Approximate length (km)	Speed classification (km/h)
Jinghu (1-18)	2011	1,318	350
Jingshen (1-19)	—	676	350
Jiaoji (21-20)	2008	362	200~250
Shijiazhuang-Dezhou	—	180	250
Shitai (22-23)	2009	190	250
Zhengxu (24-9)	—	360	350
Zhengxi (24-25)	2010	458	350
Hening (26-13)	2008	166	200
Hewu (26-27)	2008	356	200
Wuhan-Chengdu	—	1,260	200
Huhang (18-29)	2010	158	350

“—” indicates that the rail line or section is planned or not completely in operation.

TABLE 2: The given train OD patterns.

In-line train ODs	Cross-line train ODs	Train ODs for transfer connections only
	(1, 20) (1, 26) (1, 29) (3, 20) (3, 27) (3, 29)	
	(6, 19) (6, 27) (6, 29) (9, 20) (9, 26) (9, 29)	
(1, 13) (1, 18) (3, 18) (6, 13) (6, 18) (9, 18)	(13, 20) (13, 23) (13, 25) (13, 29) (18, 19)	(19, 1) (20, 21) (23, 6) (25, 9) (24, 9) (28, 13)
(13, 18)	(18, 20) (18, 23) (18, 25) (18, 24) (18, 28)	(27, 13) (26, 13) (28, 27) (28, 26)
	(18, 27) (18, 26) (20, 19) (20, 27) (20, 25)	
	(29, 25) (29, 24)	

TABLE 3: Input parameters of the models.

Parameter	Value or descriptive calculation
Split coefficient $\varepsilon(\ell_k)$	11 for in-line trains and 3 for cross-line trains
Seating capacity $\kappa(\ell_k)$	1,060 seats/train-set
Number of train stops $N(\ell_k)$	Maximum of 7 times for in-line trains, and 8 for cross-line trains
In-vehicle time $\tau^{it}(a_n(\ell_k))$	Train route length $h(v_i, v_j)$ /train travel speed (accommodates to the rail line's speed classification)
Consuming time $\tau^{vt}(a_n(\ell_k))$	Ticket fares (0.55 RMB/km for 350 km/h train, 0.45 RMB/km for other trains) $\times h(v_i, v_j)$ /time value of passenger (45 RMB/h on average)
Wait time $\tau^{rt}(a_n(\ell_k))$	$0.50/f(\ell_k)$: a fraction (taken as 0.50 here) of headway which is the inverse of a train's operating frequency
Transfer time $\tau^{tt}(a_n(\ell_k))$	30 minutes
Weights δ^{it} , δ^{vt} , δ^{rt} , and δ^{tt}	0.39, 0.28, 0.12, 0.21
Baseline value μ	0.80 for in-line travelers and 0.50 for cross-line (transfer) travelers

that, in the interaction with objective functions, the obtained train stop schedules perform even better. For instance, under the schedule for the restricted TSSP, 73 percent of trains are assigned to a number of stops less than restricted as compared with 65 percent by the nominal TSSP. This difference is due to a hard rule applied in the restricted TSSP that more nonstop trains are allowed operations, and the consequence arising from this influences the average train travel speeds in the same way. In both methods, the entirely assured connectivity for travelers is not surprising because travelers are always tracked on which train(s) they are assigned and stops are accordingly determined by defining such type of a decision variable. For example, the train “9-11-13-14-17-18” stops at four intermediate stations for estimated travelers of 10 passenger

ODs assigned on it, even though in practice, travelers of all 15 combinatorial passenger ODs can be absolutely served by that train. Therefore, from a practical perspective, the number of possible passenger OD combinations served by each train is necessarily higher than the measures from the computing results of passenger flow assignment.

The indicator of realized passenger traffic represents a notable difference between the two methods. The realized passenger traffic of our restricted TSSP is significantly larger. As already mentioned in Section 3, this is achieved by allowing transfers in travel paths. However, the gain in reducing passengers' generalized cost is not that much higher when compared to the nominal TSSP. To make a more accurate comparison, we calculate generalized cost on the scale of

TABLE 4: Train stop schedule estimated from the method of our restricted TSSP.

Train OD	Train route no.	Train stop pattern with station sequence
In-line trains		
1, 13	1	1-9-13
	2	1-2-4-8-9-10-11-12-13
1, 18	1	1-18
	2	1-3-18
	3	1-3-6-13-18
	4	1-3-5-6-7-13-18
	5	1-3-6-11-13-16-17-18
	6	1-2-3-5-7-8-9-16-18
	7	1-2-3-4-5-9-15-17-18
	8	1-2-3-6-9-10-11-13-18
	9	1-2-3-4-9-10-12-14-18
	10	1-3-6-9-11-13-15-17-18
3, 18	1	3-18
	2	3-6-18
	3	3-6-9-18
	4	3-5-6-7-8-15-17-18
	5	3-4-5-6-7-8-12-17-18
6, 13	1	6-13
	2	6-7-9-10-11-12-13
6, 18	1	6-9-18
	2	6-13-18
	3	6-7-8-9-14-16-17-18
	4	6-10-11-12-13-16-17-18
9, 18	1	9-13-18
	2	9-11-13-14-17-18
13, 18	1	13-18
	2	13-14-15-16-17-18
Cross-line trains		
1, 20	1	1-3-20
	2	1-2-3-4-5-20
1, 26	1	1-3-6-9-26
	2	1-3-5-6-8-11-26
	3	1-3-4-5-6-7-9-10-11-26
1, 29	1	1-3-6-9-13-29
3, 20	1	3-5-20
3, 27	1	3-9-13-26-27
	2	3-4-7-10-12-27
3, 29	1	3-5-6-10-11-13-18-29
19, 6	1	19-1-5-6
6, 27	1	6-8-26-27

TABLE 4: Continued.

Train OD	Train route no.	Train stop pattern with station sequence
6, 29	1	6-7-9-29
	2	6-8-13-18-29
20, 9	1	20-6-9
9, 26	1	9-11-26
9, 29	1	9-11-13-17-29
20, 13	1	20-6-13
23, 13	1	23-22-6-7-8-10-11-13
25, 13	1	25-9-13
	1	13-18-29
13, 29	2	13-14-15-16-17-18-29
	1	19-1-3-6-9-13-18
19, 18	2	19-2-7-8-9-10-11-13-17-18
	3	19-2-4-5-9-13-14-15-16-18
20, 18	1	20-6-14-15-16-17-18
23, 18	1	23-22-9-14-15-16-17-18
25, 18	1	25-9-12-13-18
	2	25-9-10-11-13-14-15-16-17-18
24, 18	1	24-9-13-18
	2	24-9-10-11-12-14-15-16-17-18
28, 18	1	28-13-17-18
27, 18	1	27-13-15-16-17-18
26, 18	1	26-13-18
	2	26-13-15-16-17-18
19, 20	1	19-3-4-5-20
20, 27	1	20-9-10-11-27
20, 25	1	20-6-8-25
25, 29	1	25-9-12-13-18-29
24, 29	1	24-9-13-18-29

an individual traveler. For the case of the restricted TSSP, the total passengers' generalized cost is 3.59392×10^6 (h) while the cost is 3.12417×10^6 (h) for the nominal TSSP. This represents a saving of approximately 0.27 hours per traveler. This small but meaningful amount of time saving is attributed to that, in the restricted TSSP, the decrease of travel time via operating nonstop trains is greatly offset by the increase of transfer time.

The effectiveness of the approach of our restricted TSSP is also supported by another operational indicator. It is apparent from Table 5 that the matching degree measured by μ for the nominal TSSP is much lower, since no limitation is imposed on μ when assigning each passenger to his optimal itinerary. An operational outcome of this result is that the average train deadhead kilometers is nearly 4 percent higher in the nominal TSSP compared to our restricted TSSP.

TABLE 5: Comparison of operational indicators between methods for our restricted TSSP and the nominal TSSP.

Operational indicator	Restricted TSSP	Nominal TSSP
Using combinatorial train stop patterns or not	Yes	Yes
Percentage of trains with number of stops less than restricted (%)	73	65
Average train travel speed (km/h)	301	294
Realized passenger traffic by train services (people)	453,205	380,772
Connectivity degree for realized passenger traffic (%)	100	100
Percentage of travelers organized onto nonstop trains (%)	33	21
Matching degree measured by μ (%)	100	51
Percentage of transfer passengers (%)	9	0

Transfer hubs in travel paths are distributed at stations mostly on the JHSR as restricted by the model. Stations 1, 3, 6, 9, and 13 assemble approximately 92 percent of transfer passengers. Due to being covered by travel paths of a larger amount of cross-line passenger ODs in this instance, stations 9 and 13 together are estimated to have a higher probability of being transfer hubs selected by approximately 65 percent of transfer passengers. Obviously, there are clear benefits in assisting rail operator to plan passenger transfer organization work as part of the TSSP and, as an added benefit, the inconvenience of making a transfer connection is reduced relative to the convenience of direct connection.

5. Conclusions

Rail operators develop train stop schedules with the goals of retaining good connectivity and rapidity to travelers while also in the face of requirements for capacity resources utilization. In this paper, we have shown how to incorporate restricted passenger flow assignment into a TSSP formulation to achieve this purpose. To this end, two procedures need to be implemented: (1) introducing the MCFP constraint intended to route passenger travel paths freely and (2) during passenger travel paths generation, establishing four criteria to produce restrictions so that the operator can collect a desirable set of travel paths. Our approach has been applied to a real-world HSR network case from the Chinese railways along with a comparison with a nominal train stop scheduling method that uses stochastic passenger flow assignment. The results showed that our approach is very competitive and obtains a train stop schedule solution of good quality in acceptable computing time. Future direction of research into efficient formulation of the TSSP can be devoted to collaboratively optimize train operating frequency which is treated as constant value in the present paper.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (Beijing Jiaotong University) under Grant no. 2013JBM042. The authors are grateful to the referees for their valuable comments and suggestions.

References

- [1] H. Guo, W. Wang, W. Guo, X. Jiang, and H. Bubb, "Reliability analysis of pedestrian safety crossing in urban traffic environment," *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.
- [2] H. W. Hamacher, A. Liebers, A. Schöbel, D. Wagner, and F. Wagner, "Locating new stops in a railway network," *Electronic Notes in Theoretical Computer Science*, vol. 50, no. 1, pp. 13–23, 2001.
- [3] J.-W. Goossens, S. van Hoesel, and L. Kroon, "On solving multi-type railway line planning problems," *European Journal of Operational Research*, vol. 168, no. 2, pp. 403–424, 2006.
- [4] Y. Y. Ulusoy, S. Chien, and C.-H. Wei, "Optimal all-stop, short-turn, and express transit services under heterogeneous demand," *Transportation Research Record*, no. 2197, pp. 8–18, 2010.
- [5] Y.-H. Chang, C.-H. Yeh, and C.-C. Shen, "A multiobjective model for passenger train services planning: application to Taiwan's high-speed rail line," *Transportation Research B*, vol. 34, no. 2, pp. 91–106, 2000.
- [6] C. Lee and W. Hsieh, "A demand oriented service planning process," in *Proceedings of the World Congress on Railway Research*, Koln, Germany, 2001.
- [7] Y. Zhang, M. Ren, and W. Du, "Optimization of high speed train operation," *Journal of Southwest Jiaotong University*, vol. 33, no. 4, pp. 400–404, 1998 (Chinese).
- [8] L. Deng, F. Shi, and W. Zhou, "Stop schedule plan optimization for passenger train," *China Railway Science*, vol. 30, no. 4, pp. 102–107, 2009.
- [9] H. Fu, L. Nie, B. R. Sperry, and Z. He, "Train stop scheduling in a high-speed rail network by utilizing a two-stage approach," *Mathematical Problems in Engineering*, vol. 20102, Article ID 579130, 11 pages, 2012.
- [10] R. Borndörfer, M. Grötschel, and M. E. Pfetsch, "A column-generation approach to line planning in public transport," *Transportation Science*, vol. 41, no. 1, pp. 123–132, 2007.
- [11] R. Borndörfer, M. Grötschel, and M. E. Pfetsch, "Models for line planning in public transport," ZIP-Report 04-10, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, Germany, 2004.

Research Article

Estimation Trajectory of the Low-Frequency Floating Car Considering the Traffic Control

Zhijian Wang, Min Li, Li Wang, and Xiaoming Liu

Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Beijing 100144, China

Correspondence should be addressed to Zhijian Wang; wzjian0722@163.com

Received 14 August 2013; Revised 30 October 2013; Accepted 18 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Zhijian Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Floating car equipped with GPS to detect traffic flow has been widely used in ITS research and applications. The trajectory estimation is the most critical and complex part in the floating vehicle information processing system. However, the trajectory estimation would be more difficult when using the low-frequency data sampling because of the high communication cost and the numerous data. Specifically, the ordinary algorithm cannot determine the specific vehicle paths with two anchor points across multiple intersections. Considering the accuracy in map matching, this paper used a delay matching algorithm and studied the trajectory estimation algorithm focusing on the issue of existence of a small road network between two anchor points. A method considering the three multiobjective factors of signal control and driving distance and number of intersections was developed. Firstly, an optimal solution set was acquired according to multiobjective decision theory and Pareto optimal principles in game theory. Then, the optimal solution set was evaluated synthetically based on the fuzzy set theory. Finally, the candidate trajectory which is the core evaluation factor was identified as the best possible travel path. The algorithm was validated by using the real traffic data in Wangjing area of Beijing. The results showed that the algorithm can get a better trajectory estimation and provide more traffic information to traffic management department.

1. Introduction

With the rapid development of economic, travel efficiently tends to decline in recent years because of the congestion of traffic road network. Road travel time used in the dynamic traffic information service represents the key parameter of traffic congestion information. The quality and timeliness of road travel time determined the success of dynamic traffic guidance system. How to accurately acquire the travel time information of road network is very critical for the traffic guidance system. With the development of GPS (Global Positioning System) and GIS (Geographic Information System) technology, floating car system as a new traffic information acquisition method has the advantage as follows: short construction period, less investment, large range cover, high precision data, and better real time. Such systems have become a new acquisition method of dynamic traffic information instead of traffic fixed detector for high maintenance cost and big investment [1–3].

Although the floating car system as the dynamic information acquisition has a short history, it has become the

research focus and is largely applied in the traffic control and traffic information service in foreign countries. In the famous ADVANCE project of American, the utilization of floating car system showed the importance in the traffic guidance system [4]. The German Aerospace Center built the floating car system based on the taxi, which can receive the mass GPS data to estimate the travel time for the traffic control center [5]. The Nagoya of Japanese has done the biggest experiment of floating car system in 2003–2007, which used the location information of 1,500 floating car to estimate and judge the real time traffic state of road network and publish the traffic information so as to distribute the traffic flow and ease the traffic congestion [6].

Some floating car system researches aimed at dealing with the urban traffic congestion were completed by the domestic traffic experts and scholars in recent years. The transportation research center of Beijing has built the road network traffic intelligent analysis system based on the 7,000 floating cars, which integrated the map matching and travel time estimation for road network state judgment and dynamic traffic information release through the real time position

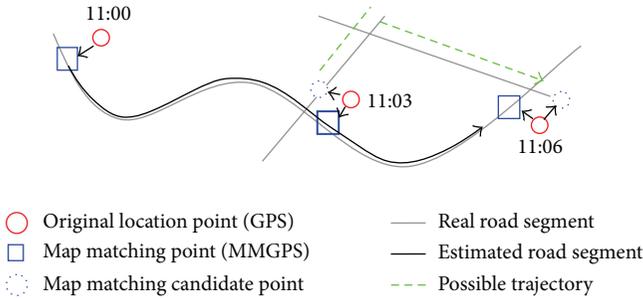


FIGURE 1: Map matching and trajectory estimation of Low-frequency GPS data.

platform [7]. Based on the 15,000 floating cars, Transport Commission of Guangzhou city built the intelligent navigation and location information service system for the 2010 Guangzhou Asian Games [8]. Similarity, Hangzhou city have installed the GPS for the 10,000 taxis to schedule the vehicle and road network state judgment, which also integrated the microwave detector data of road segment [9].

2. Research and Analysis

Although the floating car system was largely applied in traffic management, some difficulties still existed. The trajectory estimation of floating vehicles is an emphasis of the floating vehicle information processing system, in particular, in the case of using the low-frequency data sampling because of the high communication cost and significant data redundancy. This paper focuses on the trajectory estimation research, but firstly, there are two problems as follows.

2.1. The Accuracy of Map Matching before the Trajectory Estimation. In the floating vehicle information processing system, the result of accurate map matching is required before performing trajectory estimation and can largely affect the estimated results. Map matching algorithm has relation with the vehicle longitude and latitude, speed, vehicle travel angle, and the road topology composition. The floating car may run through some road segments in the GPS sampling period for the complicated topology between the adjacent GPS points. In addition, high-rise and viaduct occlusion cause the vehicle GPS data loss and frequent dynamic drift. As shown in Figure 1, if the anchor point was not accurately positioned after map matching, the trajectory estimation might come with two completely different travel paths.

2.2. The Difficulty of Estimating the Vehicle Trajectory. For the situation shown in Figures 2 and 3, if we connected simply with two anchor points, vehicle tracks will appear in the nonroad area. In Figure 2, we can replace the straight line p_1p_2 with the p_1Ap_2 as running track of the vehicle. However, the matched GPS point 1 and point 2 have a small network in Figure 3, and we need to apply the specific trajectory estimation method to determine the vehicle most possible route.

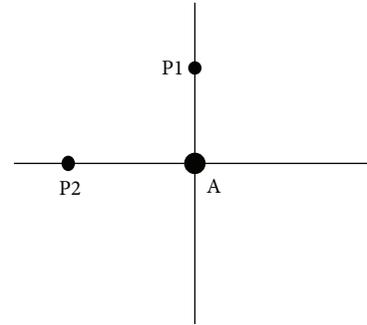


FIGURE 2: Two anchor points on the adjacent road.

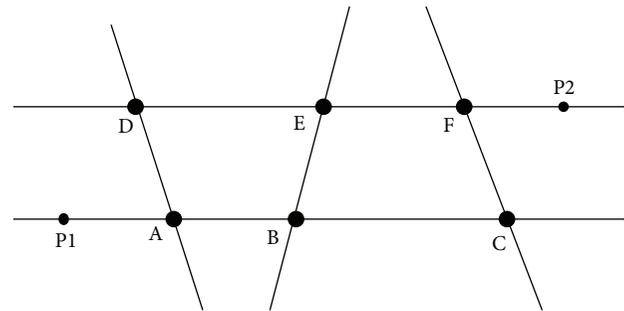


FIGURE 3: A small network between the matched GPS p_1 and p_2 .

3. Map Matching Algorithm

For the first problem of research and analysis in the above, we know that the result of accurate map matching is essential before performing trajectory estimation. Map-matching algorithm based on the matching principle are mainly geometric matching algorithm, matching algorithm based on fuzzy logic, matching algorithm based on the cost function, and matching algorithm based on pattern recognition. For map-matching accuracy problems, White et al. [10] discussed some simple map matching algorithms that can be used to reconcile inaccurate location data with an inaccurate map/network. Then he discussed point-to-point, point-to-curve, and curve-to-curve matching. But in some situations, directions do not change much as a result of small errors in the map matched location. Taylor et al. [11] proposed an improved map matching algorithm for GPS data in inaccurate conditions. The experiment of London to match the 15,000 sampling data can prove the higher accuracy algorithm. Based on analyzing the factors of affecting real time performance, robustness, and matching precision of the map matching algorithm and according to the continuity of the vehicle movement, the idea of partitioning the road net into some grids was introduced. Joshi [12] proposed a new approach to map matching for in-vehicle navigation systems. The paper presented a new metric, the rotational variation metric, which provided a new method for comparing vehicular and map paths for the purpose of map-matching. Quddus et al. [13] provided an improved map matching algorithm based on probability and statistics, the main features of which are the vehicle positioning error, the vehicle history trajectory,

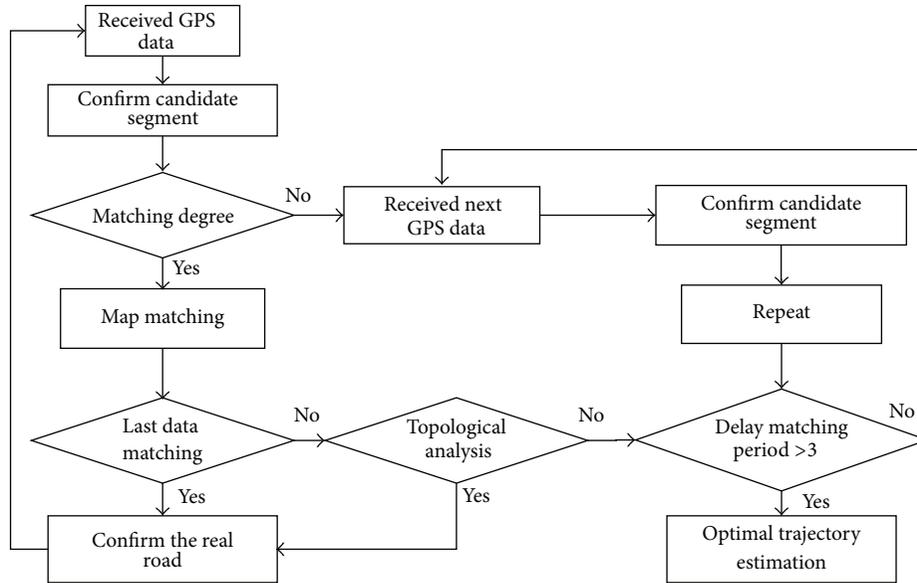


FIGURE 4: Chart of delay map matching algorithm.

and topological and vehicle speed and direction angle information all into account, solving the map matching when positioning data or the electronic map data was not accurate. In a word, now the map matching algorithm mainly aimed to solve the problem of the floating car in which region, which is influenced by the hall building occlusion and data loss. We need to determine the floating car driving in which road based on the topological relation and GPS data information so as to provide more accurate vehicle driving information.

The author [14, 15] has previously been employed the algorithm of delay map matching based on road topology structure which can be used when the vehicle cannot match the data at the first time. The actual testing data indicated that the new algorithm had a great instantaneity and accuracy in map matching, which effectively solved the problem of traffic data collection in complex intersection and area of viaduct bridge. The algorithm implementation and results are as in Figures 4, 5, 6, 7, and 8.

4. Trajectory Estimation

4.1. Analysis of Study Methods. In the second problem of research and analysis in the above, the GPS positioning cycle is relatively long so the vehicle may run through the more road section. In Figure 3, the matched GPS point 1 and point 2 have a small network, in which the starting point and ending point have more than one path. In the situation we cannot confirm the floating car driver runs which route utilizing the related traffic information. So we need to apply the specific trajectory estimation method to determine the vehicle most possible route.

Takada [16] proposed the optimal floating car trajectory function, considering the location, estimation potential points, the driver preference, and travel characteristics which the traveler prefers to run the distance shorter route. Wang [8] proposed the trajectory estimation algorithm based on

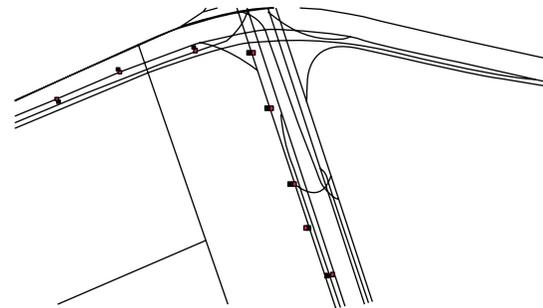


FIGURE 5: Point-map matching when around viaduct.

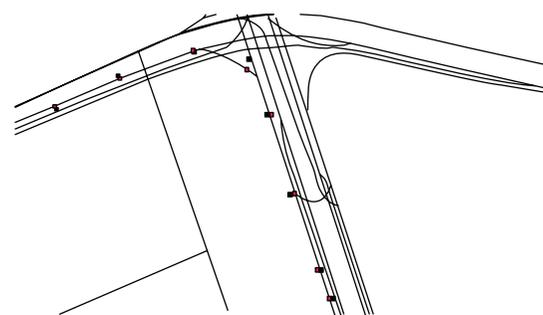


FIGURE 6: Delay map matching when around viaduct.

the mobile phone location data, considering the vehicle acceleration and deceleration to research the relationship of the travel time, which is according to the path searching to determine the vehicle trajectory. Chen [17] introduced Pareto multiobjective optimization and fuzzy theory to comprehensively judge the most likely path for each tracking problem. During this process, the driving distance and the number of passing crossroads were set as two optimal objectives. Then

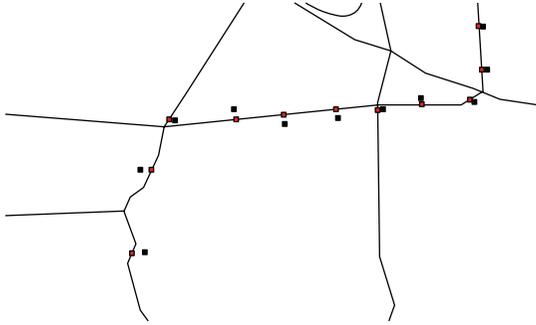


FIGURE 7: Point-map matching around intersection.

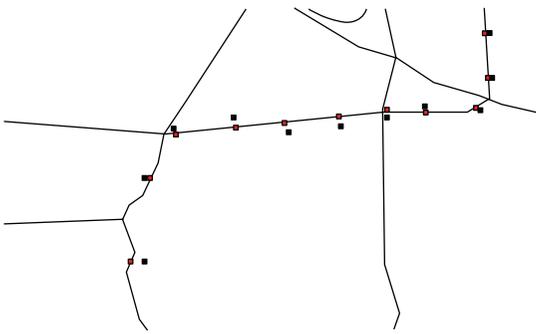


FIGURE 8: Delay map matching around intersection.

he distributed the average speed of each path proportionally to the road segments it covers, which integrated all the speed contributions a road segment collects for a final estimation of its traffic state. Kerner et al. [18] introduced a method for a reporting behavior at optimal costs of single vehicles (FCD: Floating Car Data) in road networks with the aim of a high quality of traffic state recognition which is presented. It is shown that based on minimum two FCD messages the substantial information of a typical traffic incident in a traffic center can be recognized. Byon et al. [19] introduced a method collecting and analysis traffic conditions of links by monitoring speed of probe vehicle(s) and then estimates travel time data both in static and dynamic modes. The static mode refers to the case of offline processing of GPS data from previously dispatched GPS-equipped vehicles to specified road links. The dynamic mode refers to the real time monitoring of speed on the links using a GPS-wireless Internet-equipped probe vehicle. In a word, the existing vehicle trajectory estimation algorithms usually do not consider the vehicles trajectory affected by signal control cycle, only considering the shortest path so the method is easy to cause an estimated error.

This paper focused on these issues that introduce a method of combining Pareto multiobjective optimization and fuzzy theory considering these three multiobjective factors of signal control, driving distance, and the number of intersections.

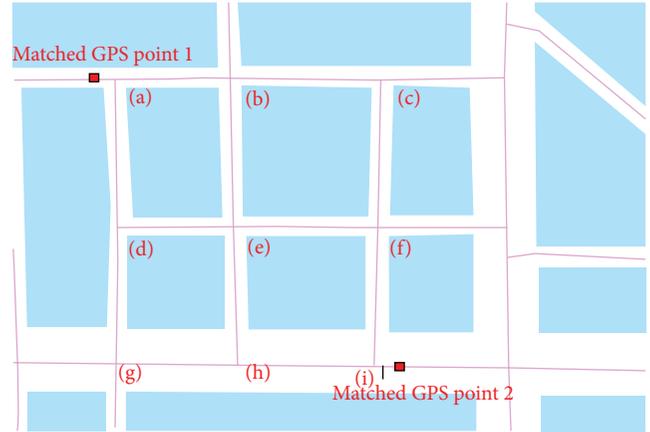


FIGURE 9: Actual road network of vehicle tracking.

4.2. Multiobjective Decision Theory and Pareto Optimal Principles

4.2.1. Multiobjective Decision Theory. When the decision object has a plurality of evaluation target, we can choose a satisfactory decision method from several feasible schemes (also known as the solution). According to the advance evaluation criterion we choose through the “priority” and “balance” to find a satisfactory solution from a set of nondominated solutions.

Multiobjective optimization problem was first developed by Italian economist Pareto L. put forward in 1896. In 1944 Von Neumann and Morgenstern [20] put forward conflicting multiobjective decision making with multiple decision-makers from the perspective of game theory. Koopmans [21] in 1951 presented in multiobjective optimization problem from the production and distribution activities analysis and introduces the concept of Pareto optimization.

There are three possible results of possible options for a multiobjective decision problem: first, the solution in which all the objectives are the best is called complete optimal solution and this situation rarely occurs; second, the solution in which all the objectives are the worst is called the inferior solutions and can be immediately eliminated; third, the solution in which the objectives have good and bad goals is called noninferior solutions, also known as the Pareto optimal or efficient solution.

Multiobjective optimization mathematical model is assuming system has y objectives $\Pi_1(\omega), \Pi_2(\omega), \dots, \Pi_y(\omega)$, the target vector $k = (k_1, k_2, \dots, k_y)$ by the y variables needs to be evaluated. If these objectives are required the maximum (or minimum) and required to satisfy the constraints set R , then the mathematical model can be expressed as $\max_{x \in R} F(x)$ or $\min_{x \in R} F(x)$, in which $F(\omega) = (\Pi_1(\omega), \Pi_2(\omega), \dots, \Pi_y(\omega))$.

Let us analyze the actual road network in Figure 9. The matched GPS point 1 and point 2 have more than 10 routes from which the driver can choose. What factor is the driver most likely to concern about when driving on the road network? According to the survey result, it is that arriving to the destination as soon as possible. Considering the influence

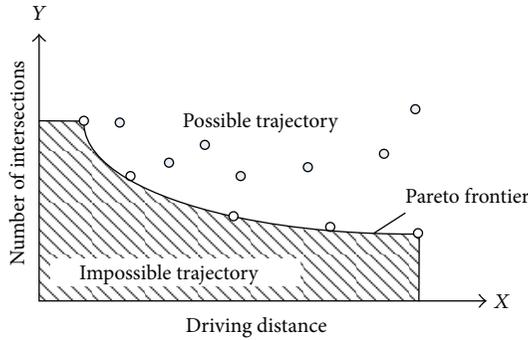


FIGURE 10: Vehicle possible trajectory in 2D plot of intersections number and driving distance.

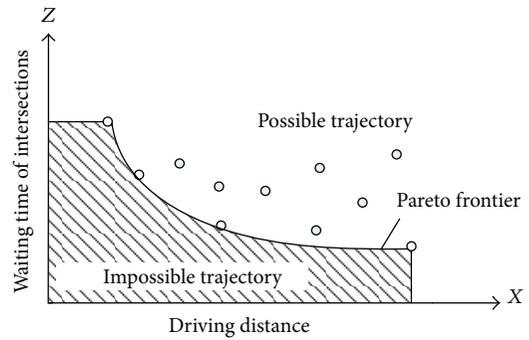


FIGURE 11: Vehicle possible trajectory in 2D plot of waiting time and driving distance.

of the floating car encountering the red-light delay time on the route choice, we ascertain the main factors of the vehicle trajectory judgement:

- (1) travel the shortest distance as far as possible;
- (2) as little as possible running through the intersections;
- (3) as much as possible to reduce the delay time of the red light.

But when there are multiple objectives, because there is a conflict between the objectives that cannot be compared, it is difficult to find a solution so that all objective functions are simultaneously optimal. We know the drivers often go more some distance in order to avoid crossing the intersection. The detour distance which different drivers can accept is not the same and is influenced by the driver's preference and the actual traffic status. So we need to find the optimal using the principle of Pareto optimal solution.

4.2.2. Pareto Optimal Principles. Pareto optimality is an important concept in game theory. For multiobjective optimization problem, there is usually a solution set by which all these solutions in terms of the objective function are beyond comparison, called the Pareto optimal solution.

In our problem, in order to find floating car driving path of achieving the Pareto optimality, a three-dimensional coordinate space needs to be established as our solution space. The x -, y -, and z -axes, respectively, determine floating car traveling distance, the number of intersections, and the waiting time of intersections.

We can describe all possible paths solution to the coordinate space. The solution space can be decomposed into three two-dimensional solution planes which represent relationship of three objects by two and two shown in Figures 10, 11, and 12. Shaded area is practically impossible to reach solution domains, namely, floating car driving distance, the number of intersections, and waiting time of intersections cannot simultaneously achieve such an ideal situation. All trajectories solution falling on the continuous curve can be considered as a set of optimal solutions. The curve is called Pareto Frontier of the current solution space, while the set of nondominated solutions is the Pareto optimal solution.

- (1) From the large number of survey data, if we reached the destination more closely, the number of crossing

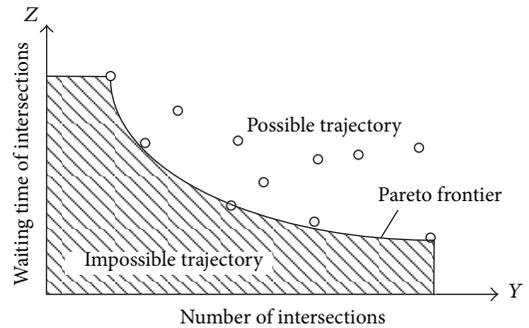


FIGURE 12: Vehicle possible trajectory in 2D plot of intersections number and waiting time.

the intersection may increase. On the contrary, the number would decrease. The number of intersections and driving distance of the floating car cannot simultaneously achieve the optimal in Figure 10.

- (2) Also we analyze the survey data and find that if we reached the destination more closely, the intersection delay time may increase. On the contrary, the delay time may decrease. The relationship between delay time and driving distance is likely as the number of intersection and driving distance. This is a typical matter of choosing the shortest distance or the shortest time in Figure 11.
- (3) As can be seen from the survey data, the number of intersection and waiting timing of crossing the intersection cannot simultaneously achieve optimal, but there is a Pareto Frontier in Figure 12.

4.3. Fuzzy Evaluation Analysis. Fuzzy comprehensive evaluation method is a comprehensive evaluation based on fuzzy mathematics method, and the method describes the fuzzy boundaries by fuzzy membership degree. The concept of fuzzy sets is proposed by the American Automatic Control Expert Chad (Zadeh) Professor [22] in 1965 to express the uncertainty of things.

In the paper, we utilize the fuzzy evaluation to analyze the driver's choice of travel trajectory, which can change the qualitative analysis into the quantitative analysis. Now we solve the problem of judging the vehicle trajectory when the

adjacent GPS matched point has a small network. According to the travel distance, delay time of intersection, and the number of crossing intersections, we build the fuzzy set respectively.

4.3.1. Travel Distance Membership Function. According to the survey data, we set the μ_d as the travel distance function and build the discourse domain. The variables are the ratio of each path distance and the trajectory tracking problem of Euclidean distance between starting point and end point. We also define the parameter μ_{d1} μ_{d2} μ_{d3} as the membership function {(travel distance short), (travel distance medium), (travel distance long)}. Membership functions are as follows:

$$\begin{aligned} \mu_{d1} &= \begin{cases} 1 & x \leq 1 \\ \frac{(1.3-x)}{0.3} & 1 < x \leq 1.3 \\ 0 & x > 1.3 \end{cases} \\ \mu_{d2} &= \begin{cases} 0 & x \leq 1 \\ \frac{(x-1)}{0.3} & 1 < x \leq 1.3 \\ \frac{(1.6-x)}{0.3} & 1.3 < x \leq 1.6 \\ 0 & x > 1.6 \end{cases} \\ \mu_{d3} &= \begin{cases} 0 & x \leq 1.3 \\ \frac{(x-1.3)}{0.3} & 1.3 < x \leq 1.6 \\ 1 & x > 1.6. \end{cases} \end{aligned} \quad (1)$$

4.3.2. The Number of Intersections Membership Function. We set the μ_n as the number of intersections and build the discourse domain. We also define the parameter μ_{n1} μ_{n2} μ_{n3} as the membership function {(intersection number less), (intersection medium), and (intersection number more)}. According to the driving experience of floating car driver the number of intersections is not more than 6. So the membership functions are built as follows:

$$\begin{aligned} \mu_{n1} &= \begin{cases} 1 & y \leq 2 \\ 2 - \frac{y}{2} & 2 < y \leq 4 \\ 0 & y > 4, \end{cases} \\ \mu_{n2} &= \begin{cases} 0 & y \leq 2 \\ \frac{y}{2} - 1 & 2 < y \leq 4 \\ 3 - \frac{y}{2} & 4 < y \leq 6 \\ 0 & y > 6, \end{cases} \\ \mu_{n3} &= \begin{cases} 0 & y \leq 4 \\ 2 - \frac{y}{2} & 4 < y \leq 6 \\ 1 & y > 6. \end{cases} \end{aligned} \quad (2)$$

4.3.3. Delay Time of Intersection Membership Function. According to the survey data, we set the μ_t as the delay time function and build the discourse domain. We also define the parameter μ_{t1} μ_{t2} μ_{t3} as the membership function {(delay time short), (delay time secondary), and (delay time long)}. According to the driving experience of floating car driver, delay time of intersection is not more than 3 minutes. So the membership functions are built as follows:

$$\begin{aligned} \mu_{t1} &= \begin{cases} 1 & z \leq 60 \\ 2 - \frac{z}{60} & 60 < z \leq 120 \\ 0 & z > 120, \end{cases} \\ \mu_{t2} &= \begin{cases} 0 & z \leq 60 \\ \frac{z}{60} - 1 & 60 < z \leq 120 \\ 3 - \frac{z}{60} & 120 < z \leq 180 \\ 0 & z > 180, \end{cases} \\ \mu_{t3} &= \begin{cases} 0 & z \leq 120 \\ 2 - \frac{z}{60} & 120 < z \leq 180 \\ 1 & z > 180. \end{cases} \end{aligned} \quad (3)$$

4.3.4. The Fuzzy Matrix Computation. After establishment of fuzzy membership function, we need to compute the fuzzy matrix to confirm the most possible vehicle trajectory. According to the fuzzy vector set as follows: $\mu_d = \{\mu_{d1} \mu_{d2} \mu_{d3}\}$, $\mu_n = \{\mu_{n1} \mu_{n2} \mu_{n3}\}$, $\mu_t = \{\mu_{t1} \mu_{t2} \mu_{t3}\}$, we combine the fuzzy vector together and get the fuzzy matrix $R = \{\mu_d^T, \mu_n^T, \mu_t^T\}^T$ considering the driver's different preferences on the road distance, delay time, and the number of intersections. So we build the weight vector $p = \{p_d, p_n, p_t\}$, where the relationship equation is $p_d + p_n + p_t = 1$. Then we need to build the fuzzy vector $Q = PR$, in which the possibility of the vehicle trajectory represented by the each element is large, medium, or small. Furthermore, In order to determine the actual path, we introduce the matching degree of vectors $\lambda = \{\lambda_1 \lambda_2 \lambda_3\}$ ($\lambda_1 + \lambda_2 + \lambda_3 = 1$), by $\alpha = Q\lambda^T$, calculated for each path of the evaluation factors α (the possibility of a true path).

5. Experimental Analysis

To verify this algorithm of vehicle trajectory estimation, we choose an actual road based on the Beijing traffic control system to acquire the delay time of intersection in real time. We know that the vehicle speed, road distance, and delay time of intersection are affected by the traffic control. So, it is necessary to consider the impact of intersection signal control when we estimate the vehicle trajectory. Although the intersection signal control information is often difficult to obtain, this paper can rely Beijing traffic control information platform in real time to obtain the appropriate traffic control

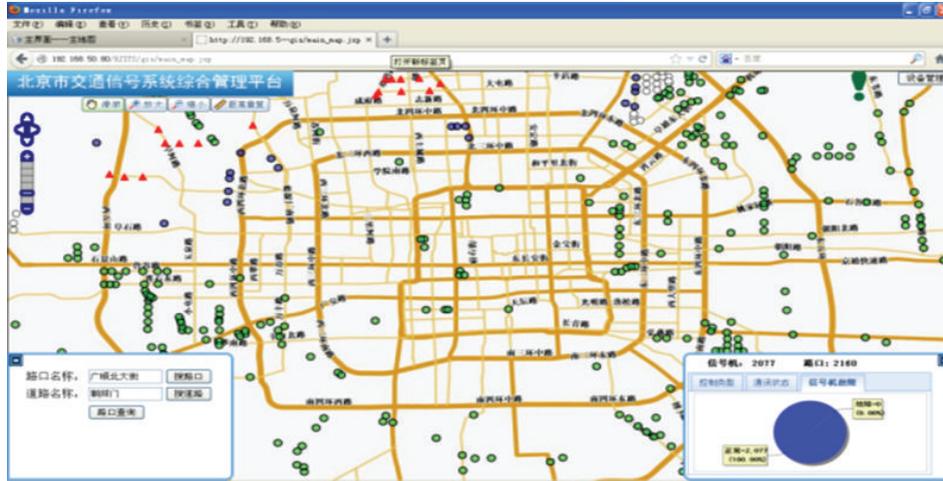


FIGURE 13: Beijing traffic control information platform.



FIGURE 14: An intersection in Shijingshan district.

information. Figures 13, 14, 15, and 16 are real time traffic timing of an intersection in Shijingshan district.

Here we determine the optimal path for an example of the urban road network in Wangjing area, Beijing city, shown in Figure 17, using the Pareto optimization and fuzzy comprehensive judgment method. The points P1 and P2 in the Figure are the starting point and end point of the undetermined path and there is a small road network which is connected by the intersections of A~I between P1 and P2. Various sections of the road network length are marked in Arabic numerals (in meters) and some intersections with traffic lights are the light controlled intersections in the figure. So we can describe all possible paths from the point P1 to P2 to a three-dimensional coordinate space be composed of traveling distance, the number of intersections, and the waiting time.

The length data of all the road section from Figure 17 is as shown in Table 1.

The light controlled intersections with traffic lights in the Figure 17 cause vehicle delay time when the vehicle passes through the intersection and have a great impact on selecting the running track for drivers. So obtaining signal control information of each intersection is very necessary. We just need to get the timing cycle data of each intersection to determine the influence of the control signal in order to improve the computational efficiency (see Table 2).

But because of different ways of the vehicle through the intersection the impact of signals control is also different. For example, the control of traffic lights is different for vehicles turning left, going straight, and turning right, and the delay caused by the intersection is also different. So when we consider the impact of lights on the path selection, we need to consider the travel direction of the path in the intersection. The average delay caused by the different traveling direction can be determined by the intersection timing cycle. If we let C as the cycle, we can set straight delay $t_s = C/4$, left delay

TABLE 1: The sections length for analytical experiment.

DL (m)	OL (m)								
	A	B	C	D	E	F	G	H	I
A	0	400	640	350	/	/	/	/	/
B	400	0	/	/	320	/	/	/	/
C	640	/	0	400	/	410	/	/	/
D	350	/	400	0	390	740	/	/	/
E	/	320	/	390	0	/	/	/	410
F	/	/	410	740	/	0	160	/	/
G	/	/	/	/	/	160	0	480	/
H	/	/	/	/	/	/	480	0	410
I	/	/	/	/	410	/	/	410	0

O: origin node, D: destination node, L: the sections length, /: nonexistent sections.



FIGURE 15: The traffic signal time program of the intersection.

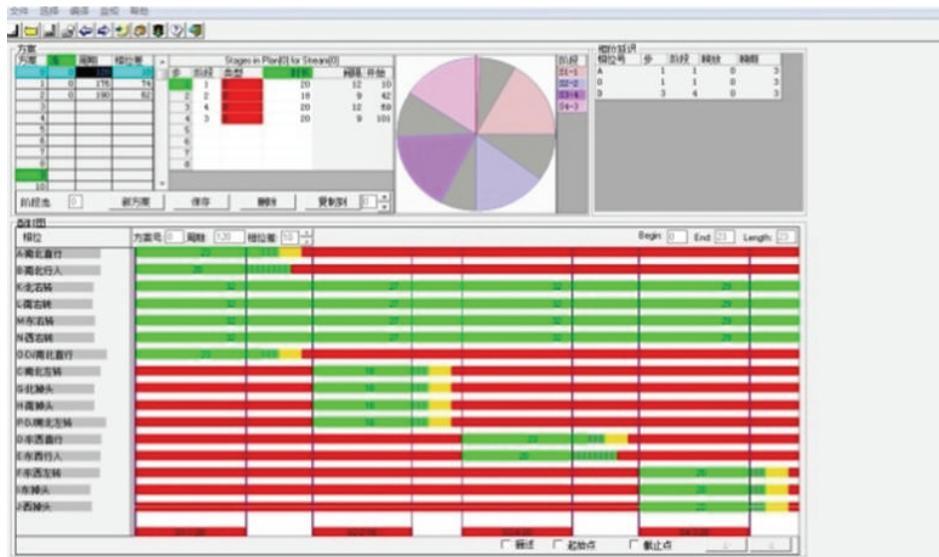


FIGURE 16: The timing plan of the intersection.

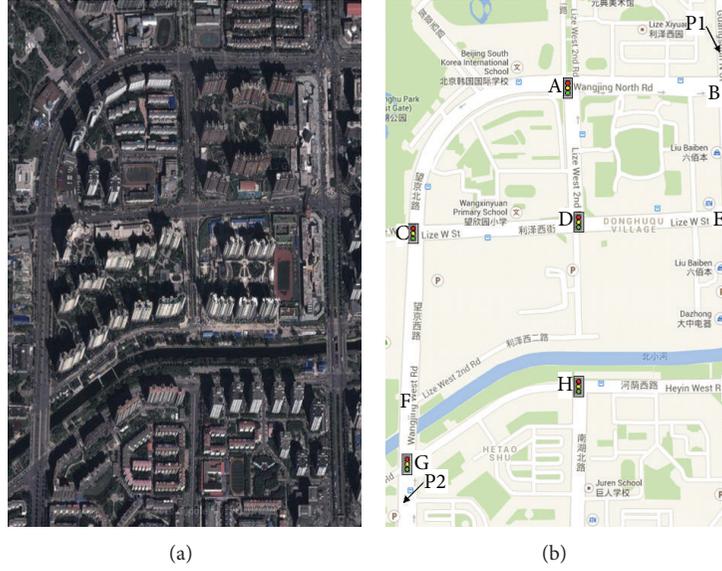


FIGURE 17: Wangjing area road network for analytical experiment.

TABLE 2: Signal timing cycle of each intersection.

		The light controlled intersections							
		A	B	C	D	E	G	H	I
cycle(s)		130	160	150	120	160	130	110	160

$t_l = 3C/8$, and right delay $t_r = 0$ (without considering the influence of signal control, only considering the impact of the intersections number).

The possible paths obtained by using the path search method are drawn in three-dimensional space composed of floating car traveling distance, the number of intersections, and the waiting time of intersections (Table 3).

Then we can obtain Pareto optimal solution using multi-objective decision theory and Pareto optimal principles, and this group of the optimal solution set forms the Pareto Frontier of the current optimization problem. The indeterminate trajectories in Pareto frontier are the numbers 1, 2, 4, and 6 that is B-A-C-F-G, B-A-D-F-G, B-E-D-F-G, and B-E-I-H-G. Then we can choose the best trajectory from the four paths on Fuzzy Comprehensive (Table 4).

Here we choose the similar weight of evaluating indicator for travel distance, intersection number, and delay time, so the weight vector $P = [0.3 \ 0.3 \ 0.4]$. Because the match degree of a path (the possibility of it as a true path) increases along with the decrease of traveling distance, the number of intersections, and the waiting time, we can construct the matching degree vector $\lambda = [0.6 \ 0.3 \ 0.1]$.

According to the judging indicators in Table 5, we can find that the first candidate trajectory's α_1 is the largest so we judge that the no. 1 path (B-A-C-F-G) is the most possible trajectory. We have done lots of the real experiment which the result showed that the driving car equipped GPS having the probability of 92 percentage selected the real driving



FIGURE 18: The most possible trajectory considering the traffic control.

trajectory comparison to the estimation trajectory by the algorithm introduced in the paper (Figure 18).

6. Conclusion

The efficient use of the floating cars data with the tracking of low-frequency sampling floating cars and traffic signal control is becoming a new hot topic. The overlong period of GPS sampling and overly complex road topology, which is the two features of current urban floating cars construction system, made it possible to use the floating car technology in transportation area. According to the map matching algorithm result, this paper developed a new algorithm based on

TABLE 3: The possible trajectory data.

Possible trajectory	Distance (meters)	Intersection numbers	Delay time (s)
1	B-A-C-F-G	5	$0 + C_A/4 + C_C/4 + 0 + C_G/4 = 102.5$
2	B-A-D-F-G	5	$0 + 3C_A/8 + C_D/4 + 0 + C_G/4 = 111.25$
3	B-A-D-C-F-G	6	$0 + 3C_A/8 + 0 + 3C_C/8 + 0 + C_G/4 = 137.5$
4	B-E-D-F-G	5	$C_B/4 + 0 + 3C_D/8 + 0 + C_G/4 = 117.5$
5	B-E-D-C-F-G	6	$C_B/4 + 0 + C_D/4 + 3C_C/8 + 0 + C_G/4 = 158.75$
6	B-E-I-H-G	5	$C_B/4 + C_E/4 + 0 + C_H/4 + 3C_G/8 = 156.25$

TABLE 4: Fuzzy matrix of membership for candidate paths.

Candidate trajectory	μ_d	μ_n	μ_t	Membership degree of fuzzy matrix
1	[0.2 0.8 0]	[0 0.5 0.5]	[0.29 0.71 0]	$\begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.5 & 0.5 \\ 0.29 & 0.71 & 0 \end{bmatrix}$
2	[0.1 0.9 0]	[0 0.5 0.5]	[0.15 0.85 0]	$\begin{bmatrix} 0.1 & 0.9 & 0 \\ 0 & 0.5 & 0.5 \\ 0.15 & 0.85 & 0 \end{bmatrix}$
4	[0.2 0.8 0]	[0 0.5 0.5]	[0.04 0.96 0]	$\begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.5 & 0.5 \\ 0.04 & 0.96 & 0 \end{bmatrix}$
6	[0.18 0.82 0]	[0 0.5 0.5]	[0 0.4 0.6]	$\begin{bmatrix} 0.18 & 0.82 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.4 & 0.6 \end{bmatrix}$

TABLE 5: Judging indicators of candidate paths.

Candidate trajectory	The fuzzy vector $Q = PR$	The evaluation factor $\alpha = Q\lambda^T$
1	[0.18 0.67 0.15]	0.324
2	[0.09 0.76 0.15]	0.297
4	[0.08 0.77 0.15]	0.294
6	[0.05 0.56 0.39]	0.237

multiobjective optimization and fuzzy theory by analyzing the characteristics of GPS data, which also integrated the trajectory estimating, driving distance, intersection time delay, and the crossroads number between the GPS matching points. The validation of the algorithm included the floating cars' trajectory estimating and the data of the signal control system. The result of the experiment achieved the author's prospective target and solved the questions mentioned before effectively. In present study, the author only conducts the experiment, estimating the floating cars' trajectory localized in urban. Therefore there are certainly some difficulties if tracking and estimating the vehicle moving no limitation in urban. In addition, because of the limited time and the length of the paper the path search algorithm and the Pareto frontier search algorithm of three-dimensional space are not efficient enough and are needed to be researched further.

However, the algorithm in the paper is reasonable and practical. The achievement of this paper is the further study of the application of floating cars' information which promote the traffic collection information technology effectively and

will lay a solid foundation to the distribution and forecasting of floating cars' travel time.

Conflict of Interests

The authors declare that they do not have any commercial or associative interests that represent a conflict of interests in connection with the work submitted.

Acknowledgments

This study was supported by the National High Technology Research and Development Program of China (863 Program) (no. 2012AA112401), Beijing Talents Qualification Project (no. 2012D005002000006), and the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institution (CIT&TCD201304002).

References

- [1] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research Part C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [2] W. Wang, X. Jiang, S. Xia, and Q. Cao, "Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation," *Safety Science*, vol. 48, no. 10, pp. 1248–1262, 2010.
- [3] W. Wang and F. Hou, "Simulation-based dependability evaluation of complex repairable system using variable time increment method," in *Proceedings of the 5th International Symposium*

- on *Intelligent Systems and Informatics (SISY '07)*, pp. 237–241, August 2007.
- [4] A. Tarko and N. Rouphail, “Travel time data fusion in ADVANCE,” in *Proceedings of the Pacific Rim TransTech Conference*, vol. 1, pp. 36–42, Seattle, Wash, USA, 1993.
- [5] R. P. Schäfer, K. U. Thiessenhusen, and P. Wagner, “A traffic information system by means of real-time floating-car data,” in *Proceedings of the ITS World Congress*, 2002.
- [6] T. Miwa, T. Sakai, and T. Morikawa, “Route identification and travel time prediction using probe-car data,” *International Journal of ITS Research*, vol. 2, no. 1, pp. 21–28, 2004.
- [7] Z. Liyu, W. Huimin, and S. Jianping, “Floating car based real-time-traffic-info collection system in Beijing,” *Urban Transport of China*, vol. 6, no. 1, pp. 77–80, 2008.
- [8] Y. Wang, *Research on the key technology of large-scale strategic traffic coordination & control system [Ph.D. thesis]*, Jilin University, 2009.
- [9] J. Zhou and Y. Li, “Hangzhou taxi operating characteristic analysis based on Floating Car Technology,” *Science & Technology Information*, vol. 3, pp. 120–122, 2011.
- [10] C. E. White, D. Bernstein, and A. L. Kornhauser, “Some map matching algorithms for personal navigation assistants,” *Transportation Research Part C*, vol. 8, no. 1–6, pp. 91–108, 2000.
- [11] G. Taylor, C. Brunson, J. Li, A. Olden, D. Steup, and M. Winter, “GPS accuracy estimation using map matching techniques: applied to vehicle positioning and odometer calibration,” *Computers, Environment and Urban Systems*, vol. 30, no. 6, pp. 757–772, 2006.
- [12] R. R. Joshi, “A new approach to map matching for in-vehicle navigation systems: the rotational variation metric,” in *Proceedings of the IEEE Intelligent Transportation Systems Proceedings*, pp. 33–38, August 2001.
- [13] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, “A general map matching algorithm for transport telematics application,” *Journal of Navigation*, vol. 13, no. 2, pp. 116–119, 2005.
- [14] Z. Wang, Z. Li, and L. Wang, “The application of GPS data processing technique in map matching,” in *Proceedings of the International Conference on Transportation, Mechanical, and Electrical Engineering*, pp. 588–591, 2011.
- [15] Z. Wang, L. Wang, and J. Wang, “Delay map matching algorithm of mass GPS data based on topological judgment,” *Journal of Southwest JiaoTong University*, vol. 47, no. 5, pp. 861–866, 2012.
- [16] H. Takada, *Road traffic condition acquisition via mobile phone location referencing [Ph.D. thesis]*, University of Waterloo, 2006.
- [17] Y. Chen, *Traffic flow analysis of urban road network [Ph.D. thesis]*, ShangHai JiaoTong University, 2008.
- [18] B. S. Kerner, C. Demir, R. G. Herrtwich et al., “Traffic state detection with floating car data in road networks,” in *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems*, pp. 44–49, September 2005.
- [19] Y.-J. Byon, A. Shalaby, and B. Abdulhai, “Travel time collection and traffic monitoring via GPS technologies,” in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC '06)*, pp. 677–682, September 2006.
- [20] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior (commemorative Edition)*, Princeton University Press, 2007.
- [21] T. C. Koopmans, “Analysis of production as an efficient combination of activities,” *Activity Analysis of Production and Allocation*, vol. 13, pp. 33–37, 1951.
- [22] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

Research Article

A Study on the Model of Traffic Flow and Vehicle Exhaust Emission

Han Xue,¹ Shan Jiang,² and Bin Liang³

¹ Institute of Development Studies, Shougang Corporation, Beijing 100082, China

² School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

³ Beijing Automotive Research Institute Company Limited, Beijing 100079, China

Correspondence should be addressed to Shan Jiang; jiangshanyou@163.com

Received 9 September 2013; Revised 16 November 2013; Accepted 2 December 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Han Xue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The increase of traffic flow in cities causes traffic congestion and accidents as well as air pollution. Traffic problems have attracted the interest of many researchers from the perspective of theory and engineering. In order to provide a simple and practical method for measuring the exhaust emission and assessing the effect of pollution control, a model is based on the relationship between traffic flow and vehicle exhaust emission under a certain level of road capacity constraints. In the proposed model, the hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxides (NO_x) are considered as the indexes of total exhaust emission, and the speed is used as an intermediate variable. To verify the rationality and practicality of the model, a case study for Beijing, China, is provided in which the effects of taxi fare regulation and the specific vehicle emission reduction policy are analyzed.

1. Introduction

The negative impact of urban road traffic is mainly on air quality [1], ecosystem, and noise level [2]. Due to the continuing increase of motor vehicles, human health [3] and environment have been severely impacted. According to the classification of air pollutant sources in urban area, motor vehicle emission accounts for more than 80% of the air pollution in major cities [4]. The statistics of Beijing show that the level of carbon monoxide (CO) and nitrogen oxides (NO_x) exceeds national standard even in the city's fourth and fifth ring roads where the average speed of vehicles is high.

Since the concept of sustainable development has been adopted into the theory and methods of urban transport systems planning, the coordination between transportation development and urban environment becomes the focus of the urban transportation research in the 21st century. In recent years, many scholars have studied vehicle exhaust emission for environment protection [5].

In order to reduce air pollution caused by motor vehicles, several control measures have been imposed in the last decade including upgradation of gasoline quality, strict environmental standards, and the promotion of new energy

vehicles. Schifter et al. studied the trends of exhaust emissions from gasoline motor vehicles in the metropolitan area of Mexico City [6]. They presented results on brand new vehicles which indicated that NO_x emission factors, though they were within the Tier I standard, deteriorated rapidly with the travel distance. Alkurdi et al. determined the profile and concentration of PAH in exhaust emissions of light- and heavy-duty vehicles running on the roads of Damascus city [7].

Ropkins et al. provided a comprehensive critical review of the techniques utilized to monitor real-world vehicle exhaust emissions [8]. Real-world measurements (measurements of exhaust emissions from vehicles in operation on the highway network) differ from laboratory-based measurements (typically using test cycles) because they have a more realistic potential to capture the range of variability typically encountered in real-world driving, including variability in driver behaviour, interactions with other road users, and interactions with highway infrastructure, all of which have the potential to influence exhaust emissions. May et al. used four independent yet complementary approaches to investigate POA gas-particle partitioning [9]: sampling artifact

correction of quartz filter data, dilution from the constant volume sampler into a portable environmental chamber, heating in a thermodenuder, and thermal desorption/gas chromatography/mass spectrometry analysis of quartz filter samples. This combination of techniques allowed gas-particle partitioning measurements to be made across a wide range of atmospherically relevant conditions. Mikulcic et al. analyzed the influence of different amounts of fuel, mass flow of the tertiary air, and the adiabatic wall condition on the decomposition rate of limestone particles, burnout rate of coal particles, and pollutant emissions of a newly designed cement calciner. Numerical models of calcination process and pulverized coal combustion were developed and implemented into a commercial computational fluid dynamics code, which was then used for the analysis [10].

These studies covered the measurement of vehicle exhaust emission factors [11, 12], the analysis or model of pollutant emission on traffic corridors [13], and so forth. However, the required variables and detection methods were complicated, and the conclusions were significant only to the sample. Due to the lack of rapid and effective macrodetection and calculation methods for overall pollution throughout the city roads, these studies cannot provide effective support to traffic regulation departments to develop policies or take interim control measures.

This paper puts forward a vehicle exhaust emission model based on the speed which is taken as an intermediate variable and the exhaust emission of hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxides (NO_x) as indexes of exhaust emission. This model provides a simple and easy method to measure and effectively control the exhaust emission.

2. An Analysis of Motor Vehicle Exhaust Emissions

2.1. Factors of Vehicle Exhaust Emissions. Motor vehicle exhaust emissions are known to contain carbon monoxide (CO), nitrogen oxides (NO_x), hydrocarbons (HC), suspended particulate matter, and a small amount of sulfur dioxide (SO_2). Since carbon monoxide (CO), nitrogen oxides (NO_x), and hydrocarbons (HC) are tested as pollutants by two Chinese national standards, namely, Emission Standard for Exhaust Pollutants from Light-duty Vehicle (GB 14761.1-93) and Emission Standard for Exhaust Pollutants from Gasoline Engine of Vehicle (GB 14761.2-93); in this paper carbon monoxide (CO), nitrogen oxides (NO_x), and hydrocarbons (HC) are selected as assessment indexes of road traffic exhaust emissions.

The emission factor is a key parameter in the calculation of vehicle pollutant emission, which is also called emission of unit mass, and it is the average emission under the influence of various factors according to the Chinese national standard, namely, Automotive Emission-Terms and Definitions (GB5181-2001). Vehicle exhaust emission is affected by many factors, including the features of vehicles (such as vehicle type, technical level, emission control devices, and operation condition), urban road conditions, maintenance frequency,

fuel type, the levels and effect of maintenance, and the characteristics of roads (altitude, temperature and humidity, road conditions, and traffic conditions). The motor vehicle operation conditions are different between urban roads and highways. The former conform to the standard conditions while the latter conform to the constant speed condition. The differences between two emission factors for different operation conditions are relatively large, so it is necessary to consider the average speed of vehicles for the calculation of vehicle emission factors.

2.2. The Selection of Testing Routes and Testing Method. The selection of testing routes was a result of consideration of the features of Beijing traffic network. Representative types of roads were chosen for testing, including the Second Ring Road, the Third Ring Road, the Fourth Ring Road, the Fifth Ring Road, the Changan Avenue, and the Jingtong Expressway. Testing time covered morning and evening rush hours as well as nonrush hour daytime period. Motorists' driving habits were also considered.

The main device for vehicle exhaust emission testing was the SEMTECHDS gas analysis system made by the Sensor Company, USA. Five automotive brands were selected as testing vehicles due to their higher market share, including Volkswagen, Honda, GM, Ford, and Nissan. GPS speedometer and slope meter auxiliary equipment were used for continuous record of each vehicle's speed and road conditions. Data were collected for ten hours of steady driving without accidents for each vehicle under testing. Vehicle's exhaust emissions were measured under the operation speed: 0 km/h ~ 90 km/h. The impact caused by the installation position of testing device on vehicles was eliminated.

3. Speed and Pollutant Emission Model

Researchers such as Richard Anthony Margiotta believed that emission factors were sensitive to the vehicle speed and their relation function could be expressed as $\text{EF}_{jk} = g(u)$ [14].

3.1. The Relationship between the Pollutant and Vehicle Speed. In order to quantify the relationship between the vehicle speed and the vehicle pollutant exhaust, the impacts of slope, sudden acceleration, and deceleration were eliminated first. That is, the data of vehicles operated at constant speeds on smooth roads were studied first. The method of data collection generated noise fluctuations in data. Therefore, high frequency of noise components was removed through curve-fitting techniques [15]. By comparing the continuous data recorded by kilometers with the corresponding time of SEMTECHDS gaseous analysis system, the exhaust emissions at each time and the corresponding data can be obtained. Figures 1, 2, and 3 reveal the relation curve between the vehicle speeds and the emission factors of CH, CO, and NO_x , respectively.

From Figure 1 to Figure 3, we can see that the trend of comprehensive emission factors of HC, CO, and NO_x is changing with the mean vehicle speed. There are several points worth noticing.

TABLE 1: Interval division of acceleration.

Slowdown		Constant speed		speedup	
Interval index	Interval range (m/s ²)	Interval index	Interval range (m/s ²)	Interval index	Interval range (m/s ²)
N_1	$-0.3 < a \leq -0.1$	0	$-0.1 < a \leq 0.1$	P_1	$0.1 < a \leq 0.3$
N_2	$-0.6 < a \leq -0.3$			P_2	$0.3 < a \leq 0.6$
N_3	$-1.0 < a \leq -0.6$			P_3	$0.6 < a \leq 1.0$
N_4	$a \leq -1.0$			P_4	$a > 1.0$

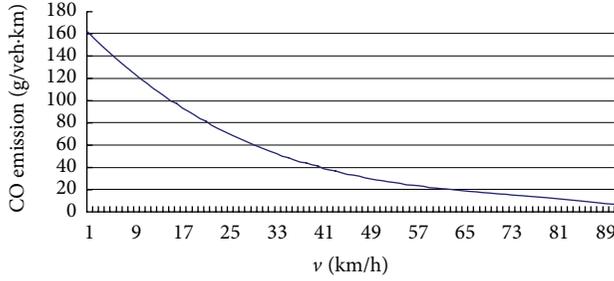


FIGURE 1: The change curve of CO comprehensive emission factor.

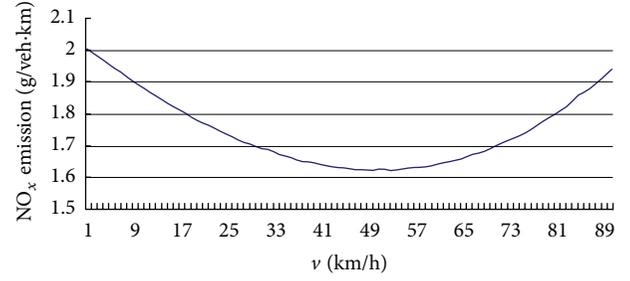
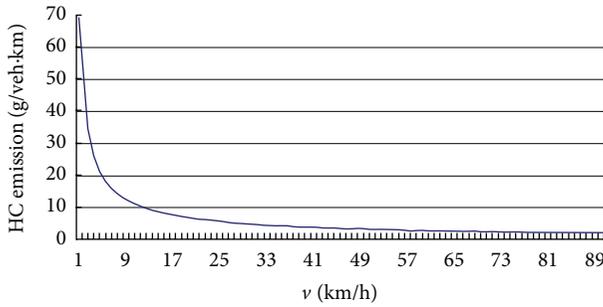

 FIGURE 3: The change curve of NO_x comprehensive emission factor.


FIGURE 2: The change curve of CH comprehensive emission factor.

Firstly, along with the increase of mean vehicle speed, vehicle emission of CH comprehensive factor decreases. Secondly, along with the increase of mean vehicle speed, vehicle emission of CO comprehensive factor gradually decreases. Thirdly, along with the increase of the mean vehicle speed, comprehensive NO_x emission factor decreases at first and then increases gradually, with 50 km/h and 60 km/h being the minimum value.

3.2. The Relationship between the Acceleration and Exhaust Emissions. As shown in Table 1, the vehicle running state could be divided into eight kinds according to the acceleration frequency under all kinds of driving conditions.

In addition, the authors selected speed ranges in the condition of low speed, medium speed, and high speed to analyze the relationship between the acceleration and exhaust emissions in order to eliminate the impact of the speed. The division of speed range is as follows: low speed was defined as within the interval (10, 20] h/km, medium speed as within the interval (40, 50] h/km, and high speed as within the interval

(60, 70] h/km. In each speed range, the authors analyzed the relationship between the acceleration and exhaust emissions.

The data show that HC, CO, NO_x, and CO₂ emission factors increase quickly with the increasing acceleration, especially in the range of high speed. When $a < -0.6$ m/s² (a denotes acceleration), the amplitude of emission factors increases slowly with increasing acceleration; when $-0.6 < a \leq 0.3$ m/s², the amplitude of emission factor increases rapidly with the increasing acceleration; when $0.3 < a \leq 1.0$ m/s², the amplitude of emission factor decreases rapidly with the increasing acceleration; when $a > 1$ m/s², the amplitude of emission factor increases with the increasing acceleration.

3.3. Fitting and Correction of the Speed and Exhaust Emission Function. In addition to the speed and acceleration, the vehicle exhaust emission is also closely related to the slope of road. Therefore, vehicle specific power (VSP) was introduced to fit exhaust emission curve.

VSP was put forward by Palacios and Luis and applied to analyze remote sensing data [16]. VSP is one of the parameters most close to the actual conditions, and it has been one of the core parameters of the next generation mobile emission model. P_{VPS} denotes the ratio of the motor vehicle output power and its quality (in kW/t). P_{VPS} combines parameters such as speed, acceleration, slope, and wind resistance, so it can greatly improve the accuracy of the fitting.

V denotes speed, a denotes acceleration, and θ denotes a road gradient expressed in radians, and the following functions are obtained:

$$E_{NO_x} = 2.0164 + [1.1a + 9.81(a \tan(\sin \theta)) - 0.0142] V + 0.0001V^2 + 0.00000053V^3,$$

$$\begin{aligned}
E_{CO} &= 167.154 + [1.1a + 9.81(a \tan(\sin \theta)) - 5.2911] V \\
&\quad + 0.0662V^2 + 0.0003V^3, \\
E_{HC} &= 68.7252V^{1.1a+9.81(a \tan(\sin \theta))-0.7760}.
\end{aligned}
\tag{1}$$

Thus, the basic functional relationship between exhaust emission and speed has been obtained. In order to more accurately reflect the actual urban road conditions of Beijing, the authors calculated the mean value of VSP for many types of motor vehicles. If the accurate VSP of a particular type of vehicle is to be obtained, more samples are needed to correct the specific parameters.

4. Speed and Traffic Flow Model

4.1. Traffic Features Analysis. Urban roads are the main infrastructure of urban transportation. Urban roads should not only satisfy the requirement of the traffic but also meet the requirement of urban land [17]. Features of urban road traffic are related to urban road grade and urban area. In general, urban road traffic at different levels has different features. The composition of urban road traffic is complicated. Generally, urban road traffic covers all sorts of motor vehicles, and the ratio of motor vehicles is related to the road grades and urban area. The traffic flow is normally distributed to time. If the influence of road intersections is not considered, vehicle's speed on main lines is mainly related to the traffic flow. In this paper, the research was conducted in Beijing area. Since the main lines in Beijing are the city's ring roads and there are no traffic lights on these ring roads, consideration of the influence of the intersections is not needed.

4.2. The Relationship between Speed and Traffic Flow. Based on the Greenshields model, the function of the relationship between speed and flow can be obtained as follows:

$$\bar{U}_i = N_i K_j \left(v - \frac{v^2}{\bar{v}} \right). \tag{2}$$

Here \bar{U}_i denotes the standard traffic flow through road i per unit time, pcu/h; N_i denotes the number of lanes on road i ; K_j denotes the traffic density of road i , pcu/km; v denotes the average speed on road i , km/h; \bar{v} denotes the free flow speed of road i , km/h.

Therefore, when the traffic flow is in a steady state, the function of the relationship between speed and traffic flow can be obtained from the above function as follows:

$$v = f(U_i) = \bar{v} \left(\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\bar{U}_i}{N_i \bar{v} K_j}} \right). \tag{3}$$

4.3. Discussion. As for transportation system which has reached a balance between supply and demand [18], traffic density and the relationship between vehicle speed and traffic flow will change while adding new traffic flow into traffic system (Figure 4). In AB area, as vehicle traffic flow is smooth,

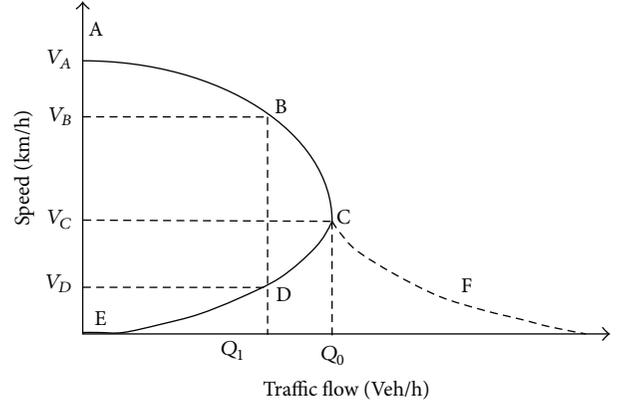


FIGURE 4: The relationship between speed and traffic flow.

vehicle's speed is always of highest level. With new vehicles entering into the transportation system, the density of traffic flow increases and enters into BC area. Gradually, traffic density increases with the spaces among vehicles reduced; Interference among vehicles increases and as a result vehicles' speed declines. As new vehicles continue to enter into the transportation system, the volume of vehicles increases and so does the traffic density until the road eventually becomes saturated. When $q_c = Q$, the transportation system is in the state of saturation [19, 20]. At this time, if timely effective measures are not taken to restrict new cars from entering into the system, the speed and traffic flow will fall further, and road traffic efficiency continues to decrease. Eventually, serious traffic congestion will happen, transportation system will be paralyzed, and traffic flow will drop to zero.

From the above analysis, it can be seen that, for a given traffic flow Q_1 , ($Q_1 < Q_0$), there will be two corresponding speeds V_B and V_D . Thus, in the following case of Beijing traffic where $Q = Q_1$ ($Q_1 < Q_0$) is given, it can be concluded that $V = V_D$, as Beijing's road is very congested.

Formula (3) yields that

$$v = f(U_i) = \bar{v} \left(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\bar{U}_i}{N_i \bar{v} K_j}} \right). \tag{4}$$

5. The Model of Traffic Flow and Vehicle Exhaust Emission

If $E_k = E_k(v)$, thus $E_k = E_k(f(U_i))$ from $v = f(U_i)$. Therefore, the model of traffic flow and vehicle exhaust emission is obtained.

In this model, carbon monoxide (CO), nitrogen oxides (NO_x), and hydrocarbons (CH) are selected as three final indexes for vehicle exhaust emission evaluation. Based on the relationship between vehicle speed and vehicle exhaust emission factors, a mathematical model of total emission and road speed is established. And based on the relationship between speed and traffic flow under a certain level of road capacity, the mathematical model of speed and traffic flow is established. Then, with speed as an intermediate variable a model is established to describe the relationship between

traffic flow and vehicle exhaust emissions under certain level of urban road capacity constraints. Although the model is based on a single road, as long as the speed of vehicles remains the same on the section road network or the entire road network, the model still applies. Thus, this model provides a theoretical basis to measure the exhaust emission within a certain region and effectively to control them.

6. Case Study

Based on the proposed model, with the given road capacity and given traffic volume (through a standard equivalent conversion), road vehicle exhaust emissions (CO, HC, and NO_x) can be analyzed for urban transportation planning and environmental impact assessment.

6.1. Background. Statistics of Beijing Municipal Bureau show that, by the end of 2012, there are 5.2 million motor vehicles in Beijing, while taxis in Beijing are about 66,000, accounting for a very small proportion. However, taxis in Beijing are exempt from the city's vehicle restriction based on the last digit on the license plate and they spend much longer hours on the road than other vehicles every day; their high level of utilization on the road and exhaust emissions have been severely criticized by people. Therefore, based on this model taxi exhaust emissions in Beijing are analyzed.

6.2. The Equivalent Conversion for Taxis and Private Cars. According to the data of the Beijing Transportation Research Center, at present, the average kilometrage of private cars is 15,000 km per year while a taxi's kilometrage reached 90,000 kilometers, which indicates that a taxi is equivalent to six standard private cars based on kilometrage. In addition, as 40% kilometrage of private cars is spent for weekend and holiday vacation travels and is spent out of the city's Sixth Ring Road, that is, out of the urban area of Beijing, the ratio of road utilization is 1:10 between taxis and private cars within the Sixth Ring Road area, that is, the urban area of Beijing.

According to "the hearing of dynamic adjustment mechanism on Beijing's taxi rents adjustment and the fuel surcharge improvement" the empty-loaded rate of taxis in Beijing was about 31% in 2012, which means for an average taxi in Beijing in 31% of the time when it was running on the streets there was no customer. Because of this, the efficiency of taxis is much lower than that of private cars. Taxi meter record data and the cost survey from Beijing Municipal Development and Reform Commission also show that taxi service efficiency is only about 70% of that of private motor vehicles.

Combining the input and output together, the authors concluded that the ratio of urban road utilization between taxis and private cars within the Sixth Ring Road is about 1:14. Therefore, the actual pollution effect of the 66,000 taxis in Beijing is equivalent to that of 924,000 standard private cars.

The data from Beijing Bureau of Statistics showed that there were 5.2 million vehicles in Beijing by the end of 2012, and the private cars were 4.16 million. Due to the vehicle restriction based on the last digit on the license plate in

Beijing, only 3.32 million vehicles are allowed to travel per day. In addition, 80% of these 3.32 million vehicles are mostly running within the Sixth Ring Road while almost all the taxis are running in the same area. Thus, the actual ratio of road utilization between taxis and private vehicles is 92.4:265, namely, 1:2.9; that is, among every four cars running in the streets in Beijing, at least one is a taxi. According to the model proposed in this paper, it can be deduced that a quarter of vehicle exhaust emissions are discharged by taxis. Given that private vehicles are mostly used for commuting to work, the proportion of the exhaust emissions of taxi is much higher during noncommuting hours.

6.3. The Taxi Planning. According to statistics, the monthly income of an average taxi in Beijing is about 16,300 Yuan, and the kilometrage is 7480 km. Taking the standard of New York City as reference, the daily kilometrage of taxi is 113 km; the present kilometrage of taxi in Beijing could be reduced by 55%. At the same time, in order to ensure the income of taxi drivers, the flag-fall price should increase by 100%, reaching 20 Yuan. According to the price elasticity of taxi demand in Beijing, the number of people using taxi will be reduced by 30%. Meanwhile, due to the wide application of the taxi reservation program, the empty-loaded rate of taxi can be controlled below 10%. Therefore, the standard equivalent quantity ratio of the taxis and private vehicles will be increased from 1:14 to 1:5. At this point, the 66,000 taxis are equivalent to 330,000 average private vehicles. The number of average vehicles within the Sixth Ring Road area is reduced to 2.6 million, which is a decrease of 27%.

6.4. The Planning Effect. According to the model, if the volume of vehicles is reduced by 27%, the speed of vehicles in the whole road network can be improved by 24%, while the average commuting time can be reduced by 19.4%. This conclusion is in accordance with the statistical data which show that; by a decrease of 12% in vehicle volume, the speed of vehicles in the whole road network can be increased by 10% in Beijing.

Furthermore, according to calculations, the exhaust emissions will be dropped as follows:

E_{NO_x} emissions will be reduced by 29.3%;

E_{CO} emissions will be reduced by 42.2%;

E_{HC} emissions will be reduced by 40.9%.

7. Conclusions

The objective of this paper is to find the relationship between vehicle exhaust emission and traffic flow under a certain level of urban road capacity constraints. The authors used speed as a bridge to describe the relationship between the two variables and established a model accordingly. The case analysis presented in the paper is to simulate the situation of taxi fare regulation in Beijing, of which the authors deduce the knock-on effects and calculate the city's vehicle exhaust emission reductions. In short, it is obvious that the impact of traffic flow on vehicle exhaust emission can be measured in

China's congested cities. Therefore the model is an effective way to control the volume of vehicles on urban roads and improve the speed and efficiency of vehicles to reduce the excessive pollution caused by vehicle emission in China.

Limitations and improvements of this model of traffic flow and vehicle exhaust emission include the following. (i) If more accurate results are required, a larger sample is required to calculate the parameters of different models to simulate actual road conditions, because emission performance of different vehicles is different (the types and ages of different vehicles determine different parameters), and (ii) in Beijing, the political center of China, traffic control based on political reasons is taken frequently and causes traffic speeds to reduce drastically or even to zero. Due to this considerable impact, the vehicle speed is too low to fit the curve, as the emissions of vehicles increased sharply. Therefore, in order to make the results more accurate, the stagnant traffic flow (speed of 10 km/h or less) should be taken into account into the model, and (iii) exhaust gas contains 150 to 200 different compounds. In addition to carbon monoxide, hydrocarbons, and nitrogen oxide compounds, particulate matter (PM) measured as either PM10 or PM2.5 (i.e., PM less than 10 μm or 2.5 μm in diameter, resp.) is also quite harmful to human body. PM is a mixture of liquid and solid particles of different sizes and chemicals that varies in composition both spatially and temporally. Epidemiological studies spanning five continents have demonstrated an association between mortality and morbidity and daily, multiday, or long-term (a period of more than a year) exposure to concentrations of pollutants, including PM. The estimated mortality impacts are likely to occur predominantly among elderly people with preexisting cardiovascular and respiratory disease and among infants. It is necessary to put PM emission into model detection range, thereby increasing the credibility of the model in the future.

References

- [1] P. Kumar, M. Ketzler, S. Vardoulakis, L. Pirjola, and R. Britter, "Dynamics and dispersion modelling of nanoparticles from road traffic in the urban atmospheric environment—a review," *Journal of Aerosol Science*, vol. 42, no. 9, pp. 580–603, 2011.
- [2] D. I. Popescu, R. E. Tuns, and I. F. Moholea, "The urban acoustic environment—a survey for road traffic noise," *Carpathian Journal of Earth and Environmental Sciences*, vol. 6, no. 1, pp. 285–292, 2011.
- [3] M. L. Burr, G. Karani, B. Davies, B. A. Holmes, and K. L. Williams, "Effects on respiratory health of a reduction in air pollution from vehicle exhaust emissions," *Occupational and Environmental Medicine*, vol. 61, no. 3, pp. 212–218, 2004.
- [4] G. A. Rhys-Tyler, W. Legassick, and M. C. Bell, "The significance of vehicle emissions standards for levels of exhaust pollution from light vehicles in an urban area," *Atmospheric Environment*, vol. 45, no. 19, pp. 3286–3293, 2011.
- [5] K. Yi, L. Z. Wang, B. X. Qi, L. M. Zhang, and S. Lei, "Investigation and analysis of motor vehicle exhaust emission in cities," *Energy Conversion and Application*, vol. 1-2, pp. 250–253, 2001.
- [6] I. Schifter, L. Diaz, S. Avalos et al., "Trends of exhaust emissions from gasoline motor vehicles in the metropolitan area of Mexico city," *International Journal of Environment and Pollution*, vol. 21, no. 2, pp. 166–174, 2004.
- [7] F. Alkurdi, F. Karabet, and M. Dimashki, "Characterization, concentrations and emission rates of polycyclic aromatic hydrocarbons in the exhaust emissions from in-service vehicles in Damascus," *Atmospheric Research*, vol. 120-121, pp. 68–77, 2013.
- [8] K. Ropkins, J. Beebe, H. Li et al., "Real-world vehicle exhaust emissions monitoring: review and critical discussion," *Critical Reviews in Environmental Science and Technology*, vol. 39, no. 2, pp. 79–152, 2009.
- [9] A. A. May, A. A. Presto, C. J. Hennigan, N. T. Nguyen, and T. D. Gordon, "Gas-particle partitioning of primary organic aerosol emissions: (1) gasoline vehicle exhaust," *Atmospheric Environment*, vol. 77, pp. 128–139, 2013.
- [10] H. Mikulcic, E. von Berg, M. Vujanovic, P. Priesching, and R. Tatschl, "Numerical analysis of cement calciner fuel efficiency and pollutant emissions," *Clean Technologies and Environmental Policy*, vol. 15, no. 3, pp. 489–499, 2013.
- [11] C. R. Fulper, S. Kishan, R. W. Baldauf et al., "Methods of characterizing the distribution of exhaust emissions from light-duty, gasoline-powered motor vehicles in the US fleet," *Journal of the Air and Waste Management Association*, vol. 60, no. 11, pp. 1376–1387, 2010.
- [12] M. M. Baum, J. A. Moss, S. H. Pastel, and G. A. Poskrebyshev, "Hydrogen cyanide exhaust emissions from in-use motor vehicles," *Environmental Science and Technology*, vol. 41, no. 3, pp. 857–862, 2007.
- [13] X. H. Tu, F. Z. Dong, Q. Feng, Y. J. Zhang, and W. Q. Liu, "On-road remote sensing of CO and CO₂ of motor vehicle exhaust emissions in Beijing using a TDLAS system," in *Optical Technologies for Atmospheric, Ocean, and Environmental Studies*, vol. 5832 of *Proceedings of SPIE*, pp. 350–358, SPIE, Beijing, China, August 2005.
- [14] R. A. Margiotta, *Improved vehicle speed estimation procedures for air quality and planning application [Ph.D. thesis]*, University of Tennessee, Knoxville, Tenn, USA, 1996.
- [15] W.-D. Ueng, J.-Y. Lai, and Y.-C. Tsai, "Unconstrained and constrained curve fitting for reverse engineering," *International Journal of Advanced Manufacturing Technology*, vol. 33, no. 11-12, pp. 1189–1203, 2007.
- [16] J. Palacios and J. Luis, *Understanding and quantifying motor vehicle emissions with vehicle specific power and Tildas remote sensing [Ph.D. thesis]*, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass, USA, 1999.
- [17] Y. X. Niu, M. Niu, and Y. Li, "Application of man-machine-environment system engineering in urban road traffic safety in China," in *Proceedings of the 9th Conference on Man-Machine-Environment System Engineering (MMESE '09)*, pp. 282–285, Dandong, China, August 2009.
- [18] F. Siebel, W. Mauser, S. Moutari, and M. Rasclé, "Balanced vehicular traffic at a bottleneck," *Mathematical and Computer Modelling*, vol. 49, no. 3-4, pp. 689–702, 2009.
- [19] M. Smith and R. Mailler, "The effect of congestion frequency and saturation on coordinated traffic routing," in *Agents in Principle, Agents in Practice*, vol. 7047 of *Lecture Notes in Artificial Intelligence*, pp. 187–201, 2011.
- [20] V. T. Arasan and P. Vedagiri, "Estimation of saturation flow of heterogeneous traffic using computer simulation," in *Proceedings of the 20th European Conference on Modelling and Simulation: Modelling Methodologies and Simulation Key Technologies in Academia and Industry, (ECMS '06)*, pp. 393–398, Bonn, Germany, May 2006.

Research Article

An Approach to Optimize the Departure Times of Transit Vehicles with Strict Capacity Constraints

Huimin Niu and Xiaopeng Tian

School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

Correspondence should be addressed to Huimin Niu; hmniumail.lzjtu.cn

Received 21 October 2013; Revised 16 November 2013; Accepted 26 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 H. Niu and X. Tian. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper focuses on an urban transit line which connects several residential areas and a workplace during the morning rush hours. The congestion is represented by some passengers who must wait for an extended duration and board the next or the third departure vehicles. This paper divides the time horizon equally into several small periods to measure the dynamic passenger demands. Under period-dependent demand conditions, a biobjective optimization model is developed to determine the departure times of transit vehicles at the start station with strict capacity constraints, in which a heuristic algorithm based on intelligent search and local improvement is designed to solve the model. The developed model can address the case in which more than two passengers arrive at a station simultaneously during one same period and calculate the number of boarded passengers. Finally, the model and algorithm have been successfully verified by a numerical example.

1. Introduction

Public transit vehicles shuttle on the lanes in accordance with a predetermined schedule to deliver the passengers from origins to destinations. The schedule for urban public transit systems is of predominant bridge between the service provider and the passenger users. At the same time, transit schedules and their compliance mirror the quality of public transit service provided. How to construct a practicable and effective schedule is a complicated task which requires the consideration of characteristics of passenger demands and benefit of the transit company as well as the operational regulations.

As a matter of fact, the key problem of designing a schedule for an urban public transit line is to determine the departure times of transit vehicles at the start station. This paper considers the scheduling problem for an urban transit line which connects several residential areas and a workplace during the morning rush hours. The spatial and temporal imbalances of demand distributions represent the transit congestion that occurs only at some stops or during some times. Under such conditions, some waiting passengers cannot board the current transit vehicle due to the limited capacity. A constant-headway schedule cannot

satisfy effectively the time-dependent passenger demands. An irregular schedule that allows uneven headways can accommodate the dynamic demand patterns and accelerate the fleet utilization that is provided by a limited number of vehicles. This schedule type reduces the passenger waiting times and economizes the operation costs as well.

The transit scheduling problem has received much attention in the past years [1–3]. LeBlanc introduced a model for determining frequencies using a modal-split assignment programming model with distinct transit routes to capture the effects of increases or decreases on individual transit line frequencies [4]. Lee et al. revealed that the total fleet size required for the operation of a bus network with mixed-size vehicles is smaller than the fleet size necessary when all fleet vehicles have the same regular size [5]. Scheduling of urban transit network can be formulated as an optimization problem of minimizing the overall transfer time of transferring passengers and initial waiting time of the passengers waiting to board a vehicle at their point of origin [6]. The fleet size required for a bus line is decreased by inserting express and partial services in the schedule with alternative fleet assignment strategies, which take the headways as inputs [7]. De Palma and Lindsey examined the schedule delay

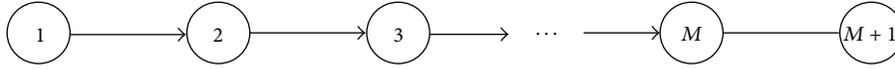


FIGURE 1: Illustration of an urban public transit line.

costs incurred from travelling earlier or later than desired and formulated an optimization model with the objective of minimizing the total riders' schedule delay costs [8]. Since passenger satisfaction with bus services has a high correlation with the attributes of the services supplied [9], Tom and Mohan formulated an optimization problem to minimize the bus operating cost and the passenger travel time by using frequency-coded genetic algorithms [10]. Assuming that no more than one passenger arrives at a station during a time interval, Niu and Zhou focused on optimizing a passenger train timetable in a heavily congested urban rail corridor [11].

A number of studies have somehow tried to model operational control strategies that are integrated with the scheduling problem to improve the efficiency of public transit systems. Site and Filippi proposed a particular service pattern including not only the short-turn strategy but also the vehicle size and frequencies that are required to supply the service over different operation periods [12]. By skipping a number of stops in high-frequency transit operations, the deadheading strategy could shorten the waiting time of passengers in the bus transport system [13]. Leiva et al. built an optimization method for designing limited-stop services, which minimizes both the travel time of passengers and the operating cost of an urban bus corridor by optimizing the transport services with different arrival frequencies of various types of vehicles [14]. Fu and Liu proposed a real-time scheduling strategy that aimed to strike an optimal balance between the benefits of operators and passengers [15]. Under the availability of real-time information and historical data of the system, the total user delay can be reduced by using a multipoint holding control strategy for a bus transit system [16]. Without considering the capacity constraints, Niu developed a non-linear programming model and proposed a bilevel genetic algorithm to design a transit schedule with uneven headways and skip-stop operations [17].

The existing studies have paid much attention to regular schedules, but they have failed to explore the scheduling problem with uneven vehicle-departure headways based on time-dependent passenger demands. Adopting a similar approach to that used by dynamic traffic assignment in urban road network, our research divides an operation day into several equal time periods to formulate the transit scheduling problem. These periods cannot be too short; otherwise, obtaining the demand records will become difficult; however, if the period length is too long, the model loses the dynamic-demands feature. Under normal conditions, such as a one-minute period, we face a challenge that more than two passengers could arrive at a station simultaneously during one same period. Such a case can surely lead to some confusion when determining the queuing order for passengers arriving at this period under strict capacity constraints.

For a congested urban public transit line, some passengers must wait for an extended duration and then board the next or the third departure vehicles. When a vehicle departs from a station, under such time-dependent and congested conditions, we should decide which passengers can board the vehicle and which cannot be on the vehicle. This issue is also associated with the calculation of the number of boarded passengers for a departure vehicle at each station. The main aim of this research is to optimize the departure times of transit vehicles under strict capacity constraints. The model proposed in this paper is a biobjective programming problem, in which a novel heuristic algorithm is developed to solve the model.

The remainder of the paper is organized as follows. Section 2 presents an analysis of the vehicle fleet operations and the passenger travel behaviors, along with the presentation of the underlying assumptions. A biobjective optimization model for the scheduling problem with strict capacity constraints for an urban public transit line is given in Section 3. In Section 4, a heuristic algorithm based on intelligent search and local improvement is developed to solve the proposed model. In Section 5, a numerical example is provided to illustrate the application of the model and algorithm. The last section concludes the paper and outlines the possibilities for future research in related areas.

2. Problem Statement

2.1. Analysis. This paper focuses on an urban transit line which connects several residential areas and a workplace during the morning rush hours. The stations on the urban public transit line are numbered as $1, 2, \dots, M, M+1$ along the vehicle's forward direction, as shown in Figure 1, where $1, 2, \dots, M$ are the boarding stations; and $M+1$ is the destination station. In this scenario, the demand distribution is represented as multisource and single sink; specifically, the passengers board at the intermediate stations and reach their common destination station.

The passenger demands associated with urban public transit systems are characterized as being time dependent and stochastic; herewith, we use $[0, T]$ to index the study time horizon. In order to measure the dynamic demands, we divide the time horizon $[0, T]$ equally into several time periods (e.g., one minute as a period), and we use τ to denote a period and \mathfrak{R} to represent the set of periods ($\tau \in \mathfrak{R}$).

Taking into account that all passengers have the same destination, this study uses $P_i(\tau)$ to indicate the number of passengers who arrive at station i ($i = 1, 2, \dots, M$) during period τ ($\tau \in \mathfrak{R}$) travelling to the destination station $M+1$. At the same time, this paper uses Ω to denote the set of vehicles that depart from the start station. Thus, the issue to be

addressed in this paper is transformed into the determination of the departure time d_j^1 of transit vehicle j ($j \in \Omega$) from station 1.

2.2. Assumptions. In order to objectively analyze the vehicle operation and the passengers' behaviour for urban public transit systems, the following assumptions are made throughout to facilitate the model formulation.

A1: all transit vehicles have the same capacity and follow strict capacity constraints. Once the number of in-vehicle passengers reaches the vehicle's full capacity, that vehicle is no longer available to any other passengers.

A2: all vehicles are scheduled according to the assumption that they travel at the same time between any two consecutive stations and they have the same dwell time for each station.

A3: the order for boarding a vehicle is randomly determined when several passengers arrive at a station simultaneously during the same period.

A4: all passengers who arrive at a station in different periods follow the principle of first-in-first-out (FIFO) when boarding a vehicle, which means that the passengers arriving within the earlier time periods board a vehicle prior to those arriving within the later periods.

A5: the total supply provided by the fleet meets the total passenger demand over the study horizon. Thus, some of the passengers may not board the current vehicle due to the oversaturation of demands, but the passengers waiting for the last vehicle could board the final run.

3. Formulation

3.1. Notation and Parameters. The following notation and parameters are used throughout this paper:

- i : index of stations, $i \in \{1, 2, \dots, M, M + 1\}$;
- \mathfrak{R} : set of periods;
- τ : index of periods, $\tau \in \mathfrak{R}$;
- Ω : set of vehicles departing from the start station;
- j : index of vehicles, $j \in \Omega$;
- d_j^i : departure time of vehicle j from station i , $i \in \{1, 2, \dots, M\}$, $j \in \Omega$;
- $P_i(\tau)$: number of passengers who arrive at station i during period τ and are travelling to the destination station, $i \in \{1, 2, \dots, M\}$, $\tau \in \mathfrak{R}$;
- Q_j^i : number of boarded passengers in vehicle j during the vehicle's stay at station i , $j \in \Omega$, $i \in \{1, 2, \dots, M\}$;
- r^i : vehicle running time from station i to station $i + 1$, $i \in \{1, 2, \dots, M\}$;
- h^i : dwell time at station i , $i \in \{2, 3, \dots, M\}$;

c : passenger loading capacity of a vehicle;

I_{\min} : prespecified minimum interval between two consecutive vehicles at the same station;

I_{\max} : prespecified maximum interval between two consecutive vehicles at the same station.

3.2. Constraints. The core decision variables are in fact the departure times for each vehicle at the start station. The detailed arrival and departure times at the other stations can be accordingly generated for scheduling applications. Given dwell time h^i and running time r^i , the departure time for each vehicle at each station should satisfy the following equation:

$$d_j^i = d_j^{i-1} + r^{i-1} + h^i. \quad (1)$$

Now, the minimum interval between two consecutive vehicles should be required to ensure the operational safety of vehicles, and the predetermined maximum interval should not be broken because the passenger waiting times at the stations cannot be too long. Thus, the departure times of two consecutive vehicles at the start station should satisfy the inequality

$$I_{\min} \leq d_{j+1}^1 - d_j^1 \leq I_{\max}. \quad (2)$$

To accurately calculate the number of boarded passengers for a given vehicle at the intermediate stations, our research introduces two parameters, namely, the boarded period b_j^i and residual number s_j^i . Figure 2 illustrates the two parameters by using the cumulative flow counts. The boarded period b_j^i indicates a period in which at least one passenger arriving at station i could board vehicle j , and no passenger at station i can board vehicle j offside this period. The residual number s_j^i denotes the number of remaining passengers for vehicle j at station i during the boarded period b_j^i .

If we use $R_{i,j}(\tau)$ to indicate the number of boarded passengers in vehicle j arriving at station i during period τ , the implication of a boarded period can be also expressed as follows:

$$b_j^i = \max \{ \tau \mid \tau \in \mathfrak{R}, R_{i,j}(\tau) > 0 \}. \quad (3)$$

The calculation of the residual number can be achieved as follows:

$$s_j^i = P_i(b_j^i) - R_{i,j}(b_j^i). \quad (4)$$

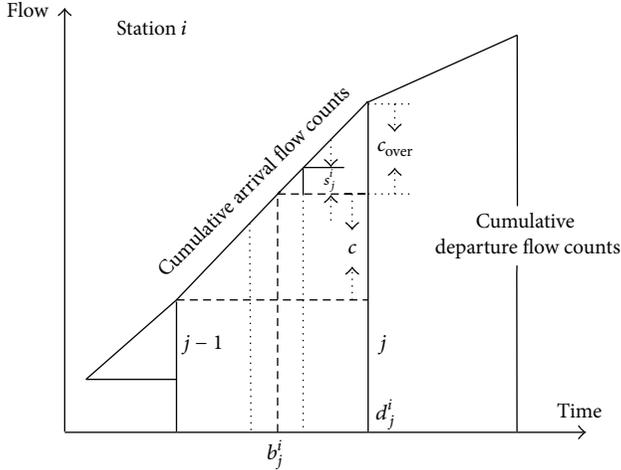


FIGURE 2: Illustration of board period and residual number.

On the other hand, the parameters of boarded period b_j^i and residual number s_j^i can be recursively calculated by

$$b_j^i = \min \left\{ \xi(d_j^i), \min \left\{ h \mid \sum_{i'=1}^{i-1} Q_{j-1}^{i'} + s_{j-1}^i + \sum_{\tau \in (b_{j-1}^i, h]} P_i(\tau) \geq c \right\} \right\},$$

$$s_j^i = \max \left\{ 0, s_{j-1}^i + \sum_{\tau \in (b_{j-1}^i, b_j^i]} P_i(\tau) - c \right\}, \quad (5)$$

where $\xi(d_j^i)$ indicates a period which covers time point d_j^i and that the boundary conditions are $b_0^1 = 0$ and $s_0^1 = 0$. The number of boarded passengers in vehicle j during the vehicle's stay at station i , Q_j^i , is thus determined by the equation

$$Q_j^i = s_{j-1}^i + \sum_{\tau \in (b_{j-1}^i, b_j^i]} P_i(\tau) - s_j^i. \quad (6)$$

3.3. Objective. The transit operation is associated with the passengers and the bus company. The objective of the passengers is to minimize the overall waiting times at the stations, while the company wants to decrease the number of required vehicles. The first objective function, which aims to minimize the number of required vehicles, is

$$\min |\Omega|. \quad (7)$$

For a period with short length, this study assumes that when passengers arrive at a station during the same period, they actually all arrive at the start of that period. Let $\varphi(\tau)$

indicate the initial moment of period τ , $\tau \in \mathfrak{R}$. When vehicle j departs from station i , the total of the newly boarded passengers' waiting times is $\sum_{\tau \in (b_{j-1}^i, b_j^i]} (d_j^i - \varphi(\tau)) \times P_i(\tau)$, and the additional waiting time associated with all remaining passengers together is $(d_{j+1}^i - d_j^i) \times s_j^i$. As a result, the second objective function, which aims to minimize the total waiting time of all passengers at all stations, is formulated as follows:

$$\min \sum_{i=1}^M \sum_{j \in \Omega} \left\{ (d_{j+1}^i - d_j^i) \times s_j^i + \sum_{\tau \in (b_{j-1}^i, b_j^i]} (d_j^i - \varphi(\tau)) \times P_i(\tau) \right\}. \quad (8)$$

4. Algorithm

The model proposed in this paper is a biobjective, non-linear programming problem, which is difficult to solve with conventional, gradient-based methods or commercial optimization solvers. As a result, a heuristic algorithm based on intelligent search and local improvement is developed to solve the proposed model. This paper first assumes that the departure times of vehicles fall on integer minute points, such as 6:31 and 6:32. The main idea of the algorithm consists of two searches. Under the condition of the vehicles used in the least numbers, the biobjective is achieved by searching for the departure times in order to minimize the current cumulative waiting times, for each run at the start station.

At the start station, all possible time points are searched forward from the earliest time, moving ahead one minute at a time. The search process requires checking the predetermined vehicle-interval constraints and counting the total number of in-vehicle passengers. If the number of passengers in the vehicle reaches the capacity as the vehicle arrives at the destination station, the vehicle should depart at the current time point; otherwise, the search should move one unit forward to check the next point until the end of the time horizon. The objectives of minimizing the number of required vehicles and the total waiting time of passengers could be achieved by taking the obtained points as the departure times of vehicles at the start station. After the forward search, checking the length of the interval between the last two consecutive vehicles is required. The corresponding points should be adjusted by the backward search if the interval constraints are not satisfied.

4.1. Forward Search. With particular respect to the algorithm, this paper uses p to indicate the search point associated with the departure time of the current vehicle, $A(j)$ to indicate the total number of boarded passengers for vehicle j as it arrives at the destination station, and $H(j) = d_j^1 - d_{j-1}^1$ to indicate the interval between vehicle j and vehicle $j-1$ at the start station.

The search area of the first vehicle is $p \in [1, I_{\max}]$ during the forward process. Once the departure time d_{j-1}^1 of vehicle $j-1$ is determined, the search area of vehicle j at the start station is $p \in [d_{j-1}^1 + I_{\min}, d_{j-1}^1 + I_{\max}]$ according to the

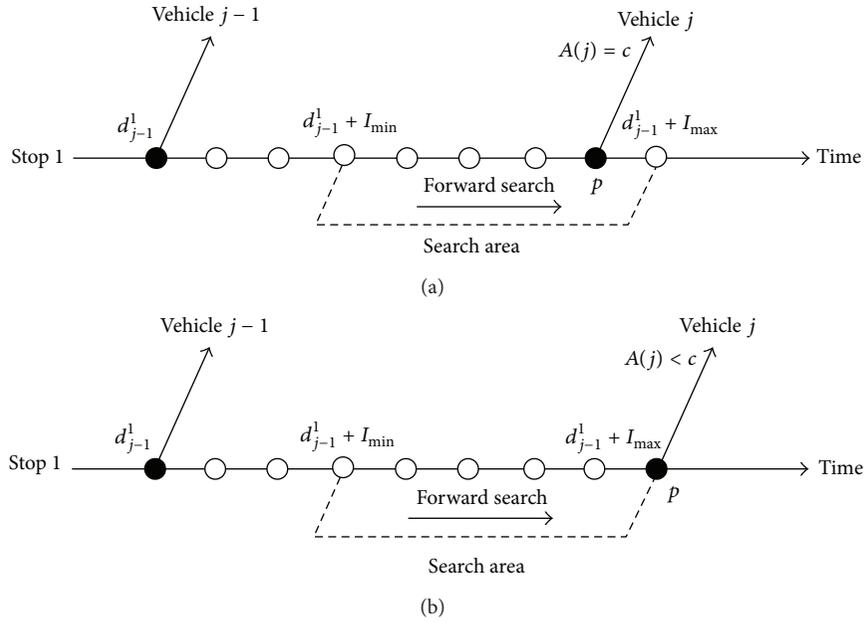


FIGURE 3: Illustration of search forward the possible times ahead one minute at a time.

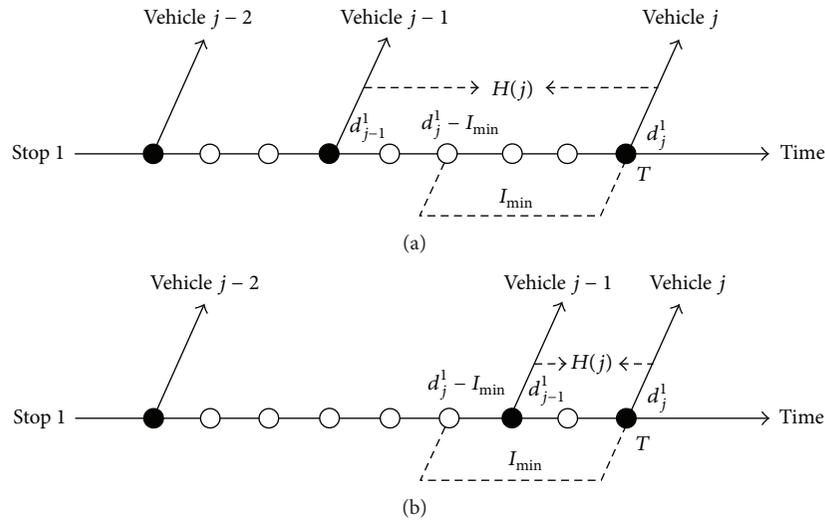


FIGURE 4: Checking the length of the interval between the last two consecutive vehicles.

headway constraints. All possible time points are searched forward from the start time $d_{j-1}^1 + I_{min}$ until a satisfactory time point is obtained. The following two cases are possible during the process of searching forward, as shown in Figure 3.

Figure 3(a) indicates that if the search reaches a point located within the predetermined area $[d_{j-1}^1 + I_{min}, d_{j-1}^1 + I_{max}]$ and vehicle j is already full (namely $A(j) = c$), the search should terminate and take the corresponding point as the departure time (namely $d_j^1 = p$). Figure 3(b) indicates that if the number of in-vehicle passengers $A(j)$ for vehicles j departing from all the search points within the predetermined area is less than the capacity c , the end moment should be taken as the departure time (namely $d_j^1 = d_{j-1}^1 + I_{max}$).

4.2. *Bound Check.* According to assumption A5 that the provided supply meets the total passenger demand, the passengers waiting for the last vehicle can board the final run. To satisfy this assumption, this paper also assumes that the last vehicle departs from the start station at T , the study's last moment. When the forward search reaches the last moment, T , either all vehicles satisfy the headway constraints, or the interval between the last two vehicles does not satisfy the minimum headway condition. As a result, we need to perform the bound check while the departure times are generated from the forward search. During the bound check process, the following two cases, shown in Figure 4, are possible.

Figure 4(a) shows a case where the bound check is not required because the interval $H(j)$ between vehicle j (the last

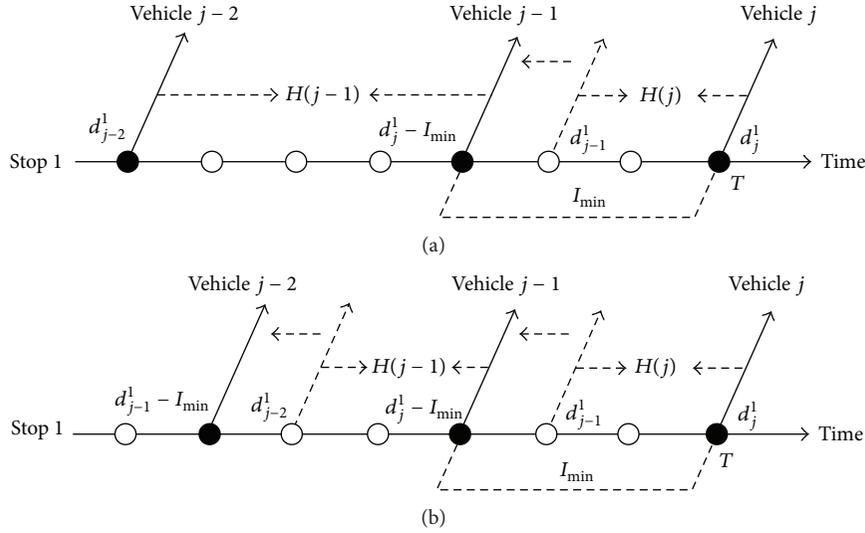
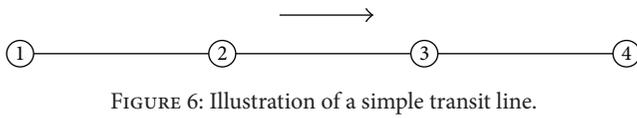


FIGURE 5: Illustration of backward adjustment in order to satisfy the minimum headway constraint.



run) and vehicle $j - 1$ at the start station has already satisfied the minimum headway constraint. Figure 4(b) shows, however, a case in which the minimum headway constraint is not satisfied.

4.3. Backward Adjustment. In the case occurring in Figure 4(b), adjustments must be made to those departure times that do not satisfy the minimum headway constraint. The adjustment should be executed from the last vehicle by moving backward the associated time point when the vehicle-interval at the start station violates the headway constraint. The interval $H(j)$ between vehicle j and vehicle $j - 1$ at the start station must be checked, as well as the interval $H(j - 1)$ between vehicle $j - 1$ and vehicle $j - 2$.

Figure 5(a) shows a case where the departure time of vehicle $j - 1$ should be moved backward to a point at which the interval between vehicle j and vehicle $j - 1$ is I_{\min} (so $H(j) = I_{\min}$), while the departure time of vehicle $j - 2$ is not required to move backward because the new interval $H(j - 1)$ between vehicle $j - 1$ and vehicle $j - 2$ still satisfies the minimum headway constraint. Figure 5(b), by contrast, shows a case where the departure time of vehicle $j - 2$ is required to move backward after an adjustment is made to the departure time of vehicle $j - 1$ because the new interval between vehicle $j - 1$ and vehicle $j - 2$ does not satisfy the minimum headway constraint.

4.4. Heuristic Algorithm. The above-described forward search, bound check, and backward adjustment form a heuristic algorithm, which is summarized as follows.

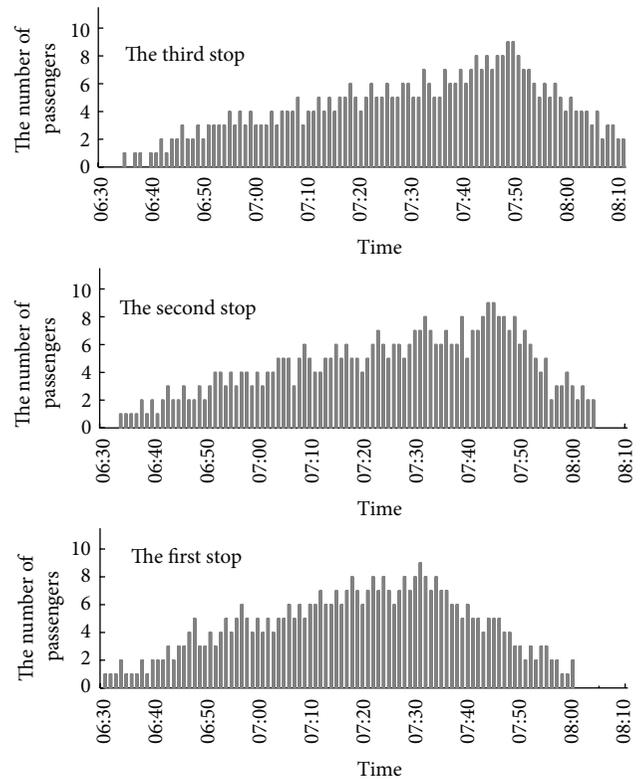


FIGURE 7: Passenger demands counted in one-minute intervals at intermediate stations.

Step 1. Set $p = 1$, $j = 1$, and $d_0^1 = 0$.

Step 2. Set $d_j^1 = p$. If $d_j^1 \geq T$, then $d_j^1 = T$ and go to Step 5; otherwise, if $d_j^1 - d_{j-1}^1 = I_{\max}$, go to Step 4, and if $d_j^1 - d_{j-1}^1 < I_{\max}$, go to Step 3.

TABLE 1: The departure times and the associated numbers of boarded and in-vehicle passengers.

Vehicle number	Departure time	The first station			The second station			The third station		
		BP/RN	Boarded passengers	In-vehicle passengers	BP/RN	Boarded passengers	In-vehicle passengers	BP/RN	Boarded passengers	In-vehicle passengers
1	6:40	10/0	13	13	15/0	17	30	22/0	26	56
2	6:47	17/0	19	19	22/0	19	38	29/2	22	60
3	6:53	23/0	22	22	28/0	22	44	34/6	16	60
4	6:58	28/0	25	25	33/0	18	43	38/9	17	60
5	7:02	32/0	18	18	37/0	18	36	44/3	24	60
6	7:06	36/0	20	20	41/0	20	40	48/3	20	60
7	7:10	40/0	22	22	45/0	20	42	52/5	18	60
8	7:14	44/0	25	25	49/0	21	46	54/12	14	60
9	7:17	47/0	20	20	52/0	15	35	59/4	25	60
10	7:20	50/0	21	21	55/0	18	39	62/0	21	60
11	7:24	54/0	30	30	59/0	23	53	64/16	7	60
12	7:27	57/0	20	20	62/0	22	42	67/17	18	60
13	7:30	60/0	23	23	65/0	19	42	70/20	18	60
14	7:32	62/0	17	17	67/0	13	30	74/5	30	60
15	7:35	65/0	22	22	70/0	19	41	76/9	19	60
16	7:38	68/0	19	19	73/0	22	41	79/16	19	60
17	7:41	71/0	16	16	76/0	26	42	81/18	18	60
18	7:44	74/0	14	14	79/0	23	37	84/11	23	60
19	7:48	78/0	18	18	83/0	24	42	88/13	18	60
20	7:53	83/0	13	13	88/0	17	30	95/2	30	60
21	8:00	90/0	14	14	95/0	16	30	102/0	14	44

TABLE 2: The departure times of vehicles at start station associated with a regular schedule.

Vehicle number	Departure time	Vehicle number	Departure time	Vehicle number	Departure time
1	6:36	8	7:08	15	7:36
2	6:42	9	7:12	16	7:40
3	6:48	10	7:16	17	7:44
4	6:52	11	7:20	18	7:48
5	6:56	12	7:24	19	7:52
6	7:00	13	7:28	20	7:56
7	7:04	14	7:32	21	8:00

Step 3. Calculate $A(j) = \sum_{s=1}^M Q_j^s$. If $A(j) = c$, go to Step 4; otherwise, set $p = p + 1$ and go to Step 2.

Step 4. Set $j = j + 1$, $p = p + I_{\min}$, and go to Step 2.

Step 5. Calculate $H(j) = d_j^1 - d_{j-1}^1$.

Step 6. If $H(j) \geq I_{\min}$, set $H(j - 1) = d_{j-1}^1 - d_{j-2}^1$ and go to Step 7; otherwise, set $d_{j-1}^1 = d_{j-1}^1 - 1$ and go to Step 5.

Step 7. If $H(j - 1) \geq I_{\min}$, then stop; otherwise, set $j = j - 1$ and go to Step 5.

5. Numerical Example

In this section, we consider a simple transit line on which there are three boarding stations and one destination station, as shown in Figure 6. The running time from station 1 to station 2 is 4 minutes; from station 2 to station 3, the time is 6 minutes; and from station 3 to station 4, the time is 15 minutes. The dwell time at each station is 1 minute. The study horizon is from 6:30 to 8:00, and the length of each subdivision is 1 minute. Each transit vehicle has a capacity of 60 passengers, and the maximum and minimum headways are 10 minutes and 2 minutes, respectively.

TABLE 3: Comparison of regular schedule with irregular schedule.

Schedule	Number of departed vehicles	Total waiting time (min)	Average waiting time (min)	Boarded passenger	Average load rate (%)
Irregular	21	1417	1.14	1240	98.41
Regular	21	3943	3.38	1168	92.70

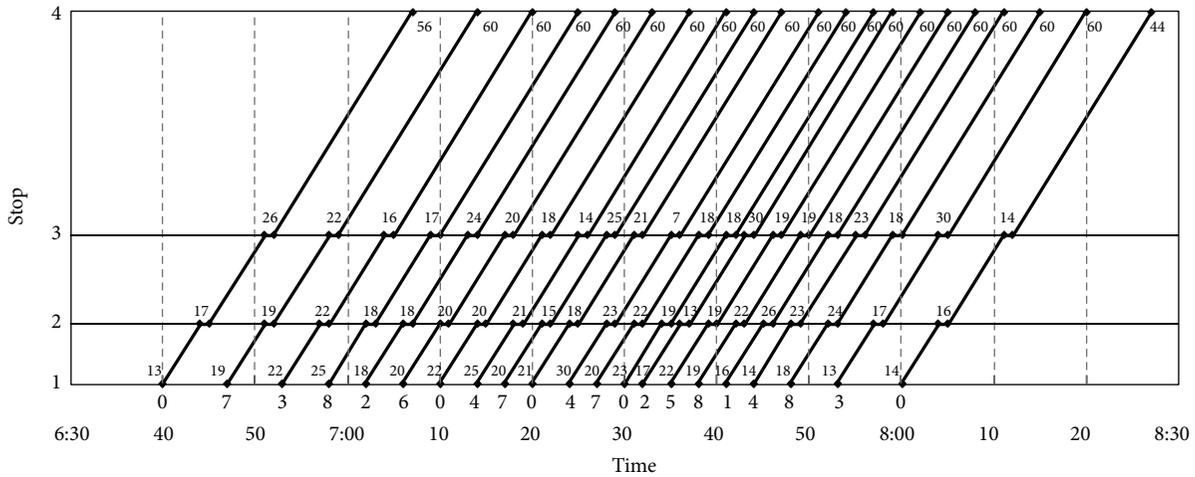


FIGURE 8: Illustration of transit schedule corresponding to heuristic algorithm.

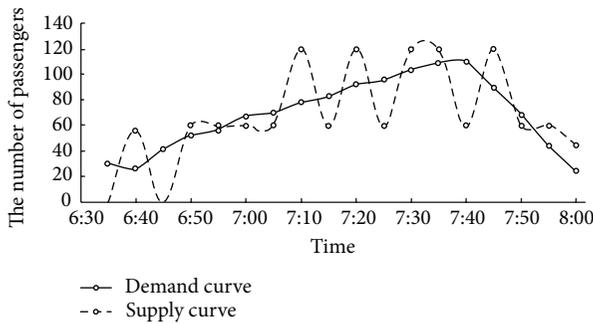


FIGURE 9: Illustration of demand and supply curves.

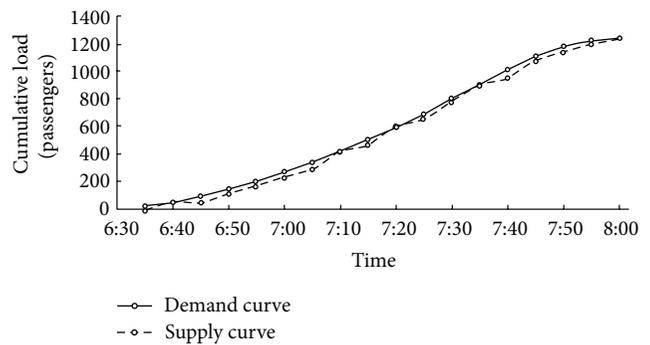


FIGURE 10: Cumulative curves of the demand and supply.

The period-dependent passenger demands, as shown in Figure 7, are illustrated by the total number of passengers arriving at the intermediate stations within each minute. By using the heuristic algorithm developed above, the departure times of vehicles at the start station and the associated numbers of boarded and in-vehicle passengers are determined; the results are given in Table 1. In this case, the necessary number of vehicles is 21.

In Table 1, BP and RN denote board period and residual number, respectively, and the total waiting time of all passengers at the stations is 1417 minutes. Using the information presented in Table 1, the corresponding schedule was created; this schedule is shown in Figure 8. Figure 8 shows that longer intervals between departure times occur from 6:30 to 7:00 than from 7:30 to 7:45. This temporal distribution of vehicles is approximately consistent with the passenger demands.

In Figure 9, the two curves illustrate the relation between the demand and supply associated with the designed schedule for the transit line. The demand curve shows the total number of passengers waiting at the stations every five minutes. The five-minute periods used at a subsequent station should be equivalent to those at the first station delayed by the running time between this subsequent station and the first station because the same vehicle arrives at that subsequent station after no less than the running time between the stations. The supply curve displays the number of boarded passengers every five minutes as the vehicles follow the designed schedule. The corresponding cumulative curves are shown in Figure 10, which indicates the correspondence between the demand and supply.

In order to further demonstrate the advantage of the method proposed in this paper, Table 2 shows a regular

schedule associated with the same fleet supply and other parameters; in this case, 72 passengers cannot board the last vehicle.

In Table 2, the intervals between the first three vehicles are 6 minutes, and those between the remaining vehicles are 4 minutes. A comparison between the designed schedule and the regular schedule is made in Table 3.

From Table 3, we can see that all passengers can board the vehicles using the designed schedule, while some cannot when using the regular one. Moreover, the average waiting time associated with the proposed method is reduced by 66.27%, which further demonstrates the efficiency of the method proposed in this paper.

6. Conclusions

This paper studies the transit scheduling problem to optimize the departure times of vehicles for a congested urban public transit line. Based on some reasonable assumptions, a biobjective optimization model has been established to minimize the number of required vehicles and the total waiting times of passengers. Under period-dependent demand and strict capacity conditions, our method has fully considered the situation in which there are more than two passengers arriving at a station simultaneously during one same period.

The proposed model is a nonlinear programming problem, which is difficult to solve with a conventional approach. Due to the potential of heuristic approaches to handle complex problems, a hybrid procedure embedded in a heuristic algorithm is applied to solve the model. The validation of the model and the algorithm has been successfully tested with the help of a numerical example. In future research, we will consider the response of passengers to the optimized schedule and extend our method to a more general transit case.

Acknowledgment

The work described in this paper was supported by the National Natural Science Foundation of China (Grant no. 71261014).

References

- [1] S. C. Wirasinghe, "Initial planning for urban transit systems," in *Advanced Modeling for Transit Operations and Service Planning*, pp. 1–29, Elsevier, Oxford, UK, 2003.
- [2] A. Ceder, *Public Transit Planning and Operation*, Elsevier, Linacre House, Jordan Hill, Oxford, UK, 2007.
- [3] V. Guihaire and J.-K. Hao, "Transit network design and scheduling: a global review," *Transportation Research Part A*, vol. 42, no. 10, pp. 1251–1273, 2008.
- [4] L. J. LeBlanc, "Transit system network design," *Transportation Research Part B*, vol. 22, no. 5, pp. 383–390, 1988.
- [5] K. K.-T. Lee, S. H. F. Kuo, and P. M. Schonfeld, "Optimal mixed bus fleet for urban operations," *Transportation Research Record*, no. 1503, pp. 39–48, 1995.
- [6] P. Chakroborty, K. Deb, and P. S. Subrahmanyam, "Optimal scheduling of urban transit systems using genetic algorithms," *Journal of Transportation Engineering*, vol. 121, no. 6, pp. 544–553, 1995.
- [7] T. A. S. Vijayaraghavan and K. M. Anantharamaiah, "Fleet assignment strategies in urban transportation using express and partial services," *Transportation Research Part A*, vol. 29, no. 2, pp. 157–171, 1995.
- [8] A. De Palma and R. Lindsey, "Optimal timetables for public transportation," *Transportation Research Part B*, vol. 35, no. 8, pp. 789–813, 2001.
- [9] P. Prioni and D. Hensher, "Measuring service quality in scheduled bus services," *Journal of Public Transportation*, vol. 3, no. 2, pp. 51–74, 2000.
- [10] V. M. Tom and S. Mohan, "Transit route network design using frequency coded genetic algorithm," *Journal of Transportation Engineering*, vol. 129, no. 2, pp. 186–195, 2003.
- [11] H. M. Niu and X. Zhou, "Optimizing urban rail timetable under time-dependent demand and oversaturated conditions," *Transportation Research Part C*, vol. 36, pp. 212–230, 2013.
- [12] P. D. Site and F. Filippi, "Service optimization for bus corridors with short-turn strategies and variable vehicle size," *Transportation Research Part A*, vol. 32, no. 1, pp. 19–38, 1998.
- [13] X. J. Eberlein, N. H. M. Wilson, C. Barnhart, and D. Bernstein, "The real-time deadheading problem in transit operations control," *Transportation Research Part B*, vol. 32, no. 2, pp. 77–100, 1998.
- [14] C. Leiva, J. C. Muñoz, R. Giesen, and H. Larrain, "Design of limited-stop services for an urban bus corridor with capacity constraints," *Transportation Research Part B*, vol. 44, no. 10, pp. 1186–1201, 2010.
- [15] L. Fu and Q. Liu, "Real-time optimization model for dynamic scheduling of transit operations," *Transportation Research Record*, no. 1857, pp. 48–55, 2003.
- [16] L. A. Koehler, W. Kraus Jr., and E. Camponogara, "Iterative quadratic optimization for the bus holding control problem," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1568–1575, 2011.
- [17] H. M. Niu, "Determination of the skip-stop scheduling for a congested transit line by bilevel genetic algorithm," *International Journal of Computational Intelligence Systems*, vol. 4, no. 6, pp. 1158–1167, 2011.

Research Article

Analysis on Residents' Travel Activity Pattern in Historic Urban Areas: A Case Study of Historic Urban Area of Yangzhou, China

Mao Ye,¹ Miao Yu,¹ Xiucheng Guo,² Yingshun Liu,¹ and Zhibin Li²

¹ Department of Transportation Engineering, Nanjing University of Science and Technology, 200 Xiaolingwei, Nanjing 210094, China

² School of Transportation, Southeast University, 2 Sipailou, Nanjing 210096, China

Correspondence should be addressed to Mao Ye; yemao0924@163.com

Received 12 September 2013; Revised 4 November 2013; Accepted 21 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Mao Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Travel behaviors and activity patterns in the historic urban area of a city are expected to be different from the overall situations in the city area. The primary objective of this study is to analyze the residents' travel activity patterns in historic urban area. Based on survey data conducted in the historic urban area of Yangzhou, the travel activities of local residents in a whole day were classified into five types of patterns. The multinomial logit (MNL) model was developed to evaluate the impacts of explanatory variables on the choices of activity patterns. The results showed that the choice of activity pattern was significantly impacted by five contributing factors including the gender, age, occupation, car ownership, and number of electric bikes in household. The other variables, which were the family population, preschoolers, number of conventional bikes in household, motorcycle ownership, and income, were found to be not significantly related to the choice of activities. The results of this study from historic urban area were compared to findings of previous studies from overall urban area. The comparison showed that the impacts of factors on activity pattern in the historic urban area were different from those in the overall area. Findings of this study provide important suggestions for the policy makings to improve the traffic situations in historic urban areas of cities.

1. Introduction

In the past decade, cities in China have experienced extremely large changes in their appearances and land use patterns. A number of new buildings and new road facilities were built in the city area. However, to keep the features of old buildings and save the panorama of history of the city, the historic urban area in a city has not been changed too much in the past years. Generally in China, the historic urban area is not only the concentrated residential area but also the political, commercial, and cultural center of the city. The historic urban area has a high population density, mixed land use pattern, and multiple types of travel activities. Thus, currently in cities of China, the historic urban areas are experiencing extremely severe traffic congestions as compared to other areas in cities. An analysis on the travel activity patterns in historic urban areas could help transportation planners better understand the characteristics of travel activities and develop strategies to reduce the congestions in such areas.

Traditionally, the aggregate methods such as the Four-Stage Method were commonly used to analyze the characteristics of travel activities [1–3]. However, the aggregate methods were based on the traffic analysis zone (TAZ) level which cannot reflect the travel behaviors of certain people in a certain area. Then the disaggregate modeling techniques, such as MNL model, MNP model, and NL model, were developed by researchers to model the travel behaviors in certain areas at an individual level [4–18]. The models developed in previous studies can be used to make a prediction on the decision of an activity or trip mode for an individual traveler in a single trip. Actually, however, travelers usually plan their travel activities with multiple purposes or destinations in a whole day simultaneously.

In recent years, many researchers have recognized the importance of multitrip patterns in travel behavior analysis. A sequence of multiple travel activities was described as an activity pattern or trip chain in previous studies [19–28].

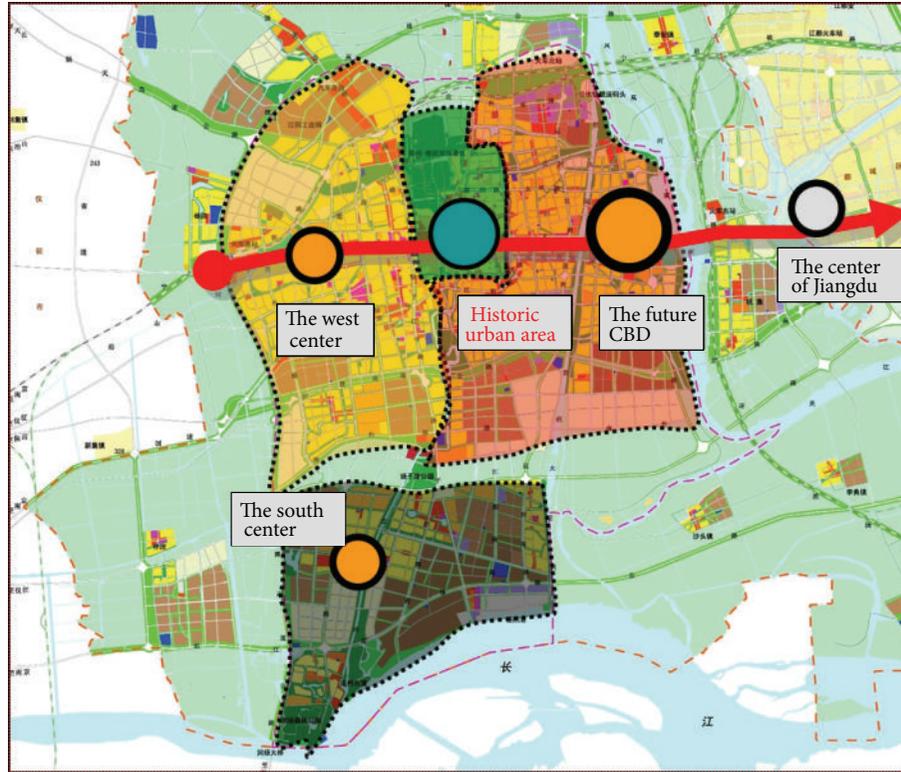


FIGURE 1: The location of historic urban area of Yangzhou.

The disaggregate modeling techniques were commonly used in the analysis of trip arrangement [9–17], trip route choice [10, 11], activity choice [18, 20, 21, 25], and trip chain pattern [19, 22–24, 27]. These studies have shown that the analysis on travelers' activity pattern (i.e., trip chains) can provide important information for transportation researchers and planners to understand the travel behaviors of travelers and improve the transportation system in city areas. The factors impacting individuals' travel behaviors include the sociodemographic characteristics, physical environments, attitudinal factors, and travel-related features. Some other factors such as land use pattern and road facility were also found to affect the residents' travel behaviors [29–33].

A review on the literature shows that though some previous studies have analyzed the characteristics of travel activity patterns, those studies generally paid attention to the overall situation in the whole city areas. None of previous studies have paid attention to the travel activity pattern in the historic urban area of a city. The travel behaviors and activity patterns within the historic urban area of a city are expected to be quite different from the overall situations in the city. For example, due to the concentrated land use pattern, the trip distance within the historic area is quite short as compared to that in the overall city area. Besides, the facilities of streets as well as attributes of trip modes are also very different from those of the overall city area. As a consequence, the findings from the overall cities in previous studies cannot accurately reflect the features of travel activity patterns in the historic urban areas. A study particularly focuses on the historic urban area which is important to help understand the unique

characteristics of travel activities in the area. In addition, the study results that particularly focus on travel behaviors in historic urban areas can provide useful suggestion for the policy makings or city planning in the historic urban areas.

The primary objective of this study is to analyze the choice of residents' travel activity pattern in historic urban areas. Based on the survey data conducted in the historic urban area of Yangzhou, the travel activities of local residents in a whole day were classified into five types of patterns. The individual and household characteristics significantly related to the choice of activity patterns were investigated using the multinomial logit modeling (MNL) technique. The findings from historic urban areas were then compared to previous studies. The remainder of this study is organized as follows. The next section introduces the data resource. Section 3 introduces the methods used for modeling. Section 4 gives the data analysis results and discusses the findings. This paper ends with brief conclusions in Section 5.

2. Data

2.1. Study Area. The historic urban area in the city of Yangzhou was considered in this study to conduct the relevant analysis on the travel activity pattern. The description of the urban area of Yangzhou are shown in Figure 1. The historic urban area of Yangzhou is surrounded by the ancient canal, moat, and internal moat which covers about 5.09 square kilometers. The historic urban area is located in the central part of Yangzhou and is the commercial, financial, cultural, and medical center of the city. The city area of Yangzhou



FIGURE 2: The corridor schematic of historic urban area of Yangzhou.

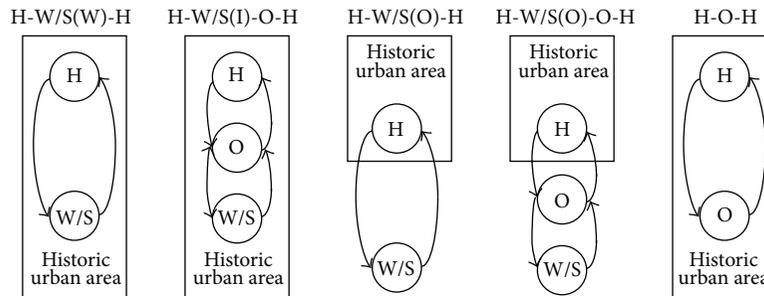


FIGURE 3: Illustrations of five types of travel activity patterns.

contains several subgroups with a distribution of a belt from east to west, as shown in Figure 2. The historic urban area is located in the middle of the belt. Because of the city structure, the traffic from east to west and west to east all travels through the historic urban area. As a result, the historic urban area has quite a large amount of traffic which makes the area become congested.

The population of the historic urban area of Yangzhou in year 2012 is about one hundred thousand. The distribution of age of residents in this area presents a pattern of dumbbell: most middle-aged people with higher economic status have moved out of the historic urban area to the new developed area of the city, and the majority of people who currently live in the historic urban area are retired or self-employed with low income. Besides, there are many elementary and middle schools in the historic urban area which makes this area have a large number of young travelers. On average, the population density in the historic urban area is much higher than the other parts of Yangzhou. The travel activities in the historic urban area are expected to be different from the other areas, which are discussed in later sections.

2.2. Data Resource. Data used for analysis were obtained from the household survey in the city of Yangzhou, China, in year 2010. The household survey was conducted by the local government in one typical weekday to draw up the planning for transportation system in the city. The survey included two parts: (1) individual and household characteristics and (2) travel information of all trips in the whole day. A total of 2,000 questionnaires were assigned randomly to residents in the historic urban area. Questionnaires were distributed and collected by the neighborhood committees. Initially, 1,528 questionnaires were obtained. Samples with the following issues were excluded from further analysis: (1) cases of missing key information (such as trip mode or trip purpose) and (2) cases that have logic problems in the encoding

process. Totally, 1,221 samples were obtained after the data selection.

In this study, a travel activity pattern (or trip chain) is defined as a sequence of trips that starts and ends at the home location in a whole day. According to previous studies, an activity pattern can be classified into different types by the complexity of activities (i.e., single and complex patterns) or trip purpose (i.e., subsistence and nonsubsistence patterns). In this study, the information of travel activity pattern was extracted from the household survey data in Yangzhou. Activity types that have small sample sizes (less than 1%) were excluded from analysis to avoid inaccurate estimates. Since this study focuses on the historic urban area in a city, the classification of activities is different from previous ones.

Finally, five major types of travel activity pattern were used for analysis, which are home-work/school (in historic urban area)-home (H-W/S(I)-H), home-work/school (in historic urban area)-other-home (H-W/S(I)-O-H), home-work/school (outside historic urban area)-home (H-W/S(O)-H), home-work/school (outside historic urban area)-other-home (H-W/S(O)-O-H), and home-other-home (H-O(I)-H). The illustrations of these activity patterns are shown in Figure 3. The descriptions of these five activity patterns are given as follows, where “h” denotes home, “w” denotes work, “s” denotes school, and “o” refers to other activities (nonsubsistence activities).

H-W/S(I)-H: there is one subsistence activity within a day. This activity pattern contains only a simple commuting activity stop. All the home and commuting stops are within the historic urban area.

H-W/S(I)-O-H: there are two types of activities within a day. This activity pattern is a combination of a simple commuting chain with a noncommuting activity stop. All the home and activity stops are within the historic urban area.

H-W/S(O)-H: there is one subsistence activity within a day. This activity pattern contains only a simple commuting activity stop. The home is located within the historic urban

area but the work/school is located outside the historic urban area.

H-W/S(O)-O-H: there are two types of activities within a day. This activity pattern is a combination of a simple commuting chain with a noncommuting activity stop. The home is located within the historic urban area, but at least one of the activity stops are outside the historic urban area.

H-O(I)-H: there is only one nonsubsistence activity within a day. This chain contains only a simple noncommuting activity stop. All the home and intermediate stop are within the historic urban area.

3. Methodology

In this study, we would like to estimate the impacts of explanatory variables on the choice of activity patterns in the historic urban area. The dependent variable includes five categories which are the five types of activity patterns. Thus, the MNL model which is based on the logistic distribution for disordered category variables is used for the data analysis. The MNL model has been widely used by researchers for the analysis of choices of trip modes and activities [26–28]. The methodology of MNL model was briefly introduced in this section.

On the basis of the random utility theory, the unity of n selects the activity pattern of i which can be expressed as

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

where U_{in} is the utility sample n for activity pattern i , V_{in} is called the systematic components of utility, and ε_{in} is the random parts.

When it is a linear relationship between V_{in} and the explanatory variables, it contains V_{in} which can be express as

$$V_{in} = \theta' X_{in} = \sum_{k=1}^K \theta_k x_{ink}, \quad (2)$$

where K is the number of explanatory variables; θ_k is coefficient to be estimated; x_{ink} is an explanatory variable.

Assuming that the random terms of the utility function subject to the double exponential distribution, the probability that the activity pattern i selected by the individual n is

$$P_{in} = \frac{\exp V_{in}}{\sum_{j \in A_n} \exp V_{jn}} = \frac{\exp \left(\sum_{k=1}^K \theta_k x_{ink} \right)}{\sum_{j \in A_n} \exp \left(\sum_{k=1}^K \theta_k x_{jnk} \right)}. \quad (3)$$

For estimation of the parameter vectors θ by maximum likelihood, the log likelihood function is shown as follows:

$$LL = \sum_{n=1}^N \left(\sum_{k=1}^K \delta_{ik} \left[\theta_k X_{kn} - LN \sum_{\forall k} \text{EXP}(\theta_k X_{kn}) \right] \right). \quad (4)$$

Since the estimation process of maximum likelihood is not the major objective of this study, the details of the estimation process are not given here. Interested readers could refer to Washington et al. [28] for more details on the estimation of MNL models.

TABLE 1: Trip characteristics between Yangzhou and its historic urban area.

Categories	Yangzhou	Historic urban area
Average daily travel times	2.81 times/day	2.99 times/day
Trip mode		
Walk	16.4%	27.4%
Electric bicycle/bicycle	53.9%	50.3%
Car	5.5%	3%
Public transit	6.3%	8.5%
Motorcycle	15.5%	9.9%
Other	2.4%	0.9%
Age		
<50 years old	70.8%	61.4%
>50 years old	29.2%	38.6%
Occupation		
Retiree	16.5%	25.4%
student/staff/service	58.0%	49.9%
Private	25.5%	24.7%
Trip purpose		
School	24.7%	19.3%
Work	6.8%	6.3%
Home	47.1%	47.3%
Other	21.4%	27.1%

4. Results and Discussion

4.1. Preliminary Analysis. In the first section we have discussed that the travel behaviors in the historic urban area of a city are expected to be quite different from the overall situations. In this section, the preliminary analysis was first conducted to compare the basic characteristics of travels between the historic urban area and the overall city. The results are shown in Table 1. The contents in Table 1 suggest that an analysis of travel activity pattern particularly on the historic urban area is necessary.

It is found that the travel behaviors are quite different between the two area scopes: (1) the average daily travel time in historic urban area (2.99 times/day) is more than that of overall area in Yangzhou (2.81 times/day). (2) The percentage of age above 50 in historic urban area (38.6%) is higher than the overall city area of Yangzhou (29.2%), and the occupations in historic urban area are also different from the overall area. (3) The percentage of walk mode in historic urban area (27.4%) is significantly higher than in the overall area of Yangzhou (16.4%), which could be because the trip distance within the historic urban area is shorter. (4) There are more public transit travels and less bicycle, car, and motorcycle in the historic urban area. (5) The trips to work/school in historic urban area are a little lower than overall area, and the trips to home in historic urban area higher.

Table 2 reports the statistics of activity patterns by type in the sample. The activity patterns “H-W/S(I)-H” and “H-O(I)-H” are the most common patterns that present 41.3 and 33.9 of all activities, respectively, which is comparable to

TABLE 2: Statistics of five types of travel activity patterns.

Activity pattern	Frequency	Percentage
H-W/S(I)-H	505	41.3
H-W/S(I)-O-H	183	15
H-W/S(O)-H	96	7.9
H-W/S(O)-O-H	23	1.9
H-O(I)-H	414	33.9
Total	1221	100

some previous studies [18, 23, 25]. The activity pattern “H-W/S(O)-O-H” occupies the least market which is only 1.9%. The information in Table 2 suggests that most of the activities were made within the historic urban area.

In the preliminary analysis we also compared the proportion of trip modes for the five types of activity patterns. The results are shown in Table 3. It is found that bicycle (including electric bike and conventional bike) is the nominate trip mode in the commuting activities (which are H-W/S(I)-H, H-W/S(I)-O-H, and H-W/S(O)-H) with the percentages over 50%. In the two nonsubsistence activity patterns (which are H-W/S(O)-O-H, and H-O(I)-H) bicycle usages are much lower. Public transit occupies a low mode share in all the activity patterns. Car is most commonly used for nonsubsistence travels outside the historic urban area. Besides, motorcycle accounts for a relative large number of travels in most of activity patterns. The information in Table 3 suggests that the mode share is different between activity patterns in the historic urban area of Yangzhou.

Our research team also investigated several personal attributes in the five types of travel activity patterns in the historic urban area. The results are shown in Figure 4. It is found in Figure 4(a) that there are more male travelers in the subsistence activity patterns H-W/S(I)-H and H-W/S(O)-H, and less males in the nonsubsistence patterns including H-W/S(I)-O-H, H-W/S(O)-O-H, and H-O(I)-H. The reason could be that females are more likely to have shopping or other maintenance travels than males. Figure 4(b) suggests that people older than 50 are more likely to have nonsubsistence activities. The percentage of population over 50 years old in activity pattern H-O(I)-H is much higher than the other patterns. The information of Figure 4 suggests that the personal (and household) characteristics have significant impacts on the choice of travel activity patterns, which are discussed in detail in the following section.

4.2. Modeling Results. The choice of travel activity pattern is impacted by the personal and household characteristics. In the household survey of this study, personal attributes include gender, occupation, and age. While household population, preschoolers, private cars, the number of bikes and electronic bikes, motorcycle, and incomes constitute household attributes. Affected by those factors, residents plan their daily activities and adopt appropriate activity patterns. The statistics of explanatory variables used for model development are shown in Table 4.

The MNL model was developed in this section to identify the impacts of explanatory variables listed in Table 4 on the choice of activity pattern. Initially, all explanatory variables were considered in the models. Variables not significantly related to the outcome were excluded from the model specification step by step. The contributing factors were kept in the model specification. The variable selection processes were repeated to carefully determine the contributing factors in the final model. The multinomial logistic module of statistical software SPSS was used to calibrate the parameters of each available in the MNL model. The estimation results of the MNL model shown in Table 5.

In the MNL model, the activity pattern H-O(I)-H is specified as the reference category. All the parameter estimates in Table 5 reflect the impacts of each variable on the changes of probabilities of the activity patterns with respect to the probability of pattern H-O(I)-H. The model estimates are interpreted as follows.

Gender is estimated to have significant effects on the probabilities of H-W/S(I)-H, H-W/S(I)-O-H and H-W/S(O)-H, activity patterns. The coefficient of gender for the H-W/S(W)-H and H-W/S(O)-H patterns is 0.76 and 0.891, and for the H-W/S(W)-H pattern is -0.559 . It suggests that male travelers prefer to choose the activity pattern of H-W/S(I)-W and H-W/S(O)-H, of which the probabilities are 2.138 and 2.437 times more than female travelers. The reason for this phenomenon is that females could be more likely to make shopping and other types of travels for maintenance and recreations. Thus females are more likely to make H-W/S(I)-H activity pattern between work and home.

The occupation variable (OCC = 1) estimated in the H-W/S(I)-H, H-W/S(I)-O-H, and H-W/S(O)-H activity patterns is negative, and OCC = 2 are all positive in the four activity patterns. The reasons for the results could be that students/staff/service participate in work/school while retirees have little such trip purposes which coincides with the real situation. The coefficients of age variable (AGE = 1 and AGE = 2) for H-W/S(I)-H, H-W/S(I)-O-H, and H-W/S(O)-H activity patterns estimated in the model are all positive. It could be because the residents younger than 50 are mainly staff and students, which could have more commuting activities than these older than 50.

Several household attributes such as car ownership and number of electric bicycles are found to have significant impacts on the probability of H-W/S(O)-H activity pattern. The coefficient of car ownership and number of electric bicycles is -2.064 and -2.177 , indicating that travelers that do not have a car or electric bicycles is less likely to choose the activity pattern H-W/S(O)-O-H. The reason could be that car and electric bicycle have relative high speeds and are quite convenient to use. Without a car or electric bicycle travelers that do not make activities with multiple purposes outside the historic urban area. Some variables including the income, household population, preschoolers, and number of conventional bicycles and motorcycles are not found to have obvious impact on the choice of travel activity pattern in the historic urban area.

The modeling results of choice of activity pattern from the historic urban area are compared to the findings from

TABLE 3: The proportion of trip modes for the five types of activity patterns.

Activity pattern	Walk (%)	Bicycle (%)	Car (%)	Public transit (%)	Motorcycle (%)
H-W/S(W)-H	5.3	65.6	4.4	6.1	18.6
H-W/S(W)-O-H	10.5	50.9	6.7	7.4	24.5
H-W/S(O)-H	0	58.6	4.0	8.1	29.3
H-W/S(O)-O-H	13.0	34.8	26.1	4.3	21.8
H-O(I)-H	35.2	39.7	9.1	10.7	5.3

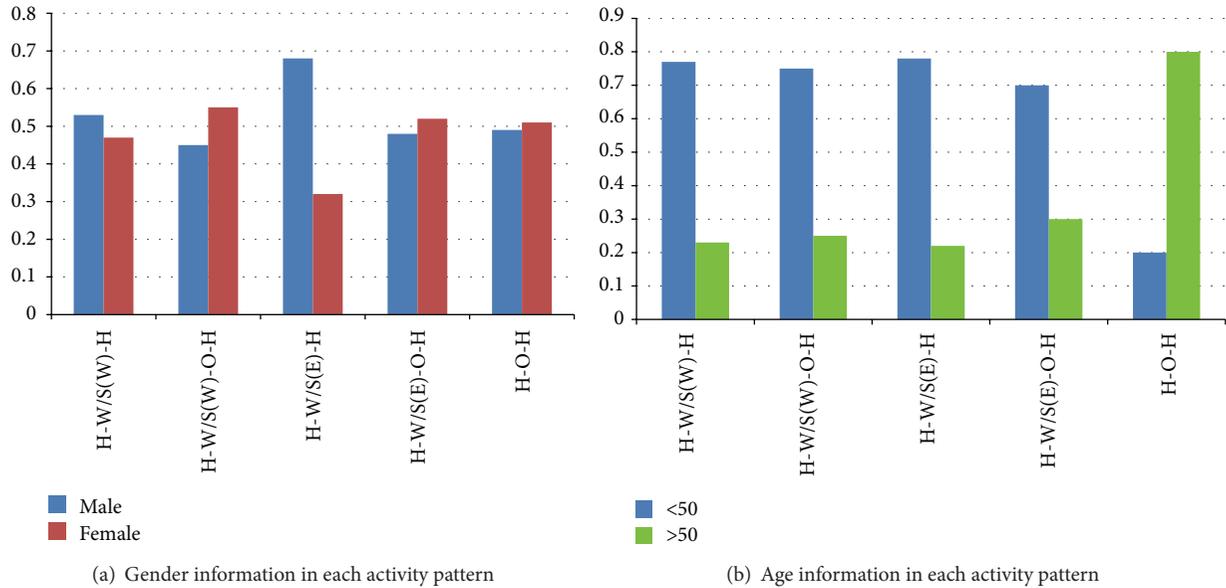


FIGURE 4: Personal attributes in the five activity patterns in historic urban area.

previous studies which focus on the overall city areas. Several consistent findings are obtained for the two study scopes. For example, both this study and previous ones report that the individual characteristics including age, gender, and occupation are significantly related to the choice of activity pattern. For example, the studies conducted by Allaman et al. [34], Lu and Pas [35], and Kuppam and Pendyala [36] all showed that the personal attributes such as gender, age, and occupation significantly affected travelers' activities. The findings from our study further confirmed the results of previous studies.

There are some inconsistent findings between our study and previous ones. For the historic urban areas as found in this study, only two household characteristics variables were found to be significantly related to activity choices. The other household related variables were not found to be contributing factors which was different as compared to previous findings for the overall city areas. For example, Kuppam and Pendyala [36] and Lu and Pas [35] found that the household income, number of bicycles in household, and number of family members were found to be contributing factors which were not found in this study. The number of children in household was found to be an important factors in the study by Golob [37] and Lu and Pas [35]. The reason for the difference of findings could be that in the historical areas, the choice

of activity pattern is restricted by the characteristics of historical areas such as limited road resource and large traffic demands. In our case of Yangzhou in 2012, the road length per capita in the overall city is 8.84 m and the road area per capita is 13.94 m², while the two measures in the historic area are 13.26 m and 19.45 m², respectively. And the average daily travel times in the overall city is 2.81 times/day, while in the historic area is 2.99 times/day. Thus, travelers with different household characteristics have to make similar activity choices, making the impacts of household factors on travel activities less significant.

The possible reason for the different findings could be that in the historic urban area most of the people have relative low income level which makes the variable of income not significant. Another reason could be that due to the limited road resource and high activity frequency the people that have higher income level do not have more selections on the travels. In historic areas the width of streets is much lower than the other places in the city area, and the buses usually run slow due to the congested traffic in historic areas. Thus, these reasons lead to the results that the number of bicycles in household is not significantly related to activity choice. Household population is not significant, probably because there are many retirees in historic urban area of Yangzhou (the proportion of people over 50 years old is 29.2% for

TABLE 4: Statistics of explanatory variables for model development.

Attributes	Variable	Definition	Frequency (%)
Personal attributes	Gender (GEN)	Male (=1)	653 (53.5%)
		Female (=0)	568 (46.5%)
	Occupation (OCC)	Retiree (=1)	273 (22.4%)
		Student/staff/service (=2)	599 (49.1%)
		Private (=0)	349 (28.6%)
	Age (AGE)	<30 (=1)	579 (47.4%)
30–49 (=2)		118 (9.7%)	
>50 (=0)		524 (42.9%)	
Household attributes	Household population (HP)	<3 (=1)	397 (32.5%)
		More than 3 (=0)	824 (67.5%)
	Preschoolers (CH)	Have (=1)	200 (16.4%)
		Without (=0)	1021 (83.6%)
	Car ownership (CAR)	Have (=1)	68 (5.6%)
		Without (=0)	1153 (94.4%)
	Number of conventional bicycle (BIC)	0 (=0)	254 (20.8%)
		1-2 (=1)	855 (70%)
		>3 (=2)	112 (9.2%)
	Number of electric bicycle (EBIC)	0 (=0)	498 (40.8%)
		1-2 (=1)	698 (57.2%)
		>3 (=2)	25 (2%)
Motorcycle ownership (MOT)	Have (=1)	477 (39.1%)	
	Without (=0)	744 (60.9%)	
Income (INC)	<¥ 10000 (=1)	190 (15.6%)	
	¥ 10000–50000 (=2)	900 (73.7%)	
	>¥ 50000 (=0)	131 (10.7%)	

TABLE 5: Estimation results of the MNL model.

	H-W/S(I)-H ^a		H-W/S(I)-O-H		H-W/S(O)-H		H-W/S(O)-O-H	
	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.	Coeff.	Sig.
Intercept	-2.329	0.002	-1.969	0.002	-3.469	0.002	-5.493	0.002
[GEN = 1]	0.76	0.001	-0.559	0.027	0.891	0.007		
[OCC = 1]	-2.604	<0.001	-1.443	<0.001	-2.076	0.008		
[OCC = 2]	2.297	<0.001	2.191	<0.001	2.687	<0.001	1.624	0.001
[AGE = 1]	1.521	<0.001	1.091	0.024	1.51	0.005		
[AGE = 2]	1.143	<0.001	1.297	<0.001	1.421	<0.001		
[HP = 1]	^b							
[CH = 1]								
[CAR = 1]							2.064	0.001
[BIC = 1]								
[BIC = 2]								
[EBIC = 1]							2.177	0.040
[EBIC = 2]								
[MOT = 1]								
[INC = 1]								
[INC = 2]								

^aThe reference category of activity pattern is H-O(I)-H.

^bEstimates of variables not significant at a 90% confidence level are not given.

the overall city area and 38.6% for the historic urban area), and they have done shopping and other travel behaviors during their travels. The families with preschoolers are always taken care of by the retirees, so the impact of preschoolers on travel activity pattern is not significant either.

5. Conclusions

This study analyzed the characteristics of residents' travel activity patterns in historic urban area of a city and evaluated the impacts of explanatory variables on the choice of activity pattern. The data used for analysis in this study were obtained from the household survey in the historic urban area of Yangzhou. Based on the data, the travel activities of local residents in a whole day were classified into five types of patterns which were the (H-W/S(I)-H), home-work/school (in historic urban area)-other-home (H-W/S(I)-O-H), home-work/school (outside historic urban area)-home (H-W/S(O)-H), home-work/school (outside historic urban area)-other-home (H-W/S(O)-O-H), and home-other-home (H-O(I)-H). The multinomial logit (MNL) model was developed in this study to evaluate the impacts of explanatory variables on the choices of activity patterns. The findings of this study were compared to previous ones.

Based on the data analysis results, the following conclusions were obtained.

- (1) Bicycle is the dominate mode of transport in historic urban area for the subsistence or commuting activity patterns H-W/S(I)-H, H-W/S(I)-O-H, and H-W/S(O)-H. Car is most commonly used for activity patterns H-W/S(O)-O-H which contains trips outside the historic urban area. Public transit is most commonly used in the single noncommuting activity pattern H-O(I)-H. The overall mode share of public transit is quite small, indicating that improving the level of service of public transit will help reduce the pressure of traffic in historic urban area.
- (2) Male travelers prefer to select the activity pattern H-W/S(W)-H and H-W/S(O)-H, and female are prone to choose the H-W/S(W)-O-H pattern. The households without car and electric bicycle seldom select the H-W/S(O)-O-H pattern. The residents who are younger than 50 and students/staff/services are more likely to choose the H-W/S(I)-H, H-W/S(I)-O-H, and H-W/S(O)-H patterns.
- (3) Among family attributes, ownership of private car and the number of electronic bicycles have significant effect on the choice of activity pattern. The other variables, which were the family population, preschoolers, number of conventional bike in household, motorcycle ownership, and income, were found to be not significantly related to the choice of activities.

Findings of this study on the travel activity patterns in historic urban areas could help transportation planners better understand the characteristics of travels and develop strategies to reduce the congestion in such areas. This study provides important suggestions for the policy makings to

improve the traffic situations in historic urban areas of cities. However, due to the limitations of the survey data, we just get the relationship between historic urban area resident activity pattern and personal and family attributes. We did not consider the trip mode selection during the analysis on travel activity pattern choice. In our future studies, we would like to analyze the choices of trip modes and activity patterns simultaneously and consider the correlations between the two choices. Besides, many travelers in historic urban area are tourists other than local residents. It is interesting to compare the characteristics of travel activities between tourists and residents. Future studies may focus on the above topics.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the Projects of the National Natural Science Foundation of China (no. 51208256 and no. 51178157). The authors would like to thank the senior students from Department of Transportation Engineering of Nanjing University of Science and Technology for their assistance in data collection and reduction. Especially, the authors thank Dr. Chen Yang and Dr. Zhibin Li in Southeast University for offering to help in this research.

References

- [1] A. B. Sosslau, A. B. Hassan, M. M. Carter, and G. V. Wickstrom, "Quick response urban travel estimation techniques and transferable parameters, user guide, NCHRP report 187," in *Proceedings of the Annual Meeting of the Transportation Research Board*, pp. 22–63, Washington, DC, USA, 1978.
- [2] M. Florian, M. Gaudry, and C. Lardinois, "A two-dimensional framework for the understanding of transportation planning models," *Transportation Research Part B*, vol. 22, no. 6, pp. 411–419, 1988.
- [3] E. Weiner, "Urban transportation planning in the United States: an historical overview, fifth edition," Tech. Rep. DOT-T-97-24, US Department of Transportation, Washington, DC, USA, 1997.
- [4] M. Ben-Akiva, J. L. Bowman, and D. Gopinath, "Travel demand model system for the information era," *Transportation*, vol. 23, no. 3, pp. 241–266, 1996.
- [5] B. De Geus, I. De Bourdeaudhuij, C. Jannes, and R. Meeusen, "Psychosocial and environmental factors associated with cycling for transport among a working population," *Health Education Research*, vol. 23, no. 4, pp. 697–708, 2008.
- [6] J. Pucher and R. Buehler, "Making cycling irresistible: lessons from the Netherlands, Denmark and Germany," *Transport Reviews*, vol. 28, no. 4, pp. 495–528, 2008.
- [7] J. Pucher, J. Dill, and S. Handy, "Infrastructure, programs, and policies to increase bicycling: an international review," *Preventive Medicine*, vol. 50, pp. S106–S125, 2010.
- [8] I. N. Sener, N. Eluru, and C. R. Bhat, "Who are bicyclists? why and how much are they bicycling?" *Journal of the Transportation Research Board*, vol. 2134, pp. 63–72, 2009.

- [9] Y. Xing, S. L. Handy, and P. L. Mokhtarian, "Factors associated with proportions and miles of bicycling for transportation and recreation in six small US cities," *Transportation Research Part D*, vol. 15, no. 2, pp. 73–81, 2010.
- [10] J. G. Su, M. Winters, M. Nunes, and M. Brauer, "Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada," *Transportation Research Part A*, vol. 44, no. 7, pp. 495–505, 2010.
- [11] C. Yang, W. Wang, X. Shan, J. Jin, J. Lu, and Z. B. Li, "Effects of personal factors on bicycle commuting in developing countries: case study of Nanjing, China," *Transportation Research Record*, no. 2193, pp. 96–104, 2010.
- [12] E. Heinen, B. van Wee, and K. Maat, "Commuting by bicycle: an overview of the literature," *Transport Reviews*, vol. 30, no. 1, pp. 59–96, 2010.
- [13] E. Heinen, K. Maat, and B. Van Wee, "The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances," *Transportation Research Part D*, vol. 16, no. 2, pp. 102–109, 2011.
- [14] J. Pucher and R. Buehler, "Analysis of bicycling trends and policies in large North American cities: lessons for New York," Final Report, University Transportation Research Center, 2010.
- [15] Z. B. Li, W. Wang, P. Liu, and D. R. Ragland, "Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities," *Transportation Research Part D*, vol. 17, no. 3, pp. 256–261, 2012.
- [16] Z. B. Li, W. Wang, P. Liu, J. Bigham, and D. R. Ragland, "Modeling bicycle passing maneuvers on multilane separated bicycle paths," *Journal of Transportation Engineering*, vol. 139, no. 1, pp. 57–64, 2013.
- [17] Z. B. Li, W. Wang, C. Yang, and D. R. Ragland, "Bicycle commuting market analysis using attitudinal market segmentation approach," *Transportation Research Part A*, vol. 47, pp. 56–68, 2013.
- [18] Z. B. Li, W. Wang, C. Yang, and G. Jiang, "Exploring the causal relationship between bicycle choice and trip chain pattern," *Transport Policy*, vol. 29, pp. 170–177, 2013.
- [19] J. G. Strathman and J. K. Dueker, "Understanding trip chaining, 1990 NPTS special reports on trip and vehicle attributes," Tech. Rep. FHWA/PL-95-033, 1-1-1-28, FHWA, US Department of Transportation, 1995.
- [20] C.-H. Wen and F. S. Koppelman, "A conceptual and methodological framework for the generation of activity-travel patterns," *Transportation*, vol. 27, no. 1, pp. 5–23, 2000.
- [21] H. Timmermans, P. van der Waerden, M. Alves et al., "Spatial context and the complexity of daily travel patterns: an international comparison," *Journal of Transport Geography*, vol. 11, no. 1, pp. 37–46, 2003.
- [22] N. McGuckin, J. Zmud, and Y. Nakamoto, "Trip-chaining trends in the United States: understanding travel behavior for policy making," *Journal of the Transportation Research Board*, vol. 1917, pp. 199–204, 2005.
- [23] M. Yang, W. Wang, X. Chen, T. Wan, and R. Xu, "Empirical analysis of: commute trip chaining case study of Shangyu, China," *Journal of the Transportation Research Board*, vol. 2038, pp. 139–147, 2007.
- [24] X. Ye, R. M. Pendyala, and G. Gottardi, "An exploration of the relationship between mode choice and complexity of trip chaining patterns," *Transportation Research Part B*, vol. 41, no. 1, pp. 96–113, 2007.
- [25] S. Krygsman, T. Arentze, and H. Timmermans, "Capturing tour mode and activity choice interdependencies: a co-evolutionary logit modelling approach," *Transportation Research Part A*, vol. 41, no. 10, pp. 913–933, 2007.
- [26] K. A. Small and C. Hsiao, "Multinomial logit specification tests," *International Economic Review*, vol. 26, no. 3, pp. 619–627, 1985.
- [27] T. Adler and M. Ben-Akiva, "A theoretical and empirical model of trip chaining behavior," *Transportation Research Part B*, vol. 13, no. 3, pp. 243–257, 1979.
- [28] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis*, CRC Press, Boca Raton, Fla, USA, 2nd edition, 2010.
- [29] M. Boarnet and R. Crane, "The influence of land use on travel behavior: specification and estimation strategies," *Transportation Research Part A*, vol. 35, no. 9, pp. 823–845, 2001.
- [30] K. K. W. Yim, S. C. Wong, A. Chen, C. K. Wong, and W. H. K. Lam, "A reliability-based land use and transportation optimization model," *Transportation Research Part C*, vol. 19, no. 2, pp. 351–362, 2011.
- [31] P. Waddell, G. F. Ulfarsson, J. P. Franklin, and J. Lobb, "Incorporating land use in metropolitan transportation planning," *Transportation Research Part A*, vol. 41, no. 5, pp. 382–410, 2007.
- [32] Q. Su, "Travel demand in the US urban areas: a system dynamic panel data approach," *Transportation Research Part A*, vol. 44, no. 2, pp. 110–117, 2010.
- [33] V. Van Can, "Estimation of travel mode choice for domestic tourists to Nha Trang using the multinomial probit model," *Transportation Research Part A*, vol. 49, pp. 149–159, 2013.
- [34] P. M. Allaman, F. C. Dunbar, and D. Steinberg, "Impacts of demographic trends on time allocations for household activities," in *Proceedings of the Annual Meeting of the Transportation Research Board*, pp. 231–244, Washington, DC, USA, 1981.
- [35] X. Lu and E. I. Pas, "Socio-demographics, activity participation and travel behavior," *Transportation Research Part A*, vol. 33, no. 1, pp. 1–18, 1999.
- [36] A. R. Kuppam and R. M. Pendyala, "A structural equations analysis of commuters' activity and travel patterns," *Transportation*, vol. 28, no. 1, pp. 33–54, 2001.
- [37] T. F. Golob, "A simultaneous model of household activity participation and trip chain generation," *Transportation Research Part B*, vol. 34, no. 5, pp. 355–376, 2000.

Research Article

Using Principal Component Analysis to Solve a Class Imbalance Problem in Traffic Incident Detection

Changjiang Zheng,¹ Shuyan Chen,² Wei Wang,² and Jian Lu²

¹ College of Civil and Transportation Engineering, Hohai University, Nanjing 210098, China

² Transportation School, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Changjiang Zheng; zhenghhu@sina.com

Received 21 August 2013; Revised 6 November 2013; Accepted 12 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Changjiang Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High imbalances occur in real-world situations when a detection system needs to identify the rare but important event of a traffic incident. Traffic incident detection can be treated as a task of learning classifiers from imbalanced or skewed datasets. Using principal component analysis (PCA), a one-class classifier for incident detection is constructed from the major and minor principal components of normal instances. Experiments are conducted with a real traffic dataset collected from the A12 highway in The Netherlands. The parameters setting, including the significance level, the percentage of the total variation explained, and the upper bound of the eigenvalues for the minor components, is discussed. The test results demonstrate that this method achieves better performance than partial least squares regression. The method is shown to be promising for traffic incident detection.

1. Introduction

Early detection of traffic incidents can minimize the delay experienced by drivers, wasted fuel, emissions, and lost productivity, while also reducing the likelihood of secondary collisions [1]. Traffic incident detection is thus a critical issue and it is important to develop mechanisms to detect traffic incidents as early as possible. The incident detection problem has received great interest from researchers and many incident detection techniques have been developed. Black and Sreedevi [2] have extensively reviewed many approaches to incident detection. Existing incident detection methods fall into the following major categories: pattern recognition, time series analysis, Kalman filters, partial least squares regression [3, 4], and data mining technologies, which include neural networks, fuzzy logic, support vector machines (SVM) [5, 6], rough set [7], ensemble learning [8, 9], and decision tree learning [10]. Data mining technologies have been shown to be the most promising techniques of the incident detection methods.

The typical classifiers in data mining, such as decision tree inductive systems or neural networks, are designed to optimize overall accuracy without accounting for the relative

distribution of each class. As a result, these classifiers tend to neglect small classes while focusing on classifying the large classes accurately [11]. Unfortunately, these classifiers perform poorly in incident detection because the real-world traffic data suffers from class imbalances that typically contain much fewer incident cases than incident-free cases. Such situations pose challenges for these classifiers. Many solutions to the class imbalance problem have been previously proposed both at the data and algorithmic levels. At the data level, resampling methods are commonly used to address the class imbalance problem [12, 13]. Although such approaches can be very simple to implement, tuning resampling methods to be effective in an application is not an easy task. At the algorithmic level, cost-sensitive learning [14], one-class classifiers [15], and ensemble-based classifiers, such as Boosting and Adaboost [16, 17], are very well-known approaches for solving dataset imbalance problems.

Typically, in a conventional multiclass classification problem, data from two (or more) classes are available and the decision boundary is supported by the presence of examples from each class. However, a one-class classifier completely neglects one of the classes and learning is accomplished using examples from a single class only at the learning stage.

Different researchers have used other terms to present similar concepts, such as outlier detection, novelty detection or concept learning [18]. One-class approaches to solving classification problems may be superior to discriminative (two-class) approaches, such as decision trees or neural networks [19]. Raskutti and Kowalczyk [15] demonstrated that one-class learning with positive-class examples can be a very robust classification technique when dealing with extremely unbalanced datasets composed of a high-dimensional noisy feature space.

When a traffic incident occurs, the associated traffic data change dramatically, so that the incident observation is different from most of the traffic data. Thus, we can detect traffic incidents by recognizing that the traffic incident data deviate significantly from normal traffic data. From this point of view, one-class classifiers or outlier detection can be employed to address the incident detection problem. A one-class classifier can be built from normal data to detect any deviation from the normal model in the observed data. Given a set of normal data to train from, together with a new piece of test data, the goal is to determine whether the test data belong to “normal” or anomalous behavior. For traffic incident detection, the task is to build an incident detector from the traffic data, that is, a predictive model capable of distinguishing between abnormal traffic states (called incidents) and normal states.

Shyu et al. [20] proposed a novel outlier detection scheme, based on principal components that was applied to intrusion detection. The underlying assumption of such a method is that intrusions appear as outliers in the normal data. Similarly, traffic incident cases also appear as outliers in the normal traffic data. In this paper, this principal component-based approach was used to detect traffic incidents. We used this approach for incident detection mainly due to the advantages of the principal component approach over many anomaly detection methods. First, the principal component approach does not make any distributional assumption. Second, this approach is typically used with high-dimensional datasets. Another benefit of this scheme is that the statistics can be computed in a shorter time during the detection stage so that the scheme can be used in real time. Last but not least, the detection model can be built using only normal cases, thereby avoiding the class imbalance problem.

This paper is organized as follows. Section 2 provides background on principal component analysis (PCA) and outlier detection. Section 3 provides details of the datasets used in the experiments, followed by an analysis of the results; the results are discussed in Section 4. Finally, the work concludes in the last section.

2. Anomaly Detection Scheme

2.1. Principal Component Analysis (PCA). Principal component analysis (PCA) describes the variance-covariance structure of a set of variables in terms of fewer new variables that are linear combinations of the original variables. The new variables are easily obtained from eigenanalysis of the covariance matrix or the correlation matrix of the original data.

If the variables are measured on scales with widely different ranges or if the units of measurement are not commensurate, it is preferable to perform PCA on the sample correlation matrix. In our study, we perform PCA on the correlation matrix of the normal group because the features of the data, such as time, volume, or speed, are measured on different scales.

Let the original dataset $X = [X_1 \ X_2 \ \cdots \ X_p]$ be an $n * p$ data matrix of n observations, each consisting of p variables, and let $R = [R_1 \ R_2 \ \cdots \ R_p]$ be the correlation matrix of X . If $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$ are the p eigenvalue-eigenvector pairs of the matrix \mathbf{R} , then the i th principal component is

$$y_i = z\mathbf{e}_i = z_1e_{1i} + z_2e_{2i} + \cdots + z_pe_{pi}, \quad (i = 1, 2, \dots, p), \quad (1)$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, $\mathbf{e}_i = (e_{1i}, e_{2i}, \dots, e_{pi})'$ is the i th eigenvector and $z = (z_1, z_2, \dots, z_p)$ is the standardized matrix defined as

$$z_k = \frac{x_k - \bar{x}_k}{s_k}, \quad (k = 1, 2, \dots, p), \quad (2)$$

where \bar{x}_k and s_k are the average and standard deviation of the observations in the k th dimension of X , respectively.

2.2. Outlier Detection. Most datasets often contain one or a few samples that do not conform to the general behavior of the dataset and which are called outliers. When an observation is different from most of the data or is sufficiently unlikely under the assumed probability model for the data, the observation is considered to be an outlier. With data on a single feature, outliers are those that are either very large or very small relative to the others. Many features correspond to a complex situation. In high dimensions, all features need to be considered together using a multivariate approach.

The procedure commonly used to detect multivariate outliers is to measure the distance of each observation from the center of the data using the Mahalanobis distance. Any observation larger than a threshold value is considered to be an outlier. The threshold is typically determined from the empirical distribution of the distances.

PCA has long been used for multivariate outlier detection. Consider the sample principal components y_i ($i = 1, 2, \dots, p$) of an observation x , the sum of squares of the standardized principal component scores is then given by

$$\sum_{i=1}^p \frac{y_i^2}{\lambda_i} = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} + \cdots + \frac{y_p^2}{\lambda_p}, \quad (3)$$

which is equivalent to the Mahalanobis distance of the observation x from the sample mean [20]. It is customary to examine the individual principal components or some functions of the principal components for outliers. Hawkins [21] obtained superior performance for a scheme that used statistics derived from principal components to detect errors in multivariate data.

Because the sample principal components are uncorrelated, under the normal assumption and assuming the sample size is large, the major components are given as follows:

$$\sum_{i=1}^q \frac{y_i^2}{\lambda_i} = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} + \dots + \frac{y_q^2}{\lambda_q}, \quad q \leq p, \quad (4)$$

corresponding to a chi-square distribution with q degrees of freedom. For a given significance level α , an observation x is an outlier if the following criterion is satisfied:

$$\sum_{i=1}^q \frac{y_i^2}{\lambda_i} > \chi_q^2(\alpha), \quad (5)$$

where $\chi_q^2(\alpha)$ is the upper α percentage point of the chi-square distribution with q degrees of freedom. The value of α indicates the error or false alarm probability in classifying a normal observation as an outlier. The number of major components, q , can be determined from the amount of the variation in the training data that are described by these components. Based on experiments, Shyu et al. suggested using q major components that can explain approximately 50 percent of the total variation of the standardized features.

In addition to the major components, Shyu et al. [20] proposed that the minor components, $\sum_{i=p-r+1}^p (y_i^2/\lambda_i)$, should be used to detect observations that do not conform to the normal correlation structure. The value of r can be determined by examining those components whose variance or eigenvalue is less than 0.20.

An observation is an outlier with respect to the correlation structure if

$$\sum_{i=p-r+1}^p \frac{y_i^2}{\lambda_i} > \chi_r^2(\alpha), \quad (6)$$

where $\chi_r^2(\alpha)$ is the critical value for a chi-square distribution with r degrees of freedom testing at a given significance level α .

The anomaly detection scheme differs from other existing approaches in the use of both the major and minor components of the data. A clear advantage of this scheme over others is that outliers can be detected based on being extreme values or not having the same correlation structure as the normal data.

2.3. Incident Detection Scheme. In this paper, the principal component classifier (PCC) mentioned above is used to detect traffic incidents. The procedure for traffic incident detection is as follows.

Step 1. Divide the original dataset into a training set and a testing set. Note that the outlier detection model is constructed from only the normal instances in the training set that correspond to nonincident traffic data.

Step 2. Perform PCA on the correlation matrix of this training set. Determine the number of major components q and the number of minor components r according to the obtained eigenvalues.

Step 3. Determine the outlier thresholds, c_1 and c_2 , for a given error probability in the distribution of the test statistics.

As previously mentioned, the collected data are assumed to conform to a multivariate normal distribution, such that the thresholds are the inverse of the chi-square cumulative distribution function; that is,

$$c_1 = \chi_q^2(\alpha_1), \quad c_2 = \chi_r^2(\alpha_2). \quad (7)$$

In practice, the normality assumption seldom holds true. Therefore, we opt to set the outlier thresholds based on empirical distributions of the test statistics rather than on the chi-square distribution. The values of α_1 and α_2 are chosen to reflect the relative importance of the types of outliers we would like to detect. Without loss of generality, we choose $\alpha_1 = \alpha_2$.

Step 4. Compute the principal component scores for the major components and the minor components for each observation x in the testing dataset.

Step 5. Classify x in the testing set as an outlier corresponding to an incident state if

$$\sum_{i=1}^q \frac{y_i^2}{\lambda_i} > c_1 \quad \text{or} \quad \sum_{i=p-r+1}^p \frac{y_i^2}{\lambda_i} > c_2. \quad (8)$$

3. Data Description

In this study, we investigated PCC performance in incident detection, as applied to real-life loop detector traffic data collected from the A12 freeway in The Netherlands. The distance between two adjacent detectors installed on A12 is approximately equal to 500 m. Individual vehicle data are collected by the detectors. More specifically, the passing time, the speed, the occupancy time, and the lane in which the vehicle is being driven are recorded for every car passing the detector.

To assess PCC performance in incident detection, two different datasets were employed. The first dataset consisted of loop detector data from many neighboring detectors installed on the A12 Dutch freeways collected during December 2007. The second data source contained information on all the registered incidents that occurred over the period considered on the A12 freeway.

3.1. Loop Detector Data. The loop detector data, in the form of lane-specific traffic volume and speed, were collected at 60-second intervals during normal conditions and incident conditions. Because one incident occurred on the road rather than in a lane, we computed the average volume and speed over all the lanes.

We cannot expect all collected data to be of a high quality. Suspicious or dirty data may be buried in a dataset. Upon close examination of the data, cases where the traffic volume and speed changed dramatically were easily found, but these cases did not always correspond to an incident case registered over that time period in the incident database. If no incident occurred during the respective time period, dramatic changes

in the traffic volume or speed may have been due to a faulty detector or transmission distortion. For anomaly detection to function effectively, such nonpertinent data must be captured and removed from the dataset to improve the veracity and reliability of the traffic information. This procedure is called data quality control or data cleaning. Various methods may be used for this purpose. More details on these methods can be found in the literature [22–24].

3.2. Incident Data. The database contained information on the incidents that occurred during December 2007 as mentioned above; the database included the following information for each incident:

- (i) the location of the incident,
- (ii) the “approximate” starting time and ending time of the incident,
- (iii) a short description of the incident (the lanes in which the incident occurred, a qualitative description of the incident, etc.).

Note, however, that the exact times (start times as well as end times) of the incidents were unknown. The database contained only the time at which the incident was reported and the time at which the end of the incident was reported. Thus, we must keep in mind that the actual starting time would have been a little earlier than the reported starting time.

We constructed the incident dataset using all the incidents with duration between 20 and 180 minutes. We obtained a total of 95 incidents that occurred on the A12 freeways during December 2007. There were 6 columns in this dataset, including the date, direction, location, start time, end time, and duration of the traffic incident.

3.3. Construction of Training and Testing Sets. Incident detection was based on section-related traffic data, which means that the traffic data were collected from two adjacent detectors: an upstream detector and a downstream detector. Each incident instance included the following items:

- (i) time (reported as hh + mm/60) of data collection,
- (ii) traffic volume and speed from the upstream detector,
- (iii) traffic volume and speed from the downstream detector,
- (iv) traffic state,

where the item “traffic state” is a label with a value of -1 or 1 , which denotes that there was no incident or that an incident occurred, respectively, as determined by the incident dataset.

Each instance in the dataset contained 5 feature values and was labeled as either a normal traffic state or an incident state. The entire dataset was divided into two parts, a training set and a testing set, which were used for calibration and testing, respectively. The training set was composed of 99961 nonincident instances and 2751 incident instances (58 incident cases), collected from 1 to 14 December; the testing set was composed of 67791 samples, including

65957 nonincident instances and 1834 incident instances (33 incident cases), collected from 15–31, December.

The outlier thresholds are determined from the normal instances in training data, and they were used to decide the label of each instance in the testing set; thus the model’s detection ability can be computed. The incident instances in the training set were of no use in the PCC. However, the incident instances were used to build other incident detection models for comparison with PCC, such as the partial least squares (PLS) model.

4. Case Studies

Incident detection results are typically evaluated using the following criteria: the detection rate (DR), the false alarm rate (FAR), and the mean time to detection (MTTD). The classification rate (CR) is an additional performance measure of interest. Detailed information on these criteria can be found in references [3–10].

The experiments were conducted within the following framework:

- (i) only major components were used to detect outliers,
- (ii) both major components and minor components were used,
- (iii) PCC was compared with PLS.

Matlab subroutines were written for these procedures and all the algorithms were run on a computer with a 1.50 GHz Intel Pentium processor and 1024 MB of memory.

4.1. Use of Only Major Components. To determine the appropriate number of major components to be used in PCC, we conducted a preliminary study by varying the percentage of total variation explained by the major components. In the distribution of the test statistics, the major components scores were closer to a log-normal distribution than to any other distribution.

The major components were considered to account for 50% up to 100% of the total variation at 5% increments and a 0.04 significance level; thus, we obtained 11 classifiers with different numbers of major components, ranging from 2 to 5. All these classifiers were evaluated using the same testing set. The results showed that as the percentage of the variation explained increased, corresponding to increasing the number of major components used, the values of the four measurements, DR, FAR, MTTD, and CR, appeared to stay the same until 90% of the variation was explained. It means that the PCC method based on the major components that can explain from 50% to 90% of the total variation has the same ability for incident detection.

When more than 90% of the variation was explained, corresponding to all the major components being included in the classifiers, the measurement values changed: DR and FAR increased, while MTTD and CR decreased. Thus, PCC with more major components yielded a higher false alarm rate and a lower classification rate, which is undesirable in incident detection. Figure 1 illustrates the changes in DR,

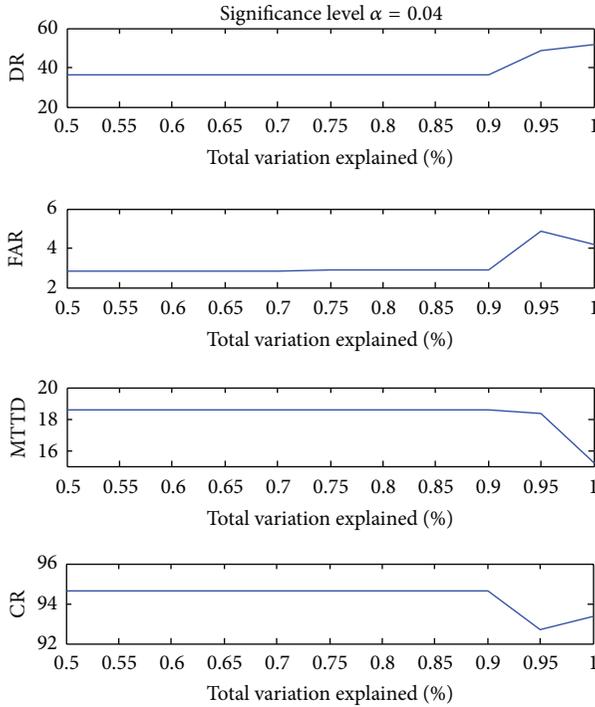


FIGURE 1: DR, FAR, MTTD, and CR versus the percentage of the variation explained.

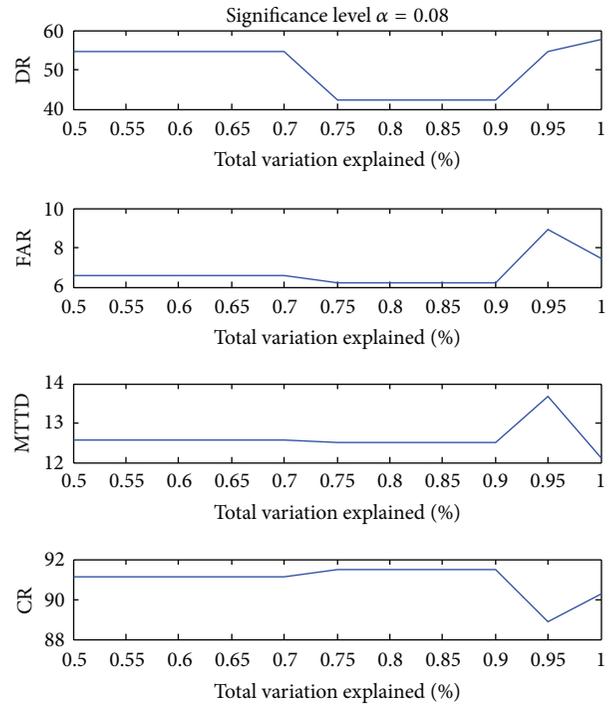


FIGURE 2: DR, FAR, MTTD, and CR versus the percentage of the variation explained.

FAR, MTTD and CR, as a function of the percentage of the variation explained for a significance level of 0.04.

The significance level was then increased from 0.01 to 0.10 in 0.01 increments and the process described above was repeated. A total of 110 classifiers were obtained. Applying all these classifiers to the same testing set, we observed that when the significance level exceeded 0.06, the PCC models with major components explaining 75% to 90% of total variation performed poorly, as evidenced by a rapid decrease in the detection rate. This result is illustrated in Figure 2 for a significance level of 0.08.

The results suggest that a value of q of 2 or 3 should be used at lower significance levels, while a q value of 2 should be used at a significance level larger than 0.06, when the classifier is only constructed using the major components.

4.2. Use of Both Major Components and Minor Components.

We next constructed outlier models using PCC with both the major and minor components; the performance of these models when applied to traffic incident detection was then evaluated. First, the percentage of the variation explained was set to 70% and the upper bound of the eigenvalues for the minor components was set to 0.2 [20]; thus, the numbers of the major and minor components were found to be 2 and 1, respectively. We constructed a series of outlier models with significance levels ranging from 0.01 to 0.10 in 0.01 increments. We always used the same values for α_1 and α_2 in (7), as we did not know in advance which type of outliers we should pay more attention to.

In the distribution of the test statistics, the major components scores followed a log-normal distribution with a freedom degree of 2, while the minor components scores followed a Weibull distribution with a freedom degree of 1: these distributions fit the test statistics more closely than any other distribution.

Figure 3 presents the testing results of these classifiers, showing how the four measurements changed with the significance level. Within increasing significance levels, DR increased while the MTTD decreased, which is desirable for an incident detection model. However, the FAR values were too high to apply the classifiers at a high significance level.

Keeping the percentage of the variation explained the same; the upper bound of the eigenvalues for the minor components was increased to 0.3 and the number of the minor components was increased from 1 to 2; we then constructed another 10 classifiers and tested their performance.

Next, we let the major components account for 90% of the total variation and let the upper bound of the eigenvalues for the minor components range from 0.2 to 0.3; we repeated the experiments above again.

Table 1 shows the testing results obtained for 4 classifiers that were built with different numbers of major components and minor components due to different values of the parameters setting. The significance level for all 4 classifiers was set to 0.08. The first column is the parameters setting; for example, (70%, 0.2) indicated that the major components accounted for 70% of the total variation and that the upper bound of the eigenvalues for the minor components was 0.2. The column “numbers” corresponds to the number of major and minor components used.

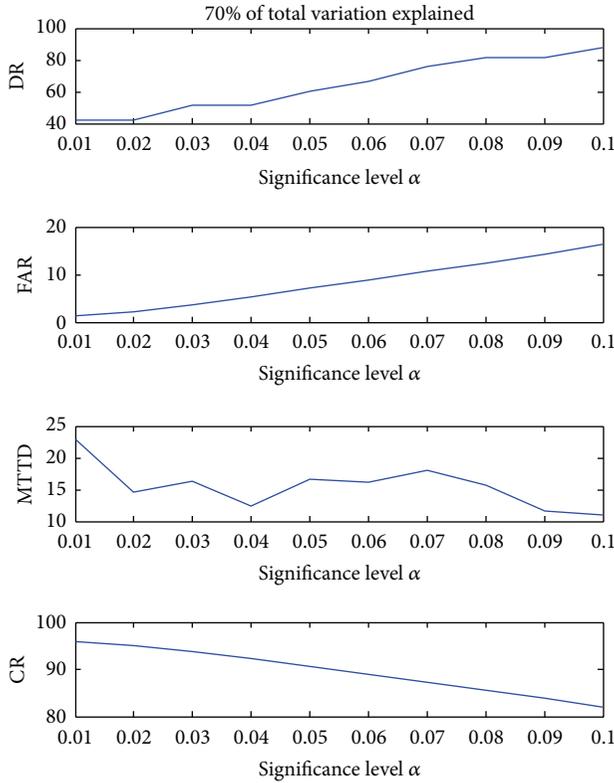


FIGURE 3: DR, FAR, MTTD and CR versus significance level.

TABLE 1: Comparison of the performance of classifiers built with different parameters.

Classifiers		DR	FAR	MTTD	CR
Parameters	Numbers				
(70%, 0.2)	(2, 1)	81.82	12.48	15.63	85.59
(70%, 0.3)	(2, 2)	72.73	12.04	12.21	85.97
(90%, 0.2)	(3, 1)	72.73	12.10	16.29	85.93
(90%, 0.3)	(3, 2)	63.64	11.65	11.76	86.32
Only major components		54.55	6.60	12.56	91.11

In this table, the best result in each column is shown in bold (not including the last row). The classifiers with the parameter pair (90%, 0.3) outperformed the other classifiers in terms of the FAR, MTTD, and CR values. However, the detection rate for the parameter pair (90%, 0.3) did not perform acceptably. The table shows that as the percentage of the variation explained increased, DR tended to decrease, while FAR and CR exhibited small fluctuations and performed at a similar level. Another observation is that a classifier yielded good MTTD values when more minor components were used by increasing the upper bound of the eigenvalues for the minor components.

For comparison, we also list the testing results achieved by a classifier that only used the major components in the last line of Table 1. The significance level was also set to 0.08 and the major components accounted for 70% of the total variation. Although this classifier achieved the best values

TABLE 2: Performance of PLSR built with a different proportion of incident instances.

Proportion	DR	FAR	MTTD	CR
35	12.12	0.90	5.00	96.49
40	36.36	4.22	15.00	93.42
45	69.70	13.07	11.04	85.12
50	93.94	32.50	7.03	66.72
55	100.00	53.80	2.58	46.69

for FAR and CR, the DR values were too low to be used in practice. A classifier using minor components can thus dramatically improve the detection rate of models, as has been proven by Shyu et al. [20].

4.3. *Comparison with PLS.* Wang et al. [3, 4] developed automatic incident detection (AID) models based on partial least squares regression, which were compared to the use of support vector machines for freeway incident detection. Here, we compare PCC and PLS for incident detection.

The training dataset had 102712 instances, with only 2751 incident instances compared to 99961 nonincident instances. The class distribution was therefore highly skewed as the frequency of the main class was more than 97%. The PLSR model is sensitive to imbalanced training data; that is, the proportion of incident samples in the training set strongly influences the detection performance [3]. To address this problem, we discarded random nonincident instances while retaining all the incident instances to increase the proportion of incident instances. In this way, we obtained a series of new training datasets with the proportion of incident instances ranging from 35% to 55% in 5% increments. Then, we built a PLSR model for each new set. Next, we tested the detection performance of the PLSR models with the testing dataset mentioned above.

Table 2 shows the testing results obtained with the PLSR model. Using the dataset containing 40% to 50% of incident instances to build the PLSR resulted in relatively good model performance. If there were too few incident instances, the PLSR model would produce DR values that would be too low; if there were too many incident instances, the PLSR model would produce FAR values that would be too high. Taking all these criteria into consideration, we chose the classifier modeled on the training set with 45% incident instances for further study. The testing results of this classifier are shown in bold.

Twelve classifiers were then constructed by allowing the significance level to range from 0.07 to 0.09, adjusting the percentage of the variation explained from 0.7 to 0.9, and increasing the upper bound of the eigenvalues for the minor components from 0.2 to 0.3. The detection performances of the 12 classifiers are presented in Table 3, along with the average detection performance; the PLSR performance is listed in the last row for comparison.

The results show the variation in classifier performance. Table 3 shows that the upper bound of the eigenvalues for the minor components strongly influenced the DR and MTTD values. To increase its value, the MTTD values decreased,

TABLE 3: Performance comparison between PCC and PLSR.

Significance	Classifiers		DR	FAR	MTTD	CR
	Explained	Bound				
0.7	70%	0.2	75.76	10.78	18.08	87.16
		0.3	69.70	10.75	12.65	87.17
	90%	0.2	69.70	10.40	18.57	87.53
		0.3	63.64	10.31	12.29	87.58
0.8	70%	0.2	81.82	12.48	15.63	85.59
		0.3	72.73	12.04	12.21	85.97
	90%	0.2	72.73	12.10	16.29	85.93
		0.3	63.64	11.65	11.76	86.32
0.9	70%	0.2	81.82	14.37	11.70	83.81
		0.3	72.73	13.42	11.92	84.68
	90%	0.2	75.76	13.93	11.64	84.22
		0.3	66.67	13.01	11.18	85.05
Average of PCC			75.76	13.93	11.64	84.22
PLSR			69.70	13.07	11.04	85.12

which was a desirable result. Unfortunately, the DR values also decreased as the MTTD values decreased. Comparing the PCC and PLSR model results, PCC appeared to exhibit poorer performance in terms of the FAR, MTTD, and CR values, but yielded higher average DR values. However, we should keep in mind that the highest performing PLSR model was compared with the PCC model results. If the average PLSR performance was compared to PCC performance, PLSR performance would be observed to be slightly inferior to PCC performance.

5. Conclusion

The detection of traffic incidents, congestion, and other traffic operational problems is a very important component of traffic system operation. The task of incident detection can be regarded as constructing classifiers from imbalanced or skewed datasets. In such problems, almost all the instances are labeled as a nonincident class, while far fewer instances are labeled as an incident class, which is usually the more important class. The learned classifier determines whether an incident occurs using traffic flow measurements, so that classifying the traffic state is any efficient means of using the class imbalance problem to solve the incident detection problem.

In this study, we determined the applicability of principal component analysis. PCA is often applied to reduce the dimensionality of a problem, as well as to detect outliers. In this paper, PCA was used to construct a simple classifier to detect incidents. This classifier consisted of two simple functions, the major components and the minor components. Only a few parameters, the significance level, the percentage of the total variation explained by the major components, and the upper bound of the eigenvalues for the minor components, needed to be retained for future detection. The influence of these parameters on detection performance was discussed in our experiments.

PCA performance was tested on a real dataset collected from the A12 freeway in The Netherlands. The PCC results were compared to the results of the standard linear PLSR method. The experimental results showed that PCC outperformed the PLSR model for AID.

Although the test results showed that PCC can achieve good incident detection performance, there is still room for improvement; the FAR values were too high and the MTTD values were too high with PCC. These results may be attributed to the poor quality of real traffic data. Real traffic data are always easily contaminated by a high noise level in the dataset, due to missing values, transcription errors, incomplete information, and the absence of standard formats. Learning from noisy data is a challenging and practical issue for real-world data mining applications. Common practices include data cleaning, error detection, and classifier ensembles [25, 26].

Refinements in data quality show promise in terms of improving incident detection performance. Note that common data cleaning methods are not suitable for AID data because these methods recognize incident instances as noise and consequently delete the incidents based on their rarity. Therefore, further research should focus on developing suitable algorithms for data cleaning.

Acknowledgments

The authors would like to express their gratitude to Dr. Sascha Hoogendoorn in DVS Center for Transport and Navigation, Ministry of Transport, Public Works and Water Management, The Netherlands, and Dr. Yusen Chen at the Delft University of Technology for their assistance in obtaining traffic data. This work was supported by the National Natural Science Foundation of China under Grant no. 61074141, the National High Technology Research and Development Program of China (863 Program) under Grant no. 2012AA112304, and the Natural Science Foundation of Jiangsu Province, China, under Grant no. BK2011745.

References

- [1] B. Hellinga and G. Knapp, "AVI based freeway incident detection," in *Proceedings of the TRB Annual Meeting*, pp. 1-26, 2000.
- [2] J. Black and I. Sreedevi, "Detection Algorithms," http://www.calccit.org/itsdecision/serv_and_tech/Incident_management/incident_management_overview.html.
- [3] W. Wang, S. Chen, and G. Qu, "Incident detection algorithm based on partial least squares regression," *Transportation Research Part C*, vol. 16, no. 1, pp. 54-70, 2008.
- [4] W. Wang, S. Chen, and G. Qu, "Comparison between partial least squares regression and support vector machine for freeway incident detection," in *Proceedings of the 10th International IEEE Conference on Intelligent Transportation Systems, ITSC 2007*, pp. 190-195, Seattle, Wash, USA, October 2007.
- [5] F. Yuan and R. L. Cheu, "Incident detection using support vector machines," *Transportation Research Part C*, vol. 11, no. 3-4, pp. 309-328, 2003.

- [6] R. L. Cheu, D. Srinivasan, and E. T. Teh, "Support vector machine models for freeway incident detection," in *Proceedings of Intelligent Transportation Systems*, vol. 1, pp. 238–243, 2003.
- [7] S.-Y. Chen, W. Wang, and G.-F. Qu, "Traffic incident detection based on rough sets approach," in *Proceedings of the 6th International Conference on Machine Learning and Cybernetics (ICMLC '07)*, pp. 3734–3739, August 2007.
- [8] S. Chen, W. Wang, G. Qu, and J. Lu, "Application of neural network ensembles to incident detection," in *Proceedings of the IEEE International Conference on Integration Technology (ICIT '07)*, pp. 388–393, Shenzhen, China, March 2007.
- [9] S. Chen, W. Wang, and H. van Zuylen, "Construct support vector machine ensemble to detect traffic incident," *Expert Systems with Applications*, vol. 36, no. 8, pp. 10976–10986, 2009.
- [10] S. Chen and W. Wang, "Decision tree learning for freeway automatic incident detection," *Expert Systems with Applications*, vol. 36, no. 2, pp. 4101–4105, 2009.
- [11] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Dimitris kanellopoulos, and panayiotis pintelas. Handling imbalanced datasets: a review," *GESTS International Transactions on Computer Science and Engineering*, no. 1, pp. 25–36, 2006.
- [12] N. Japkowicz and S. Stephen, "The class imbalance problem: a systematic study," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–450, 2002.
- [13] A. Estabrooks, T. Jo, and N. Japkowicz, "A multiple resampling method for learning from imbalanced data sets," *Computational Intelligence*, vol. 20, no. 1, pp. 18–36, 2004.
- [14] C. Elkan, "The foundations of cost-sensitive learning," in *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pp. 973–978, Seattle, Wash, USA, 2001.
- [15] B. Raskutti and A. Kowalczyk, "Extreme rebalancing for svms: a case study," *SIGKDD Explorations*, vol. 6, pp. 60–69, 2004.
- [16] R. E. Schapire, "A brief introduction to boosting," in *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, pp. 1401–1406, 1999.
- [17] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," in *Proceedings of the 13th International Conference on Machine Learning*, M. Kaufmann, Ed., pp. 148–156, San Francisco, Calif, USA, 1996.
- [18] S. S. Khan and M. G. Madden, "A survey of recent trends in one class classification," *Lecture Notes in Computer Science*, vol. 6206, pp. 188–197, 2010.
- [19] N. Japkowicz, "Concept-learning in the presence of between-class and within-class imbalances," in *Proceedings of the Fourteenth Conference of the Canadian Society for Computational Studies of Intelligence*, pp. 67–77, 2001.
- [20] M. -L. Shyu, S. -C. Chen, K. Sarinnapakorn et al., "A novel abnormal detection scheme based on principal component classifier," in *The IEEE International Conference on Data Mining series (ICDM ,03)*, 2003.
- [21] D. M. Hawkins, "The detection of errors in multivariate data using principal components," *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 340–344, 1974.
- [22] S. Turner, L. Albert, B. Gajewski, and W. Eisele, "Archived intelligent transportation system data quality: preliminary analyses of San Antonio TransGuide data," *Transportation Research Record*, no. 1719, pp. 77–84, 2000.
- [23] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, 2001.
- [24] S. Chen, W. Wang, and H. van Zuylen, "A comparison of outlier detection algorithms for ITS data," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1169–1178, 2010.
- [25] Y. Zhang, X. Zhu, X. Wu, and J. P. Bond, "ACE: an aggressive classifier ensemble with error detection, correction and cleansing," in *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '05)*, pp. 310–317, chn, November 2005.
- [26] V. Brusica, J. Zeleznikow, T. Sturniolo, E. Bono, and J. Hammer, "Data cleansing for computer models: a case study from immunology," in *Proceedings of 6th International Conference on Neural Information Processing (ICONIP '99)*, pp. 603–6609, 1999.

Research Article

State of Charge Estimation Based on Microscopic Driving Parameters for Electric Vehicle's Battery

Enjian Yao,^{1,2} Meiyang Wang,^{1,2} Yuanyuan Song,² and Yang Yang²

¹ MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China

² School of Traffic and Transportation, Beijing Jiaotong University, Haidian District, Beijing 100044, China

Correspondence should be addressed to Meiyang Wang; 12120914@bjtu.edu.cn

Received 27 September 2013; Revised 17 November 2013; Accepted 25 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Enjian Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, battery-powered electric vehicle (EV) has received wide attention due to less pollution during use, low noise, and high energy efficiency and is highly expected to improve urban air quality and then mitigate energy and environmental pressure. However, the widespread use of EV is still hindered by limited battery capacity and relatively short cruising range. This paper aims to propose a state of charge (SOC) estimation method for EV's battery necessary for route planning and dynamic route guidance, which can help EV drivers to search for the optimal energy-efficient routes and to reduce the risk of running out of electricity before arriving at the destination or charging station. Firstly, by analyzing the variation characteristics of power consumption rate with initial SOC and microscopic driving parameters (instantaneous speed and acceleration), a set of energy consumption rate models are established according to different operation modes. Then, the SOC estimation model is proposed based on the presented EV power consumption model. Finally, by comparing the estimated SOC with the measured SOC, the proposed SOC estimation method is proved to be highly accurate and effective, which can be well used in EV route planning and navigation systems.

1. Introduction

In recent years, energy and environment have suffered from heavy pressure caused by rapidly increased gasoline and diesel powered vehicles, and the growing concern about energy reserves and environmental quality of cities has called for sustainable transportation technologies and motivated active research on vehicles with alternative energy sources [1, 2]. The electric vehicles (EVs), which run on electricity, have received wide attention due to less pollution during use, low noise, and high energy efficiency [3, 4].

Although EV has considerable advantages, the popularization of EV is still hindered by limited battery capacity and relatively short cruising range. Up to now, the battery management system (BMS) has been developed as one of the EV's key technologies, and the accurate estimation of SOC of battery is viewed as a critical part of BMS. Since the battery is a strong nonlinear and time-variability system for its complicated electrochemical process [5, 6], the SOC is affected

by many factors such as open-circuit voltage, self-rescue effect, temperature, charge and discharge efficiency, and circle life. Although a great many SOC estimation methods have been explored by researchers, the accurate estimation of SOC is very difficult and complex due to limited battery models and parametric uncertainties [7].

According to different methodologies, the SOC estimation methods can be classified into four categories, including direct measurement, book-keeping estimation, adaptive system, and hybrid methods [8]. The direct measurement methods refer to measuring the SOC based on the battery properties (battery voltage, battery impedance, etc.), which mainly include open-circuit voltage (OCV) method [9, 10], terminal voltage method [11], impedance measurement method [12, 13], and impedance spectroscopy method [14, 15]. The book-keeping is a method based on current measurement and integration, which uses battery discharging current data and other relevant data (self-discharge rate of battery, temperature, charge/discharge efficiency, etc.) as inputs [16]. The two

main book-keeping estimation methods are Coulomb counting method and modified Coulomb counting method [17]. With the development of artificial intelligence, various new adaptive systems for SOC estimation have been explored, which mainly include back propagation (BP) neural network [18, 19], radial basis function (RBF) neural network [20], fuzzy logic methods [21, 22], support vector machine [23], fuzzy neural network [24], and Kalman filter [25, 26]. In addition, the hybrid methods are explored by some researchers, which take advantage of multiple SOC estimation methods to improve the estimation accuracy [27].

Although extensive research has been undertaken in exploring approaches to estimate the SOC of battery based on the complex external characteristic of battery as mentioned above, few papers have been devoted to studying the relationship between the SOC of EV's battery and microscopic driving parameters. This means that the SOC at some specific points, which is necessary for EV route planning, cannot be predicted before the travel using the current estimation methods. Therefore, in an attempt to overcome the limitations of current SOC estimation methods for EV's battery, this paper proposes a SOC estimation approach based on the forecasted microscopic driving parameters, which can predict the power consumption and SOC accurately before departure and then reduce the risk of running out of electricity before arriving at the destination or charging station.

This paper is organized into four sections. In the first section, the background and significance on SOC estimation models are provided, along with a review of existing studies on SOC estimation. In the second section, by analyzing the impacts of initial SOC for each second and microscopic driving parameters on power consumption rate, the EV power consumption rate models for different operation modes are established. On that basis, the SOC estimation model is proposed, which uses the initial SOC, instantaneous speed, and acceleration as input variables. In the third section, the model parameters are calibrated and the estimation results are discussed by comparing the measured SOC with estimated SOC. The conclusions and future work are then provided in the last section.

2. Modeling

Since the SOC of battery is decided by the initial SOC and power consumption during the discharge time, in order to establish the relationship between SOC and microscopic driving parameters, it is important to study the impact of microscopic driving parameters on power consumption of EV. Therefore, in this paper, the EV power consumption rate models for different operation modes are set up firstly, which give a full consideration to the influence of the initial SOC for each second, instantaneous speed, and acceleration on the power consumption rate. Then, the SOC estimation model is proposed based on the presented EV power consumption model.

2.1. Data Source. The data used for establishing and verifying the power consumption rate models and SOC estimation

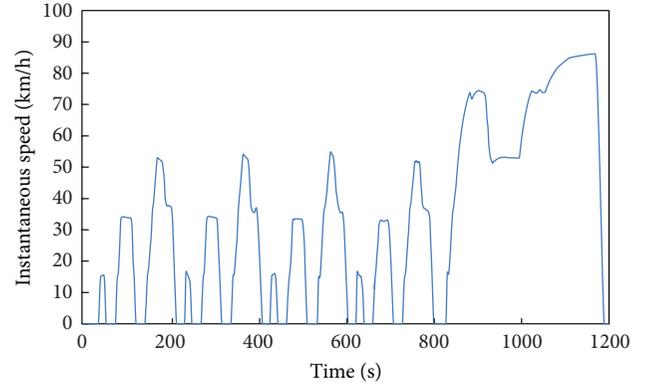


FIGURE 1: New European Driving Cycle.

TABLE 1: Correlation coefficient.

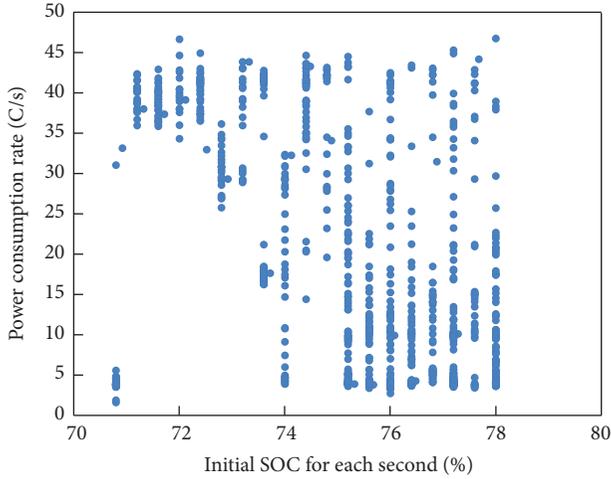
Correlation coefficient	Parameters		
	Initial SOC for each second (%)	Speed (km/h)	Acceleration (m/s^2)
Power consumption rate	-0.136	0.765	0.410

model are collected by a chassis dynamometer test with New European Driving Cycle (NEDC). NEDC is a driving cycle consisting of four repeated ECE-15 driving cycles and an Extra-Urban Driving Cycle (EUDC), as shown in Figure 1, and NEDC is supposed to represent the typical usage of a car in China and Europe, and designed to assess the emission levels of car engines and fuel economy in passenger cars (excluding light trucks and commercial vehicles). The data include four NEDCs and contain time, vehicle speed, battery working current, battery voltage, and so forth. In the paper, the data of the first three NEDCs are used to set up the EV power consumption rate models and SOC estimation model, and the data of the rest one driving cycle are used to verify the accuracy of the proposed models.

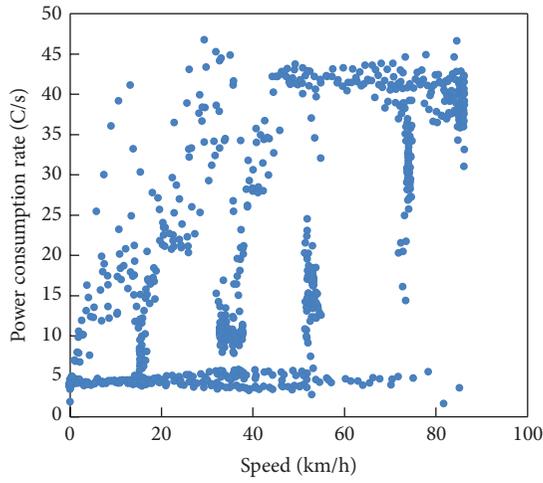
2.2. Model Methodology. Both the initial SOC for each second and microscopic driving parameters have a significant impact on the power consumption rate. Therefore, in this section, by analyzing the correlation between the power consumption rate and initial SOC for each second and microscopic driving parameters, the EV power consumption rate models are built up according to different operation modes. On that basis, the SOC estimation model is established.

2.2.1. Modeling Power Consumption Rate for EVs. The correlation coefficients between the EV power consumption rate and initial SOC for each second, instantaneous speed, and acceleration are summarized in Table 1. It can be concluded that there is a relatively strong correlation between power consumption rate and these parameters.

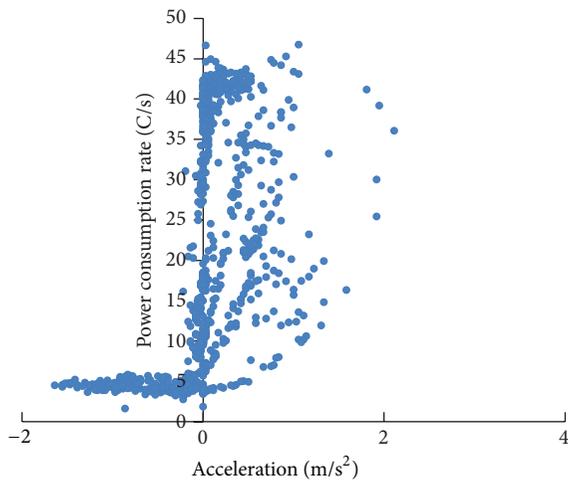
Further, the variation characteristics of power consumption rate with these parameters are analyzed (shown in Figure 2). According to Figure 2(a), the change rules of power consumption rate vary in different initial SOC for



(a)



(b)



(c)

FIGURE 2: Variation characteristics of power consumption rate with initial SOC for each second and microscopic driving parameters.

each second. Figures 2(b) and 2(c) illustrate graphically the nonlinear relationship between power consumption rate and microscopic driving parameters. It is clear that power consumption rate generally tends to increase with the increase of speed and acceleration. In addition, the gradient of the power consumption rate for decelerating mode is generally smaller than that for accelerating mode, in which it should be noted that the power consumption rate varies in different operation modes. Therefore, in this paper, the running status of EVs is divided into four operation modes, including accelerating, decelerating, cruising, and idling, and then four power consumption rate models are established, respectively.

Based on the analysis above, this paper attempts to establish the EV power consumption rate models for different operation modes that require initial SOC for each second, instantaneous speed and acceleration as independent variables.

The derivation of the model structure involves experimentation with numerous polynomial combinations of the initial SOC for each second, speed, and acceleration. Based on the VT-Micro model of fuel consumption and emission for gasoline and diesel powered vehicles [28, 29], the final regression model includes the linear initial SOC for each second coupled with a combination of linear, quadratic, and cubic speed and acceleration terms because it indicates a relatively high goodness of fit for the collected data as (1). Since the above model may produce negative dependent variable values in a few instances, a data transformation method [29] is applied to the model presented in (1), and the exponential function model presented in (2) is obtained. Furthermore, the power consumption rate for idling mode is estimated using the average value of the power consumption of every second:

$$q = \begin{cases} a_0 \times S + \sum_{i=0}^3 \sum_{j=0}^3 (a_{i,j} \times v^i \times a^j), & a > 0 \\ b_0 \times S + \sum_{i=0}^3 \sum_{j=0}^3 (b_{i,j} \times v^i \times a^j), & a < 0 \\ c_0 \times S + \sum_{i=0}^3 (c_i \times v^i), & a = 0, v \neq 0 \\ \bar{q}, & a = 0, v = 0, \end{cases} \quad (1)$$

$$q = \begin{cases} \exp \left(l_0 \times S + \sum_{i=0}^3 \sum_{j=0}^3 (l_{i,j} \times v^i \times a^j) \right) & a > 0 \\ \exp \left(m_0 \times S + \sum_{i=0}^3 \sum_{j=0}^3 (m_{i,j} \times v^i \times a^j) \right) & a < 0 \\ \exp \left(n_0 \times S + \sum_{i=0}^3 (n_i \times v^i) \right) & a = 0, v \neq 0 \\ \bar{q} & a = 0, v = 0, \end{cases} \quad (2)$$

where q is the power consumption rate, namely, the power consumption of each second, C/s; S is the initial SOC for each second, %; $(a_0, a_{i,j})$, $(b_0, b_{i,j})$, and (c_0, c_i) are the coefficients of accelerating, decelerating, and cruising, respectively,

TABLE 2: Results of parameters calibration for power consumption rate models.

Variable	Accelerating coefficients (<i>t</i> -value)	Decelerating coefficients (<i>t</i> -value)	Cruising coefficients (<i>t</i> -value)	Idling coefficients
Constant	1.862 (14.816)	1.098 (18.314)	0.602 (2.704)	3.62
<i>S</i>	0.226 (2.132)	0.317 (4.743)	0.633 (3.002)	—
<i>v</i>	—	0.025 (42.068)	0.037 (10.492)	—
<i>v</i> ²	—	—	—	—
<i>v</i> ³	—	1.184E – 006 (12.008)	–1.062E – 006 (–3.314)	—
<i>a</i>	—	—	—	—
<i>a</i> ²	—	—	—	—
<i>a</i> ³	—	—	—	—
<i>va</i>	0.098 (23.0313)	0.115 (49.117)	—	—
<i>va</i> ²	—	0.152 (35.344)	—	—
<i>va</i> ³	–0.030 (–6.753)	0.053 (28.475)	—	—
<i>v</i> ² <i>a</i>	—	—	—	—
<i>v</i> ² <i>a</i> ²	—	—	—	—
<i>v</i> ² <i>a</i> ³	—	—	—	—
<i>v</i> ³ <i>a</i>	—	—	—	—
<i>v</i> ³ <i>a</i> ²	–5.054E – 005 (–13.667)	–2.324E – 006 (–13.310)	—	—
<i>v</i> ³ <i>a</i> ³	2.965E – 005 (12.637)	—	—	—
Adjusted <i>R</i> ²	0.947	0.894	0.927	—

v is the instantaneous vehicle speed, km/h; *a* is the acceleration, m/s²; *S* is the initial SOC for each second, %; *R*² is the coefficient of determination.

for polynomial model, $i, j = 0, 1, 2, 3$; $(l_0, l_{i,j})$, $(m_0, m_{i,j})$, and (n_0, n_i) are the coefficients of accelerating, decelerating, and cruising respectively, for exponential model, $i, j = 0, 1, 2, 3$; v is the instantaneous speed of EV, km/h; a is the instantaneous acceleration, m/s²; \bar{q} is the average power consumption of every second for idling, C/s.

2.2.2. Estimating SOC for EVs. Based on the power consumption rate models proposed above, the power consumption during the discharge time can be obtained by the time integral as

$$Q = \sum_{i=1}^n q_i \times t, \quad (3)$$

where Q is the power consumption during the discharge time, C; n is the discharge time, s; q_i is the power consumption of each second, namely, the power consumption rate, C/s; t is the time interval, 1 s.

Thus, the SOC estimation model based on the initial SOC and microscopic driving parameters can be calculated by

$$S_{n+1} = f(S_0, v, a) = S_0 - \frac{Q}{3600 \times Q_r} = S_0 - \frac{\sum_{i=0}^n q_i \times t}{3600 \times Q_r}, \quad (4)$$

where S_{n+1} is state of charge after the discharge time n , %; S_0 is the initial SOC, %; Q_r is the nominal capacity of battery given by the manufacturer, Ah.

3. Results and Analysis

In this section, using SPSS software, the parameters of EV power consumption rate models for different operation

modes are calibrated, respectively, based on regression analysis method. Further, the estimated power consumption and SOC are compared with the measured power consumption and SOC to verify the accuracy of the proposed EV power consumption rate estimation models and the SOC estimation model.

3.1. Parameter Calibration. Based on the data of the first three driving cycles mentioned above, using regression analysis method, the parameters for accelerating, decelerating, cruising, and idling are calibrated, respectively. Since the use of polynomial speed and acceleration terms may result in multicollinearity between the independent variables due to the dependency of the variables, stepwise regression analysis method is utilized in the parameters calibration. However, the study shows that heteroskedasticity exists in the accelerating model calculated with ordinary least square (OLS). Therefore, for accelerating mode, weighted least square (WLS) is used for model calibration. The results of parameters calibration for power consumption rate models are shown in Table 2.

According to Table 2, the statistical results indicate a relatively high goodness of fit for power consumption rate estimation (the adjusted R^2 values in excess of 0.89 and the absolute t -values of the variables in excess of 1.96 for all the operation modes).

3.2. Estimation Results. On the one hand, in order to evaluate the accuracy of the proposed power consumption rate estimation models, the measured power consumption rate data are compared with the regression models estimation through

TABLE 3: Results of MAPE for different operation modes.

Operation modes	Accelerating	Decelerating	Cruising	Idling
MAPE	8.74%	9.92%	9.45%	9.67%

calculating the mean absolute percentage error (MAPE). MAPE is calculated by

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Q_i - Q'_i}{Q_i} \right| \times 100\%, \quad (5)$$

where Q_i is the measured power consumption, C ; Q'_i is the estimated power consumption using the proposed models, C ; n is the number of tested data samples.

With the purpose of ensuring consistency in comparison, in this paper, the data of the rest one driving cycle mentioned above are utilized to verify the accuracy of proposed models. The MAPEs for second-by-second power consumption illustrate that the models are relatively accurate with MAPE ranging from 8.74% to 9.92% (shown in Table 3).

Further, in order to evaluate the applicability of the proposed models in EV route planning and navigation systems, the MAPE of a specific route is also taken as evaluation indicator. The rest one driving cycle is divided into different time intervals (2 min, 4 min, 6 min, 8 min, 10 min, 12 min, 14 min, 16 min, 18 min, and 20 min) and each time interval is regarded as a specific planned route of EV. The power consumption for each time interval is estimated using the proposed models, and then the MAPEs for the planned route are calculated under different time intervals (shown in Figure 3). It is clear that the MAPEs for different time intervals are all less than 4.50% and the MAPE generally tends to decrease with the increase of length of time interval. Therefore, for the data of NEDC used in this paper, the presented power consumption rate models for different operation modes can estimate the EV power consumption with sufficient accuracy.

On the other hand, the accuracy of the presented SOC estimation model is analyzed and discussed. Due to the limited battery capacity and relatively short cruising range, it is necessary to obtain several specific SOC values of EV's battery during the route planning. Therefore, the several specific SOC values are calculated using the above SOC estimation model. To evaluate the accuracy and applicability of the proposed model in EV route planning and navigation systems, the estimated SOC is compared with the measured SOC (shown in Figure 4). The relative errors for different time points between the measured SOC and estimated SOC are all quite small. Therefore, for the data of NEDC used in the paper, the presented model is appropriate for estimating the SOC with sufficient accuracy and can be well applied in EV route planning and navigation systems.

4. Conclusions

The paper presents a novel SOC estimation approach based on the data collected by a chassis dynamometer test with NEDC, which is characterized by fully considering the impacts of microscopic driving parameters on SOC. Firstly,

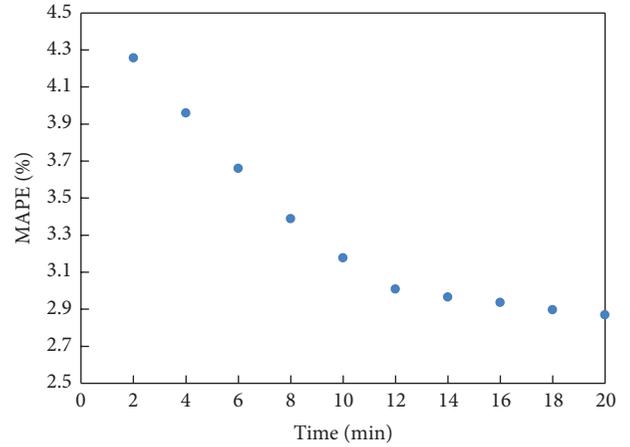


FIGURE 3: MAPEs of power consumption rate model under different time intervals.

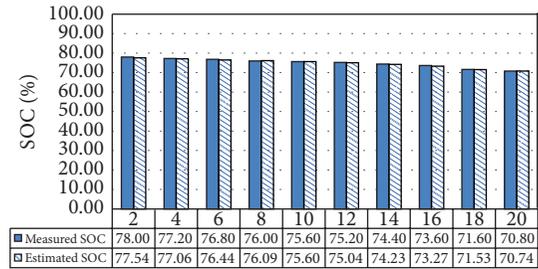


FIGURE 4: Comparison between measured SOC and estimated SOC.

through the variation characteristics analysis of power consumption rate with initial SOC for each second, instantaneous speed, and acceleration, a set of EV power consumption rate models with exponential function are established for different operation modes. Further, using the presented EV power consumption rate models, the SOC estimation model that requires the initial SOC and microscopic driving parameters as input variables is set up. Then, the parameters of EV power consumption rate models for different operation modes are calibrated, respectively, using the regression analysis method. For the data of NEDC used in the paper, the power consumption rate models for different operation modes are relatively accurate with MAPE ranging from 8.74% to 9.92%, and the SOC estimation model can be well applied in EV route planning and navigation systems to estimate the SOC of battery. In the future practical application, the models presented in this paper can be utilized in conjunction with microscopic traffic simulation technologies or GPS to further demonstrate their application to routing planning and dynamic route guidance for EV.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the National 973 Program of China (no. 2012CB725403) and the Research Fund for the Doctoral Program of Higher Education of China (no. 20130009110002).

References

- [1] J. A. Calvo, C. Álvarez-Caldas, J. L. San Román, and P. Cobo, "Influence of vehicle driving parameters on the noise caused by passenger cars in urban traffic," *Transportation Research D*, vol. 17, no. 7, pp. 509–513, 2012.
- [2] H. S. Hamut, I. Dincer, and G. F. Naterer, "Performance assessment of thermal management systems for electric and hybrid electric vehicles," *International Journal of Energy Research*, vol. 37, no. 1, pp. 1–12, 2013.
- [3] L. He, W. Chen, and G. Conzelmann, "Impact of vehicle usage on consumer choice of hybrid electric vehicles," *Transportation Research D*, vol. 17, no. 3, pp. 208–214, 2012.
- [4] Z. P. Wang, P. Liu, J. Cui, Y. Xi, and L. Zhang, "Research on quantitative models of electric vehicle charging stations based on principle of energy equivalence," *Mathematical Problems in Engineering*, vol. 2013, Article ID 959065, 10 pages, 2013.
- [5] H. He, R. Xiong, and J. Fan, "Evaluation of lithium-ion battery equivalent circuit models for state of charge estimation by an experimental approach," *Energies*, vol. 4, no. 4, pp. 582–598, 2011.
- [6] R. T. Doucette and M. D. McCulloch, "Modeling the prospects of plug-in hybrid electric vehicles to reduce CO₂ emissions," *Applied Energy*, vol. 88, no. 7, pp. 2315–2323, 2011.
- [7] A. A. A. Elgammal and A. M. Sharaf, "Self-regulating ppaper swarm optimized controller for (photovoltaic-fuel cell) battery charging of hybrid electric vehicles," *IET Electrical Systems in Transportation*, vol. 2, no. 2, pp. 77–89, 2012.
- [8] W. Y. Chang, "The state of charge estimating methods for battery: a review," *ISRN Applied Mathematics*, vol. 2013, Article ID 953792, 7 pages, 2013.
- [9] S. Lee, J. Kim, J. Lee, and B. H. Cho, "State-of-charge and capacity estimation of lithium-ion battery using a new open-circuit voltage versus state-of-charge," *Journal of Power Sources*, vol. 185, no. 2, pp. 1367–1373, 2008.
- [10] S. Abu-Sharkh and D. Doerffel, "Rapid test and non-linear model characterisation of solid-state lithium-ion batteries," *Journal of Power Sources*, vol. 130, no. 1-2, pp. 266–274, 2004.
- [11] S. Sato and A. Kawamura, "A new estimation method of state of charge using terminal voltage and internal resistance for lead acid battery," in *Proceedings of the Power Conversion Conference*, pp. 565–570, Osaka, Japan, April 2002.
- [12] S. Rodrigues, N. Munichandraiah, and A. K. Shukla, "Review of state-of-charge indication of batteries by means of A.C. impedance measurements," *Journal of Power Sources*, vol. 87, no. 1-2, pp. 12–20, 2000.
- [13] F. Huet, "A review of impedance measurements for determination of the state-of-charge or state-of-health of secondary batteries," *Journal of Power Sources*, vol. 70, no. 1, pp. 59–69, 1998.
- [14] R. Li, J. Wu, H. Wang, and G. Li, "Prediction of state of charge of lithium-ion rechargeable battery with electrochemical impedance spectroscopy theory," in *Proceedings of the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA '10)*, pp. 684–688, Taichung, Taiwan, June 2010.
- [15] K. Bundy, M. Karlsson, G. Lindbergh, and A. Lundqvist, "An electrochemical impedance spectroscopy method for prediction of the state of charge of a nickel-metal hydride battery at open circuit and during discharge," *Journal of Power Sources*, vol. 72, no. 2, pp. 118–125, 1998.
- [16] H. J. Bergveld, W. S. Kruijt, and P. H. L. Notten, *Battery Management Systems, Design By Modeling*, vol. 1 of *Philips research book series*, Kluwer Academic, Boston, Mass, USA, 2002.
- [17] K. S. Ng, C.-S. Moo, Y.-P. Chen, and Y.-C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Applied Energy*, vol. 86, no. 9, pp. 1506–1511, 2009.
- [18] W. Y. Chang, "State of charge estimation for LiFePO₄ battery using artificial neural network," *International Review of Electrical Engineering*, vol. 7, no. 5, pp. 5874–5800, 2012.
- [19] T. Weigert, Q. Tian, and K. Lian, "State-of-charge prediction of batteries and battery-supercapacitor hybrids using artificial neural networks," *Journal of Power Sources*, vol. 196, no. 8, pp. 4061–4066, 2011.
- [20] G. Hongyu, J. Jiuchun, and W. Zhanguo, "Estimating the state of charge for Ni-MH battery in HEV by RBF neural network," in *Proceedings of the International Workshop on Intelligent Systems and Applications (ISA '09)*, pp. 1–4, Wuhan, China, May 2009.
- [21] P. Singh, R. Vinjamuri, X. Wang, and D. Reisner, "Design and implementation of a fuzzy logic-based state-of-charge meter for Li-ion batteries used in portable defibrillators," *Journal of Power Sources*, vol. 162, no. 2, pp. 829–836, 2006.
- [22] S. Malkhandi, "Fuzzy logic-based learning system and estimation of state-of-charge of lead-acid battery," *Engineering Applications of Artificial Intelligence*, vol. 19, no. 5, pp. 479–485, 2006.
- [23] T. Hansen and C.-J. Wang, "Support vector based battery state of charge estimator," *Journal of Power Sources*, vol. 141, no. 2, pp. 351–358, 2005.
- [24] Y.-S. Lee, W.-Y. Wang, and T.-Y. Kuo, "Soft computing for battery state-of-charge (BSOC) estimation in battery string systems," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 1, pp. 229–239, 2008.
- [25] L. Xu, J. Wang, and Q. Chen, "Kalman filtering state of charge estimation for battery management system based on a stochastic fuzzy neural network battery model," *Energy Conversion and Management*, vol. 53, no. 1, pp. 33–39, 2012.
- [26] F. Sun, X. Hu, Y. Zou, and S. Li, "Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles," *Energy*, vol. 36, no. 5, pp. 3531–3540, 2011.
- [27] J. Wang, B. Cao, Q. Chen, and F. Wang, "Combined state of charge estimator for electric vehicle battery pack," *Control Engineering Practice*, vol. 15, no. 12, pp. 1569–1576, 2007.
- [28] K. Ahn, A. A. Trani, and H. Rakha, "Microscopic fuel consumption and emission modeling," in *Proceedings of the 78th Annual Meeting of the Transportation Research Board (CD-ROM)*, Washington, DC, USA, 1999.
- [29] K. Ahn, H. Rakha, A. Trani, and M. van Aerde, "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels," *Journal of Transportation Engineering*, vol. 128, no. 2, pp. 182–190, 2002.

Research Article

Optimizing Schedules of Rail Train Circulations by Tabu Search Algorithm

Mingming Chen and Huimin Niu

School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

Correspondence should be addressed to Huimin Niu; hmniu@mail.lzjtu.cn

Received 21 October 2013; Revised 16 November 2013; Accepted 25 November 2013

Academic Editor: Wuhong Wang

Copyright © 2013 M. Chen and H. Niu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper develops an integer programming model for the scheduling problem in train circulations on an intercity rail line. The model that aims to minimize the sum of interval time for any two consecutive tasks is proposed to characterize the train operation process. Two main constraints, namely, time-shift and equilibrium constraint, are considered to get the feasible and practical solution of train schedules. A heuristic procedure using tabu search algorithm is also designed to solve the model by introducing the penalty function and a neighborhood search method with the trip exchange and insert strategy. A computational experiment performed on test instances provided by two major stations on the Beijing–Tianjin Intercity Railway in China illustrates the proposed model and algorithm.

1. Introduction

Transit scheduling problem is a major area in operations research because of the complexity of problems that arise from various transit modes, such as airlines, railways, maritime, and urban transit. Vehicle scheduling and crew scheduling are two main problems that arise in this area. Generally, these two problems are considered separately, where the first is the vehicle scheduling problem and the second is the crew scheduling problem [1–3].

Recently, most scholars have focused on the two problems simultaneously. In [4], a single depot case with a homogeneous fleet of vehicles was considered and an exact approach was proposed to solve the simultaneous vehicle and crew scheduling problem in urban mass transit systems. An integrated approach to solve a vehicle scheduling problem and a crew scheduling problem on a single bus route was presented in [5]. An integrated vehicle and crew scheduling problem was described using an integer linear programming formulation combining a multicommodity network flow model with a set partitioning/covering model in [6]. An approach was presented to solve the bus crew scheduling problem that considers early, day, and late duty modes with time-shift and work intensity constraints in [7]. The authors in [8] proposed an integrated vehicle-crew-roster model with

days-off pattern, which aimed to simultaneously determine minimum cost vehicle and daily crew schedules.

In the field of rail transit, train and crew scheduling problems are key steps in the rail operational process. The train scheduling problem involves assigning trains to a set of trips generated by a train timetable. The crew scheduling problem involves assigning crews to trains that operate at a given schedule. The authors in [9] proposed a phase-regular scheduling method and applied a regular train-departing interval and the same train length for each period under the period-dependent demand conditions. In [10], a binary integer programming model incorporated with passenger loading and departure events was built to optimize the passenger train timetable in a heavily congested urban rail corridor. The authors in [11] established an optimization model based on maximum passenger satisfaction for train operations at the junction station in passenger dedicated lines. The authors in [12] described research in progress that would determine the minimum circulation of trains needed to execute a given timetable with given bounds on demand and capacities.

Train and crew scheduling is an NP hard problem. Generally, the difficulties stem from a large set of complex and conflicting restrictions that must be satisfied by any solution. Most of these restrictions are reflected in a sizable number

of operational conditions that involve trips in daily train timetables, train numbers, train capacities, crew numbers, and certain constraints related to time-shift, equilibrium, and work intensity. The authors in [13] proposed an algorithm which was based on local optimality criteria in the event of a potential crossing conflict to solve the train scheduling problem. A model designed to optimize train schedules on single line rail corridors was described in [14]. In [15], a multiobjective optimization model was developed for the passenger train-scheduling problem on a railroad network which included single and multiple tracks, as well as multiple platforms with different train capacities. To minimize shortages in capacity during rush hours, the authors in [16] described a model that could be used to find an optimal allocation of train types and subtypes for lines.

Meanwhile, various optimization models that relate to many aspects of train and crew schedules in railways are being studied extensively. The column generation approach is an effective algorithm for solving these problems. For example, the authors in [17] developed a column generation approach for a rail crew rescheduling problem. The authors in [18] presented a column generation approach based on the decomposition algorithm, which would achieve high-quality solutions at reasonable runtimes.

In recent years, a number of studies have paid more attention to developing a heuristic algorithm for the train scheduling problem. An algorithm that combined a compact assign and a matrix, as well as an operational time strategy was proposed in [19]. The authors in [20] developed two-solution approaches based on a space-time network representation that would operate a predetermined set of train duties to satisfy the strict day-off requirement for crew in railways. On the premise of unfixed train used sections, the authors in [21] built an optimized train operation and maintenance planning model using an algorithm with a penalty function and a 3-opt neighborhood structure to solve the model. A particle swarm optimization algorithm with a local search heuristic was presented to solve the crew scheduling problem in [22].

For an overview of the above papers, most studies on constructing the train scheduling model have been paid to the factors associated with train numbers, train capacities, interval time of trains, and so on. Moreover, interval time often includes night waiting time, which allows a train to conduct the trip tasks for the following day when the shortest layover time is less than the interval time of two consecutive trips. In fact, most trains run on the intercity line with high frequency, which is highly similar to bus transit. The origin and destination stations are generally equipped with train bases. In this paper, the factor of night waiting time is neglected, and the train scheduling problem is merely based on the train timetables in one day.

This paper is organized as follows. In Section 2, the rail train scheduling problem and an optimization model that minimizes the total interval time cost are described. A tabu search algorithm is presented in Section 3. In Section 4, a numerical example is provided to illustrate the application of the model and algorithm. The last section draws conclusions and discusses future research directions.

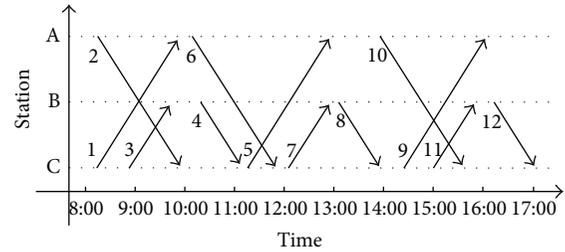


FIGURE 1: A simple train timetable.

2. Rail Train Scheduling Model

2.1. Problem Description. This paper considers train scheduling on a bidirectional intercity railway line with several stations. The location and number of trains available in each station are known. Every day, the trains are arranged from the designated stations to perform a set of trips. For each trip, the departure and arrival times and locations, which are determined by train timetables, are also known. Figure 1 shows a simple train timetable with three stations and 10 trips.

Train scheduling aims to assign a number of timetabled trips to a set of trains with the objective of minimizing total train operation costs and satisfying a range of constraints, including labor union agreements, government regulations, and company policy. For clarity, the following assumptions have to be considered.

- (1) Any two consecutive trips arranged for a train should have compatible terminals. Deadhead trips are not considered in this paper. Thus, the arrival location of a trip should be similar to the departure location of the next trip (e.g., Trips 1 and 6).
- (2) Any two consecutive trips arranged for a train should be compatible in time. A lower bound, called layover time, is present when a train arrives at a terminus. During this time, the trains wait for passengers to alight, board, turnaround, and so on. If the interval time of two trips cannot exceed the shortest layover time, the two trips for any train is voided. If we assume that the shortest layover time is 15 minutes, then we can say that Trips 4 and 5 cannot be compatible.
- (3) The train scheduling problem is solved as a daily problem in which every train schedule is assumed to be obtained from a daily train timetable. For two consecutive trips, the departure time of one trip should not be earlier than the arrival time of the next trip (e.g., Trips 1 and 2).
- (4) The number of required trains cannot exceed the prescribed maximum number of trains.

Figure 2 illustrates train schedules corresponding to the data from the train timetable in Figure 1. Three trains are arranged for 12 trips. Each row corresponds to the trip tasks of a train. For instance, the trip tasks of the first train (no. 1) are in a chain with C-A, A-C, C-B, B-C, and C-A. The columns correspond to the trips in a timetable. Notably, the number of columns that correspond to different trains differs from one

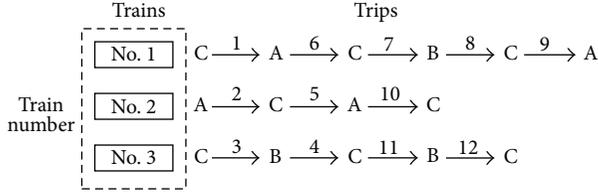


FIGURE 2: Illustration of train schedules.

another. For any trip task of a particular train, the departure station of the first trip and the arrival station of the last trip are not necessarily the same.

2.2. Objective Function. The train is an important piece of equipment used in intercity railways, which is expensive to produce or purchase. To complete train timetables, the required number of trains should be as few as possible. Furthermore, the interval time of trains, a key factor in measuring their circulation efficiency, can be described as follows: a train runs in a section and reaches an intermediate station in one trip. It operates and waits at the station and then departs from the station for the following trip task. The interval time includes the operation time and waiting time at the station. When the train arrives and departs from the terminus, the interval time can be called the turn-back time. Therefore, the optimal objective of rail train scheduling is to determine the minimum number of trains required and ensure the minimum cost of interval time.

As the sum of interval time and running time of trains is a multiple of the number of trains, and the running time of each train is a constant number, the objective of having the smallest number of trains is equivalent to the objective of having minimal cost in interval time. In this paper, the optimal objective is to minimize the total interval time, which can be formulated as follows:

$$\min \sum_{k=1}^m \sum_{i=1}^n \sum_{j=1}^n c_{ij} \cdot y_{ij}^k, \quad (1)$$

where m represents the number of trains required in a day, n represents the number of trips provided by a train timetable in a day and c_{ij} is the interval time between trip i and trip j . y_{ij}^k is a binary 0-1 variable that indicates the status of trip i and trip j conducted by train k . y_{ij}^k has two values: $y_{ij}^k = 1$ if trip j is the next trip after trip i conducted by train k , and $y_{ij}^k = 0$ otherwise.

The interval time of any two trips in train timetables is significantly influenced by factors related to the trip, such as origin station, destination station, and arrival and departure time at the two stations. For any two trips, if the destination station of the former trip is different from the origin station of the latter trip, or if the interval time is less than the shortest layover time T_0 , then the two trips cannot be arranged using the same train. Obviously, the two trips cannot satisfy the time-shift constraint. In this paper, the interval time can be set to an infinite positive number M . If the interval time of the two trips meets the time-shift constraint, then it can be

calculated using the formula $d_j - a_i$. Therefore, interval time c_{ij} between trips i and j can be expressed as follows:

$$c_{ij} = \begin{cases} d_j - a_i, & z_i = s_j, d_j - a_i \geq T_0, \\ M, & z_i \neq s_j \text{ or } z_i = s_j, d_j - a_i < T_0, \end{cases} \quad (2)$$

where z_i is the destination station of trip i , s_j is the origin station of trip j , a_i is the arrival time of trip i at destination station, and d_j is the departure time of trip j at origin station.

2.3. Constraints

- (1) To ensure that each trip can be conducted using only one train, we formulate

$$\sum_{k=1}^m x_i^k = 1, \quad \forall i. \quad (3)$$

- (2) The relationship between decision variable x_i^k and auxiliary variable y_{ij}^k can be formulated using the following equation:

$$y_{ij}^k = x_i^k \cdot x_j^k, \quad \forall k. \quad (4)$$

- (3) The two consecutive trips taken by the same train should satisfy the time-shift constraint, which means that the arrival time when the former trip ends at the destination station must be earlier than the departure time when the latter trip starts at the origin station. The minimum difference should be not less than T_0 ,

$$I(d_j - a_i - T_0) \geq y_{ij}^k, \quad \forall k, \quad (5)$$

where $I(x)$ represents sign function, which is calculated using (6),

$$I(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (6)$$

- (4) To ensure that the running mileage between different trains is equal, we have to make sure that the actual operation time between different trains does not vary significantly because of the constant speed of each train. The parameter G_k , which represents the total operation time of train k , can be calculated using the following equation:

$$G_k = \sum_{i=1}^n x_i^k (a_i - d_i). \quad (7)$$

Therefore, we can calculate the maximum and minimum operation times of trains, and the difference should not exceed the maximum value of the operation time between any two trains T_1 ,

$$\max_{k,k'} |G_k - G_{k'}| \leq T_1. \quad (8)$$

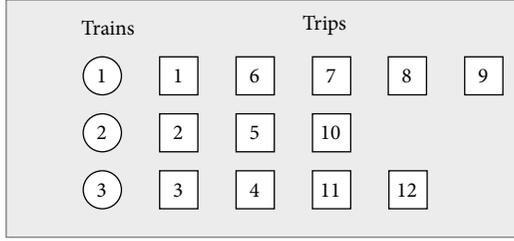


FIGURE 3: Expression of solution.

- (5) The decision variable x_i^k is a 0-1 variable to indicate whether trip i is conducted by train k . x_i^k has two values: $x_i^k = 1$ if trip i is conducted by train k , and $x_i^k = 0$ otherwise. Thus,

$$x_i^k \in \{0, 1\}, \quad \forall i, k. \quad (9)$$

3. Algorithm Design

The tabu search algorithm uses a neighborhood search procedure to iteratively move from one potential solution to an improved solution, until a stopping criterion is satisfied. Tabu search is a metaheuristic local search algorithm that can be used to solve combinatorial optimization problems. The major advantages of this algorithm are its simplicity, speed, and flexibility, and the scheduling model for rail train circulations in this paper is a complex zero-one programming problem. Thus, the tabu search algorithm can be used easily. The main parameters of the algorithm are designed as follows.

3.1. Expression of Solution. The two-dimensional integer array encoding method can be used to solve the train scheduling problem. In this method, rows represent trains, and columns represent trips. The trips are numbered according to departure time in ascending order. For example, in the train operation data presented in Figure 2, the trip chains of each train can be expressed as follows: train 1: 1-6-7-8-9, train 2: 2-5-10, and train 3: 3-4-11-12. The expression of solution is shown in Figure 3.

Based on the values in the two-dimensional array, the decoding process is the inverse of the encoding process. For example, Figure 3 contains 3 trains and 12 trips. The numbers of trips conducted by train 2 are 2, 5, and 10. Thus the variables are $x_2^2 = 1$, $x_5^2 = 1$, $x_{10}^2 = 1$, $y_{2,5}^2 = 1$, and $y_{5,10}^2 = 1$. Other trains are decoded using the same method used in the former method.

3.2. Generation of Initial Solution. The initial solution is the starting point of the algorithmic search. A superior initial solution enables the algorithm to arrive quickly at the optimal solution. In the process of generating the initial solution, the time-shift constraint should be satisfied. The procedure of the algorithm is as follows.

Step 1 (initialization). The set of trips to be conducted by train k is set $P_k = \emptyset$, for all k . Let $i = 1$, $k = 1$, and $a_0 = -T_0$.

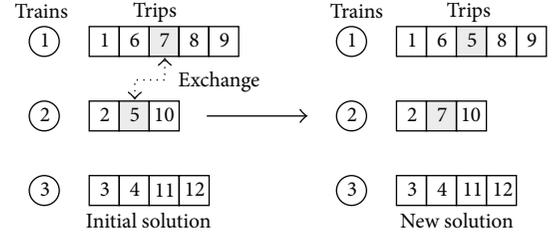


FIGURE 4: Trips exchange strategy.

Step 2. The train number k' and consecutive trip number i' corresponding to trip i are determined. If $P_k = \emptyset$, then let $k' = k$ and Step 4 is performed. Otherwise, $i' = \min\{\alpha_l \mid l \in \{1, 2, \dots, k\}\}$, $k' = \{l \mid i' = \alpha_l, l \in \{1, 2, \dots, k\}\}$, where $\alpha_l = \max\{s \mid s \in P_l\}$, and go to Step 3.

Step 3. The time-shift constraint is verified. If $d_i - a_{i'} \geq T_0$ and $s_i = z_{i'}$, then go to Step 4. Otherwise, $k \leftarrow k + 1$, and go to Step 2.

Step 4. Let $P_{k'} = P_{k'} \cup \{i\}$, $i \leftarrow i + 1$; go to Step 5.

Step 5. If $i > n$, then the algorithm ends and the results are obtained. Otherwise, go to Step 2.

3.3. Neighborhood Structure. The neighborhood structure uses trips exchange and insert strategies between different trains. The trips exchange strategy can be described as follows: a single exchange point on both the trip chains of the two parents is selected. The trip number of that point is swapped between the two parent organisms. The resulting organisms are the children. For example, trip 7 of train 1 (1-6-7-8-9) is exchanged with trip 5 of train 2 (2-5-10), and new solutions can be obtained. The trip chain of train 1 becomes 1-6-5-8-9 and that of train 2 becomes 2-7-10, as shown in Figure 4.

The trips insert strategy can be described as follows: two consecutive trips from train 1 are inserted into the trip chain of train 2. The insert point depends on the departure time in the trips chain of train 2 in ascending order, and two new trip chains for the trains result. For example, if trip 5 of train 1 (1-5-7-8-10) is inserted into the trip chain of train 2 (2-6-9), then new solutions can be obtained. The trip chain of train 1 becomes 1-8-10 and the trip chain of train 2 becomes 2-5-7-6-9, as shown in Figure 5. The trips exchange or insert operation cannot be performed if the trip chain of any crew cannot satisfy the time-shift constraint using any operation.

3.4. Evaluation of Solution. To search for better solutions in the algorithmic iterative process, we have to evaluate the solution. It must simultaneously calculate the value of the objective function and consider the constraints. Given that the initial solution satisfies the time-shift constraint and the new solution generated in the neighborhood search also satisfies the former constraint, we can consider only the equilibrium constraint as the only factor in the former steps that should be punished.

TABLE 1: Departure and arrival time of trips.

No.	DT	AT									
1	6:25	6:58	38	10:40	11:13	75	14:40	15:13	112	18:25	18:58
2	6:30	7:03	39	10:45	11:23	76	14:25	14:58	113	18:30	19:08
3	6:40	7:18	40	10:45	11:18	77	14:50	15:23	114	18:35	19:08
4	6:45	7:23	41	11:00	11:33	78	14:35	15:13	115	18:50	19:28
5	7:10	7:43	42	10:55	11:28	79	15:00	15:33	116	19:05	19:38
6	7:05	7:38	43	11:20	11:53	80	14:45	15:18	117	19:00	19:33
7	7:20	7:58	44	11:10	11:48	81	15:05	15:43	118	19:15	19:53
8	7:25	8:03	45	11:25	11:58	82	14:55	15:33	119	19:10	19:43
9	7:40	8:13	46	11:25	11:58	83	15:15	15:48	120	19:30	20:03
10	7:35	8:08	47	11:30	12:08	84	15:15	15:48	121	19:30	20:03
11	7:55	8:28	48	11:35	12:08	85	15:20	15:53	122	19:40	20:13
12	7:45	8:18	49	11:50	12:23	86	15:25	15:58	123	19:40	20:18
13	8:00	8:38	50	11:55	12:33	87	15:35	16:13	124	19:55	20:28
14	7:55	8:33	51	12:00	12:33	88	15:35	16:13	125	20:05	20:38
15	8:10	8:43	52	12:20	12:53	89	15:50	16:23	126	20:10	20:43
16	8:20	8:53	53	12:20	12:53	90	15:50	16:23	127	20:15	20:48
17	8:25	9:03	54	12:30	13:03	91	15:55	16:28	128	20:30	21:03
18	8:30	9:08	55	12:25	12:58	92	16:05	16:43	129	20:20	20:53
19	8:40	9:13	56	12:40	13:18	93	16:05	16:38	130	20:40	21:18
20	8:45	9:18	57	12:45	13:23	94	16:15	16:48	131	20:30	21:08
21	8:45	9:18	58	12:55	13:28	95	16:15	16:48	132	20:50	21:23
22	9:00	9:38	59	13:10	13:43	96	16:20	16:53	133	20:55	21:28
23	9:05	9:38	60	13:05	13:38	97	16:30	17:08	134	21:00	21:33
24	9:20	9:53	61	13:20	13:53	98	16:30	17:03	135	21:05	21:38
25	9:10	9:43	62	13:10	13:43	99	16:45	17:18	136	21:15	21:53
26	9:30	10:03	63	13:25	13:58	100	16:40	17:18	137	21:20	21:53
27	9:25	10:03	64	13:15	13:48	101	17:05	17:43	138	21:30	22:03
28	9:40	10:13	65	13:50	14:28	102	17:00	17:38	139	21:35	22:08
29	9:35	10:08	66	13:30	14:03	103	17:25	17:58	140	21:40	22:13
30	9:50	10:28	67	14:05	14:38	104	17:20	17:53	141	21:50	22:23
31	9:55	10:28	68	13:40	14:13	105	17:35	18:13	142	21:55	22:28
32	10:00	10:33	69	14:10	14:43	106	17:25	17:58	143	22:00	22:33
33	10:10	10:48	70	13:45	14:23	107	17:50	18:23	144	22:10	22:43
34	10:10	10:43	71	14:25	15:03	108	17:40	18:18	145	22:15	22:48
35	10:20	10:53	72	14:05	14:38	109	18:00	18:38	146	22:25	22:58
36	10:25	11:03	73	14:35	15:08	110	18:05	18:43	147	22:45	23:18
37	10:35	11:08	74	14:15	14:48	111	18:15	18:48	148	23:00	23:33

Notes: "DT" stands for departure time and "AT" stands for arrival time. Trains with odd numbers run from Tianjin Station to Beijing Station, and trains with even numbers run from Beijing Station to Tianjin Station.

In this study, the parameters in α can be considered as the punish factors and the value becomes a large positive number. If the solution can satisfy the equilibrium constraint, then the values of the fitness function and the objective function become equal. Otherwise, the value of the fitness function becomes significantly larger than that of the objective function, which means that the set of values for the decision variables cannot create a feasible solution. The fitness function can be formulated as follows:

$$f = Z + \alpha \cdot \max \left\{ \max_{k,k'} |G_k - G_{k'}| - T_1, 0 \right\}, \quad (10)$$

where $Z = \sum_{k=1}^m \sum_{i=1}^n \sum_{j=1}^n c_{ij} \cdot y_{ij}^k$ represents the value of the objective function.

3.5. Other Parameters. The record of the tabu table is the transform (exchange or insert) node, and tabu length has a fixed value. Regulation is selected based on the value of evaluation as the aspiration criterion. In other words, the solution of the objective can be free if it is better than any of the best candidate solutions currently known. The stopping criterion is based on the value of the fitness function. If the best value does not change after a given number of iterations, then the algorithm stops the calculation.

TABLE 2: The optimal results for train scheduling.

Train number	Trips	Operation time (min)
01	1-16-31-46-61-76-91-106-121-136	335
02	2-17-32-47-62-77-96-107-122-137	340
03	3-18-33-48-63-78-93-108-123-138	360
04	4-19-38-49-64-85-100-109-126-139	345
05	5-20-35-50-65-80-95-110-125-140	345
06	6-21-42-51-66-81-92-111-130-141	345
07	7-22-37-52-67-82-97-112-127-142	350
08	8-23-34-53-72-83-98-113-134-143	340
09	9-24-39-54-69-84-99-114-129-144	335
10	10-15-40-55-70-79-94-115-128-145	340
11	11-26-41-56-71-86-101-116-131-146	350
12	12-29-44-59-74-89-104-117-124	302
13	13-28-43-58-73-88-103-118-133-148	345
14	14-27-36-57-68-87-102-119-132-147	360
15	15-30-45-60-75-90-105-120-135	307

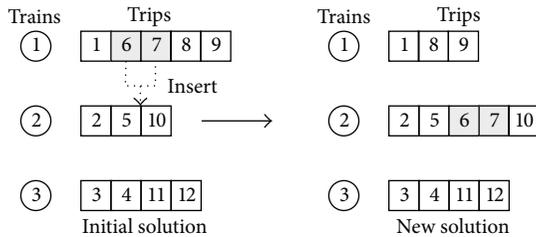


FIGURE 5: Trips insert strategy.

4. Numerical Example

The Beijing-Tianjin intercity rail line is a major railway system that serves passengers who travel between the cities of Beijing and Tianjin in China. The line starts from Beijing South Railway Station and ends at Tianjin Railway Station. It has a total length of 119.4 km and covers 74 trips in two directions every day. The main pieces of information about each trip, such as origin and destination stations as well as departure and arrival time, are presented in Table 1.

The shortest layover time for trains at switchback station is 15 minutes, and the maximum deviation value of any two trains is set to 90 minutes. The parameters for the tabu search algorithm are as follows: the tabu length is 6; the punish factor is 10,000; and the given number of iterations without improving the solution is 100. The optimal solution can be calculated using the VC++ program, as shown in Table 2. The objective value of the optimal solution is 5,099 minutes.

Table 2 indicates that the total number of trains is 15 in two directions. The maximum value of train operation time is no more than 360 minutes (nos. 3 and 14). The minimum value of train operation time is 302 minutes (no. 12). The deviation value of operation time between train nos. 3 and 12 is 58 minutes, which is less than the specified value of 90 minutes. Thus, the equilibrium constraint is satisfied.

5. Conclusions

In this paper, we present an optimized scheduling method of rail train circulations on a bidirectional intercity railway line. A binary integer programming model with the objective of minimizing the train fleet and the total interval times is presented to show the schedule process. To work out a practical schedule, the time-shift and equilibrium constraints are also considered. We have developed a tabu search algorithm to solve the proposed scheduling model.

Computational results with the trip data on Beijing-Tianjin have shown that the algorithm can produce high-quality train scheduling solutions. Furthermore, this method can be widely applied to rail train circulations characterized by high-density trips and large-quantity trains. During the peak period of an intercity railway line, because of the unbalanced feature of trip numbers in both two directions, future research work may focus on considering the rail train circulations with deadhead strategy.

Acknowledgment

The work accomplished for this paper was supported by the National Natural Science Foundation of China (no. 71261014).

References

- [1] G. Desaulniers, J. Lavigne, and F. Soumis, "Multi-depot vehicle scheduling problems with time windows and waiting costs," *European Journal of Operational Research*, vol. 111, no. 3, pp. 479–494, 1998.
- [2] J. E. Beasley and B. Cao, "A dynamic programming based algorithm for the crew scheduling problem," *Computers & Operations Research*, vol. 25, no. 7-8, pp. 567–582, 1998.
- [3] Y. D. Shen, K. K. Peng, K. Chen, and J. P. Li, "Evolutionary crew scheduling with adaptive chromosomes," *Transportation Research B*, vol. 56, no. 12, pp. 174–185, 2013.

- [4] K. Haase, G. Desaulniers, and J. Desrosiers, "Simultaneous vehicle and crew scheduling in urban mass transit systems," *Transportation Science*, vol. 35, no. 3, pp. 286–303, 2001.
- [5] D. Huisman, R. Freling, and A. P. M. Wagelmans, "Multiple-depot integrated vehicle and crew scheduling," *Transportation Science*, vol. 39, no. 4, pp. 491–502, 2005.
- [6] M. Mesquita and A. Paias, "Set partitioning/covering-based approaches for the integrated vehicle and crew scheduling problem," *Computers & Operations Research*, vol. 35, no. 5, pp. 1562–1575, 2008.
- [7] M. M. Chen and H. M. Niu, "A model for bus crew scheduling problem with multiple duty types," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 649213, 11 pages, 2012.
- [8] M. Mesquita, M. Moz, A. Paias, and M. Pato, "A decomposition approach for the integrated vehicle-crew-roster problem with days-off pattern," *European Journal of Operational Research*, vol. 229, no. 2, pp. 318–331, 2013.
- [9] H. M. Niu and M. H. Zhang, "An optimization to schedule train operations with phase-regular framework for intercity rail lines," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 549374, 13 pages, 2012.
- [10] H. M. Niu and X. S. Zhou, "Optimizing urban rail timetable under time-dependent demand and oversaturated conditions," *Transportation Research C*, vol. 36, no. 16, pp. 212–230, 2013.
- [11] Z. Huang and H. Niu, "Study on the train operation optimization of passenger dedicated lines based on satisfaction," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 451201, 11 pages, 2012.
- [12] A. Schrijver, "Minimum circulation of railway stock," *CWI Quarterly*, vol. 6, pp. 205–221, 1993.
- [13] X. Cai and C. J. Goh, "A fast heuristic for the train scheduling problem," *Computers & Operations Research*, vol. 21, no. 5, pp. 499–510, 1994.
- [14] A. Higgins, E. Kozan, and L. Ferreira, "Optimal scheduling of trains on a single line track," *Transportation Research B*, vol. 30, no. 2, pp. 147–158, 1996.
- [15] K. Ghoseiri, F. Szidarovszky, and M. J. Asgharpour, "A multi-objective train scheduling model and solution," *Transportation Research B*, vol. 38, no. 10, pp. 927–952, 2004.
- [16] E. Abbink, B. Van den Berg, L. Kroon, and M. Salomon, "Allocation of railway rolling stock for passenger trains," *Transportation Science*, vol. 38, no. 1, pp. 33–41, 2004.
- [17] D. Huisman, "A column generation approach for the rail crew re-scheduling problem," *European Journal of Operational Research*, vol. 180, no. 1, pp. 163–173, 2007.
- [18] S. Jütte and U. W. Thonemann, "Divide-and-price: a decomposition algorithm for solving large railway crew scheduling problems," *European Journal of Operational Research*, vol. 219, no. 2, pp. 214–223, 2012.
- [19] S. Qu and X. Xu, "Scheduling model of multiple units on intercity railway," *Journal of Tongji University*, vol. 38, no. 9, pp. 1298–1302, 2010.
- [20] G. Sahin and B. Yüceoglu, "Tactical crew planning in railways," *Transportation Research E*, vol. 47, no. 30, pp. 1221–1243, 2011.
- [21] F. Shi, W.-L. Zhou, Y.-W. Yu, and L. Qing, "Optimized model and algorithm of motor train-sets scheduling for dedicated passenger lines," *Journal of the China Railway Society*, vol. 33, no. 1, pp. 7–13, 2011.
- [22] A. Azadeh, M. H. Farahani, H. Eivazy, S. Nazari-Shirkouhi, and G. Asadipour, "A hybrid meta-heuristic algorithm for optimization of crew scheduling," *Applied Soft Computing*, vol. 13, pp. 158–164, 2013.

Research Article

Empirical Study of Travel Time Estimation and Reliability

Ruimin Li,¹ Huajun Chai,² and Jin Tang¹

¹ Department of Civil Engineering, Institute of Transportation Engineering, Tsinghua University, Beijing 100084, China

² Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, USA

Correspondence should be addressed to Ruimin Li; lrmin@tsinghua.edu.cn

Received 25 August 2013; Accepted 24 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Ruimin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper explores the travel time distribution of different types of urban roads, the link and path average travel time, and variance estimation methods by analyzing the large-scale travel time dataset detected from automatic number plate readers installed throughout Beijing. The results show that the best-fitting travel time distribution for different road links in 15 min time intervals differs for different traffic congestion levels. The average travel time for all links on all days can be estimated with acceptable precision by using normal distribution. However, this distribution is not suitable to estimate travel time variance under some types of traffic conditions. Path travel time can be estimated with high precision by summing the travel time of the links that constitute the path. In addition, the path travel time variance can be estimated by the travel time variance of the links, provided that the travel times on all the links along a given path are generated by statistically independent distributions. These findings can be used to develop and validate microscopic simulations or online travel time estimation and prediction systems.

1. Introduction

Traffic congestion during peak hours has become unavoidable in numerous cities worldwide because of the rapid increase in car ownerships and the lack of resources for proportionately increasing the supply capacity of road systems. This problem is causing travel time to be highly unreliable. Travel time and travel time reliability have become important performance measures in assessing traffic system conditions. A previous study [1] suggests that travel time reliability may be more important than travel time savings; that is, road users may choose a reliable route over an unreliable one despite the longer travel time of the former. The reliability and variability of travel time have attracted significant attention in the past decade.

Traffic systems are complex and stochastic. For instance, a number of notable sources of traffic congestion (traffic incidents, work zones, bad weather, special events, and traffic demand fluctuations) can disrupt system performance and lengthen travel time. A high variability indicates the unpredictability of travel time and the reduced reliability of traffic service [2]. However, the accurate estimation and prediction of travel time are essential to traffic operation and traveler information systems.

The insufficient amount of actual travel time data has caused several previous studies to use either loop detector data to analyze travel time reliability or microsimulation technique. In the past decade, numerous cities have implemented various travel time direct measurement techniques, such as automatic number plate readers (ANPR), automated vehicle identification (AVI) systems, GPS-equipped vehicles, smart phone devices, and Bluetooth [3]. All of these techniques provide accurate individual vehicle travel time data for analysis. For instance, gathering travel times with moving GPS-equipped vehicles produces accurate, continuous, and automated point-to-point measures, which are more representative of road performance compared with the point estimates of speed from fixed detectors. However, the sampling rate is occasionally low (no more than 10% of the traffic flow) without the wide use of GPS equipment in private cars.

The remainder of the paper is organized as follows. Section 2 provides a brief review of the literature with respect to modeling travel time distribution. A description of the data set used in this study and data preprocessing are then presented in Section 3. Section 4 explains the methodology used to investigate the statistical properties of the travel time distributions and results and discusses whether the normal

distribution can be used for all travel times to obtain the mean and variance, which usually indicates the travel time variability. Finally, Section 5 provides the conclusion of the study.

2. Literature Review

Travel time distribution is an important basis for modeling travel time variability and reliability, which can be measured using several travel time distribution properties, such as standard deviation and coefficient of variation. Previous studies on travel time variability and reliability assumed that travel time distribution may follow either normal distribution or log-normal distribution [4], particularly for freeways. However, several studies suggested that travel time data are skewed and that such data possess long upper tail [5, 6].

Numerous studies have been conducted to investigate the probability distribution of travel time on freeways and signalized arterial roads. Recently collected empirical travel time data exhibit positive skews and long tails. Therefore, normal distribution is not suitable for these data. During the 1950s, Wardrop [7] reported that travel times follow a skewed distribution with a long tail. Using the data of trip times to and from work on 25 routes in Detroit collected over 20 months; Herman and Lam [8] subsequently verified this observation with an empirical study on travel time variability. This study showed that the to-work trip time histogram tended to follow a normal distribution, whereas the from-work trip time histogram closely resembled a uniform distribution. Using the travel time data collected in Michigan, Polus [9] concluded that travel times fit a gamma distribution relatively well. Richardson and Taylor [10] analyzed the travel time data in Melbourne and found that the observed travel time variability may be represented by a log-normal distribution.

Dandy and McBean suggested a log-normal or gamma distribution [11] and proposed that the log-normal distribution is a useful descriptor of in-vehicle travel times. A log-normal fit has been derived in several studies, such as those by Mogridge and Fry [12] and Arroyo and Kornhauser [13]. Montgomery and May [14] conducted a survey in Leeds city and found that individual travel times were fit to a log-normal distribution with means and dispersions that are considerably more stable on a number of routes than others. Rakha et al. [15] studied the AVI data from San Antonio through goodness-of-fit tests and demonstrated that normal distribution is inconsistent with the field travel time observations. A previous study [16] showed that a lognormal distribution is highly representative of travel times, particularly under steady state conditions, and that a mixture distribution is appropriate for modeling travel times under nonsteady state conditions. Chen et al. [17] modeled the travel rate (travel time per unit distance) distribution and found that a log-normal distribution exhibits good approximation performance for the travel rate distribution at an a.m. peak hour on Wednesday.

To obtain a better view of the travel time distribution, Susilawati et al. [18] analyzed the GPS travel time data from the Adelaide database and found that the Burr distribution fits the empirical travel time data, that such distribution

can be used to represent the observed data at two urban arterial roads in Adelaide, and that the bimodal distribution is appropriate for travel time distribution in a number of short arterial links [19]. Fosgerau and Fukuda [20] studied the minute-by-minute travel times for a congested urban road over five months and found that a stable distribution fits the standardized travel time of a link or a sequence of links. A number of recent studies have focused on the link travel time distribution prediction to generate reliable real-time traffic condition forecasts [21, 22].

In most of these studies, the data derived from GPS-based probe cars only provide the travel time for the probe vehicles, resulting in a small sampling rate of traffic flow and preventing the collection of a large sampling data set for various routes and times of day. Moreover, this low sampling rate may not reflect the real distribution of travel time during a short period, such as 15 min. With the rapid development of traffic flow data collection techniques, the ANPR system allows traffic management engineers to record the time when a vehicle passes a specific location on roads. The time differences between continuous locations can be directly used as the travel time of this vehicle. The ANPR system provides almost the entire sampling rate of the traffic flow, except for the identification error of the camera. These accurate travel time data are beneficial to the studies on travel time variations in urban environments.

3. Data Preprocessing

Different travel time monitoring techniques have been developed over the past decades. These techniques include ANPR cameras, Bluetooth scanners, GPS-based in-car devices, smart phones, and speed sensors. The first two have been proven to be promising methods. Various models for estimating or predicting travel time distribution have been proposed on the basis of the data obtained from these techniques, and these models have exhibited excellent performance on freeways.

Approximately 200 ANPR identification detectors are currently being mounted throughout Beijing city to collect vehicle passing time. These detectors allow the travel time on the target road link to be obtained through the comparison and analysis of the passing time between two consecutive detectors.

However, the major problem of ANPR in urban networks lies in the difficulty in determining whether a vehicle has traveled exactly along the route between two locations without making unexpected stops. Thus, a number of invalid travel times from individual vehicles are inevitably observed. These invalid travel times do not represent the average traffic conditions on the link considered at the time the vehicle was detected. For instance, an invalid travel time that is considerably longer than the average travel time of vehicles can be observed from a vehicle making unexpected stops between two detection stations or from a bus that must stop at bus stops to load and unload passengers. Such travel times must be removed from the dataset to avoid bias in the analysis results.

The data must be preprocessed to remove the outlier value before the estimation of the data distribution. Quartile screening method using quartile interval is applied in this study to reflect the variation scale.

Specifically, the data interval is the difference between the upper and lower quartiles. If the data lies outside the interval, then it will be identified as abnormal data and deleted:

$$\begin{aligned} Z &= [Q_{0.25} - 1.5R, Q_{0.75} + 1.5R], \\ R &= Q_{0.75} - Q_{0.25}, \end{aligned} \quad (1)$$

where Z indicates the valid data interval, $Q_{0.25}$ and $Q_{0.75}$ are the lower and upper quartile values (the lowest and highest 25% of the data), respectively, and R is the quartile range.

Figure 1 presents a comparison diagram that shows the before and after preprocessing results obtained through quartile screening method for the travel time dataset of a single day in link 69–68. Several scattered and unreasonable values were discarded.

4. Estimation

4.1. Estimation of Travel Time Distribution. The influence of signal timing and other parameters causes the travel time on arterial road links to show unimodal, bimodal, or multimodal distribution shape. Therefore, in this study, the distribution patterns of the travel times of different links were analyzed in order of priority. A total of 17 unimodal distribution links were selected: seven signalized arterials with lengths ranging from 417 m to 2028 m and 10 nonsignalized urban expressways with lengths ranging from 1600 m to 4100 m.

A 15 min interval is used to study how the travel time distribution on actual road varies over time. This interval provides sufficient data for most times of a day, and these data can be used for travel time distribution estimation. In addition, this interval is short enough to capture short-term variations in travel time.

The data collected on June 16, 2011, were selected for use in this paper. For each road link, a total of 96 15 min travel time datasets may be used for analysis. Normal, log-normal, gamma, and Weibull distributions are fitted to these 15 min travel time datasets, and the Chi squares are selected for testing goodness-of-fit.

To determine the relationship between the distribution pattern and the congestion degree of the road link, travel speed \bar{v}_j is used to indicate the congestion degree of the road link. Then, the average travel speed on each link is calculated using the following formula for each 15 min interval:

$$\bar{v}_j = \frac{l_j}{\bar{t}_j}, \quad j = 1, 2, \dots, 17, \quad (2)$$

where l_j is the length of link j and \bar{t}_j is the estimated time on link j .

The relationship between the travel speed \bar{v}_j and the travel time can be studied. The traveler information system of the Beijing Traffic Management Bureau was used to classify the traffic conditions of various types of roads in Beijing according to the average travel speed, as shown in Table 1.

As shown in Table 1, the numbers and percentiles of the different distributions for each traffic condition were observed for each road type, that is, arterials and urban expressways. A total of 660 15 min intervals without adequate data to test the distribution are not shown in Table 1.

Table 1 shows that when the average speed on certain road links is relatively low (i.e., congested traffic), the Weibull distribution percentile is the largest (arterial: 0.432; urban expressway: 0.378). The average travel speed increase causes a decrease in the Weibull distribution percentile and an increase in log-normal distribution percentile. Under free flow traffic conditions in urban expressways, the log-normal distribution percentile is 0.569, indicating that the log-normal distribution best fits the distribution of the 15 min travel time for most intervals under free flow conditions. Evidently, different traffic conditions have different travel time distributions, suggesting that we should adopt different distributions for travel time reliability and variability estimation.

4.2. 15 min Travel Time Estimation. For a 15 min travel time and its variance estimation, the specific distribution of travel time for certain road links during certain times must first be obtained using distribution fitting process; then, the average travel time and the standard deviation of travel time can be derived from the corresponding distribution function. However, in actual application, the parameter estimation of log-normal, gamma, and Weibull distributions is complicated, thereby highly constraining online real-time application. By contrast, the parameter estimation of normal distribution is simple. Therefore, the average travel time and variance estimation results for the normal distribution were compared with those for the other distributions (i.e., log-normal, gamma, and Weibull). In actual application, such as travel time estimation for dynamic traffic assignment or travel time reliability estimation, if the difference is within an acceptable range, then normal distribution can be considered a substitute for other distributions to obtain higher real-time calculation efficiency and reduce computer workload.

First, we consider the “Nanheyan Intersection -> Wangfujing Intersection” as an analysis example. The dataset consists of four 15 min intervals from 8 AM to 9 AM on June 16, 2011. The average travel time and variance estimation results of different distributions in various 15 min intervals are shown in Table 2:

$$\begin{aligned} \delta_1 &= \left| \frac{t_i - t_{\text{normal}}}{t_i} \right| \times 100\%, \\ \delta_2 &= \left| \frac{\sigma_i - \sigma_{\text{normal}}}{\sigma_i} \right| \times 100\%, \end{aligned} \quad (3)$$

where i represents the log-normal, gamma, or Weibull distribution; t_i and σ_i indicate the estimated average travel time and standard deviation of travel time from i distribution, respectively; t_{normal} and σ_{normal} indicate the estimated average travel time and standard deviation of travel time from normal distribution, respectively; δ_1 and δ_2 are the absolute percentage errors of the estimated average travel time and standard deviation of travel time, respectively.

TABLE 1: Distribution form of 15 min travel time under different traffic conditions.

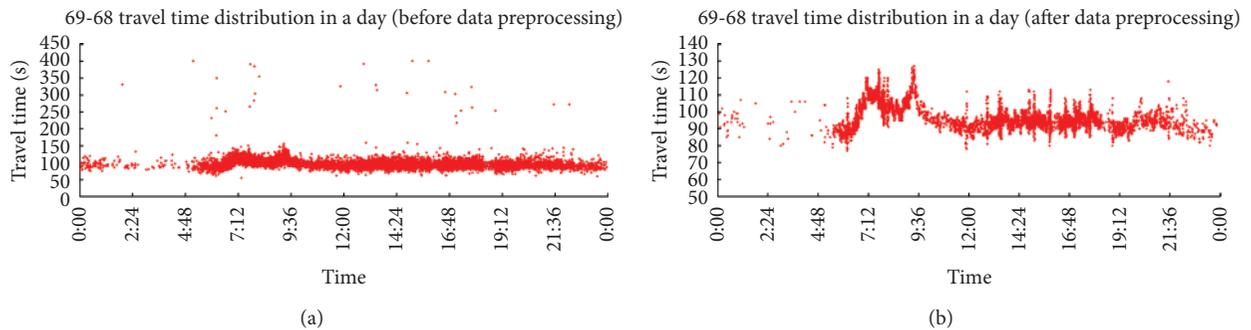
Signalized arterial link	Congested <10 km/h		Slow 10 km/h to 30 km/h		Free flow >30 km/h	
	Number	Percentile	Number	Percentile	Number	Percentile
Normal distribution (1)	12	0.136	23	0.155	13	0.169
Log-normal distribution (2)	23	0.261	55	0.372	35	0.455
Gamma distribution (3)	15	0.170	22	0.149	17	0.221
Weibull distribution (4)	38	0.432	48	0.324	12	0.156

Urban Expressway	Congested <20 km/h		Slow 20 km/h to 50 km/h		Free flow >50 km/h	
	Number	Percentile	Number	Percentile	Number	Percentile
Normal distribution (1)	10	0.222	48	0.142	16	0.050
Log-normal distribution (2)	12	0.267	149	0.442	181	0.569
Gamma distribution (3)	6	0.133	98	0.291	115	0.362
Weibull distribution (4)	17	0.378	42	0.125	6	0.019

TABLE 2: Errors between estimation of normal distribution and other distributions.

	8:00–8:15				8:15–8:30			
	T (average travel time) (seconds)	Error δ_1	σ (standard deviation of travel time)	Error δ_2	T (average travel time) (seconds)	Error δ_1	σ (standard deviation of travel time)	Error δ_2
Normal distribution	46.587	—	31.803	—	47.688	—	34.109	—
Log-normal distribution	45.593	2.18%	29.036	9.53%	46.592	2.35%	31.789	7.30%
Gamma distribution	46.587	0	28.03	13.46%	47.688	0	30.241	12.79%
Weibull distribution	47.143	1.18%	29.891	6.40%	48.257	1.18%	31.988	6.63%

	8:30–8:45				8:45–9:00			
	T (average travel time) (seconds)	Error δ_1	σ (standard deviation of travel time)	Error δ_2	T (average travel time) (seconds)	Error δ_1	σ (standard deviation of travel time)	Error δ_2
Normal distribution	73.982	—	32.15	—	56.626	—	28.849	—
Log-normal distribution	75.431	1.92%	43.737	26.49%	56.929	0.53%	33.011	12.61%
Gamma distribution	73.982	0	36.347	11.55%	56.626	0	29.279	1.47%
Weibull distribution	74.214	0.31%	30.92	3.98%	56.964	0.59%	28.201	2.30%

FIGURE 1: Travel time distribution of link 69-68 before and after preprocessing. *Note.* 68, 69 is the ID number of the intersections; link 69-68 refers to the unidirectional link from intersection 69 to intersection 68.

The following observations are drawn from Table 2. (1) For the estimation of average travel time, the errors between the estimation results obtained from the normal distribution and those obtained from other complex distributions are insignificant, with the maximum being 2.35%. (2) For the estimation of standard deviation of travel time, the errors are large (ranging from 1.47% to 26.49%). Therefore, whether the normal distribution can substitute other distributions for parameter estimation must be decided on the basis of the accuracy requirements for actual application.

Without loss of generality, 17 road links observed on June 16, 2011, were used to determine the estimation error between the best-fitting and normal distributions. A total of 962 valid 15 min travel time datasets were left for analysis after the invalid datasets were removed. The errors of average travel time estimation and standard deviation estimation derived using normal distribution with other distributions are shown in Figures 2(a) and 2(b), respectively.

As shown in Figures 2(a) and 2(b), the average travel time estimation using normal distribution is either more or less than that using other distributions, and the maximum error is $\pm 2\%$. However, the errors between standard deviation estimation with different distributions are relatively large, with the maximum error reaching -90% . The mean absolute relative error (MARE) was only 6.9%. The standard deviation estimation errors are mostly distributed within $(-0.1, 0.1)$, accounting for over 85%; that is, under most traffic situations, the accuracy of the average travel time estimation and standard deviation estimation through normal distribution estimation for various road links is acceptable.

These results suggest that the average travel time under most traffic situations can be estimated using normal distribution. For the estimation of the standard deviation of travel time, that is, travel time variability or reliability, the errors are occasionally large. For high accuracy requirements, for example, more than 95%, the accuracy of normal distribution estimation cannot be ensured.

4.3. Trip Travel Time Estimation. In actual trips, travelers are more concerned with the travel time of a path or trip, which is the travel time from the original point to the destination. For a path s composed of m links that form the set $L(s)$, the path travel time can be computed using the following two methods.

Method 1. If there is a set $V(s)$ of n vehicles travelling along the path s , then the path travel time at a certain time interval can be calculated as follows:

$$\bar{T} = E_i \{T_i\} = \frac{\sum_{i=1}^n T_i}{n}, \quad \forall i \in V(s), \quad (4)$$

where $E_i\{*\}$ denotes the conditional expectation over all the realizations i and T_i is the travel time experienced by a vehicle $i \in V(s)$ along the path s .

Similarly, the path travel time variance can be computed using the path travel time realizations [16]:

$$\sigma_t^2 = \frac{\sum_{i \in V(s)} T_i^2}{n} - \bar{T}^2. \quad (5)$$

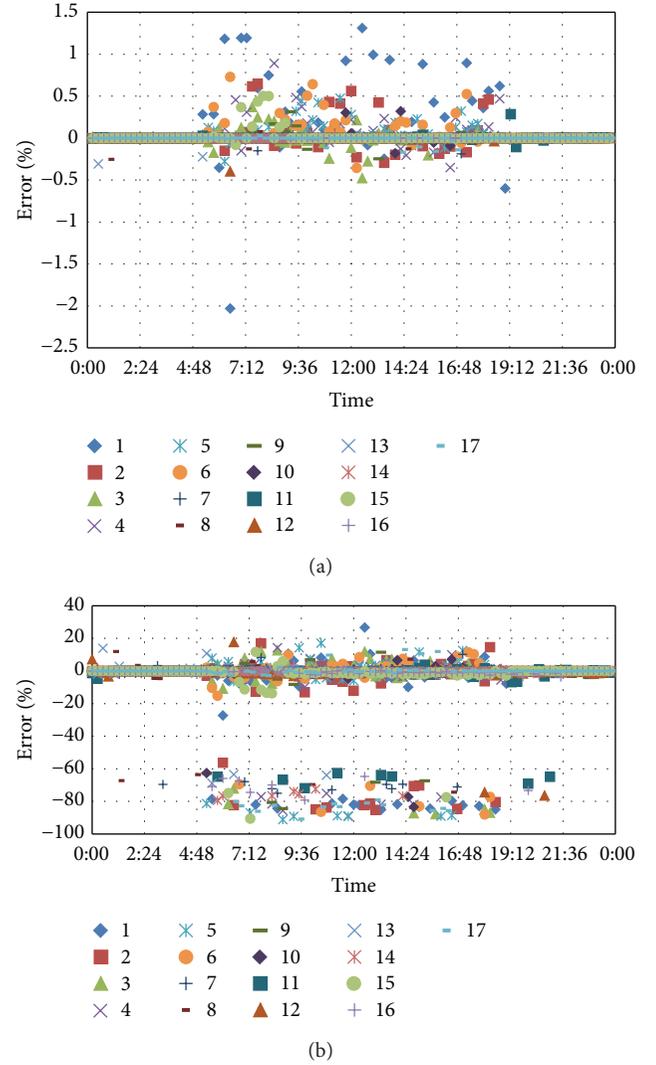


FIGURE 2: (a) Average travel time estimation errors using normal distribution with other distributions. (b) Errors of standard deviation estimation using normal distribution with other distributions. *Note.* The numbers at the bottom of this figure indicate the road link number.

This equation serves as the ground truth used for comparison purposes.

Method 2. The majority of vehicles do not travel the entire trip length; thus, the trip travel times from various road segments or link travel times should be estimated. Without loss of generality, the path travel time can be determined as the sum of the experienced link travel times along the set of links constituting the path, as shown in the following:

$$\bar{T} = \sum_{j \in L(s)} \bar{t}_j, \quad (6)$$

where \bar{t}_j is the travel time experienced by vehicles along link $j \in L(s)$.

The expected travel time along a link can be computed as follows:

$$\bar{t}_j = E_i \{t_{ij}\} = \frac{\sum_{i=1}^p t_{ij}}{p}, \quad \forall j \in L(s), \quad (7)$$

where a total of p vehicles are traveling along link j .

The travel time variance for link j can be computed as follows:

$$\sigma_j^2 = E_i \{t_{ij} - \bar{t}_j\}^2 = E_i \{t_{ij}^2\} - \bar{t}_j^2 = \frac{\sum_{i \in V(s)} t_{ij}^2}{n} - \bar{t}_j^2. \quad (8)$$

For method 2, three methods were used to compute the path travel time variance from the link travel time variance [16].

Method 2.1. The current state-of-the-practice technique for estimating path travel time variability assumes that the travel times on all the links along a given path are generated by statistically independent distributions. Therefore, the path variance can be computed as the summation of the link travel time variances for all links along a path as follows (Method 2.1, noted with subscript 1):

$$\sigma_{1t} = \left(\sum_j^m \sigma_j^2 \right)^{1/2}. \quad (9)$$

Method 2.2. Sherali et al. [23] used the maximum and minimum link travel time coefficients of variations (CV) to construct bounds on the path CV because the CV is independent of the length of the link. The following equation is derived to estimate the path variance (Method 2.2, noted with subscript 2):

$$\begin{aligned} \sigma_{2t} &= \bar{T} \cdot \text{COV}_t = \bar{T} \cdot \left(\frac{\sum_{j \in L(s)} \sigma_j^2}{\sum_{j \in L(s)} \bar{t}_j^2} \right)^{1/2} \\ &= \frac{\bar{T}}{\left(\sum_{j \in L(s)} \bar{t}_j^2 \right)^{1/2}} \sigma_{1t}, \end{aligned} \quad (10)$$

where COV_t is the coefficient of variation.

Method 2.3. Another method for estimating the trip variance [16] (Method 2.3, noted with subscript 3) is to compute the expected path CV as the conditional expectation over all realizations j of the various roadway links making up a path. The path variance estimation can be computed as follows:

$$\sigma_{3t} = \frac{\bar{T}}{m} \left(\sum_{j \in L(s)} \frac{\sigma_j}{\bar{t}_j} \right). \quad (11)$$

The majority of vehicles do not travel the entire trip length; thus, the travel time data of a path within a short-term interval is limited. To retrieve a significant amount of travel time data, the sampling time interval is set to 30 mins in this study. Five paths are used for analysis. The main

consideration in choosing these five paths are as follows. (1) In the observed period, the datasets for these paths have an enough number of valid records to support the analysis. (2) These paths have consecutive links within the path. The dataset has a travel time for each link and a travel time for the entire path, thus allowing the correlation analysis of the links and error estimation. The data of the selected path consist of 24 hour data obtained from September 3, 2008. This dataset is then divided into 48 groups of values for every 30 mins.

The five paths travel time estimation error is illustrated in Figures 3(a) and 3(b). The standard path travel time for error estimation is directly obtained from the estimation of path travel time and its standard error, as described in method 1 and shown in (4).

As shown in Figures 3(a) and 3(b), most of the related errors of the five paths are lower than 10%, except for a few values exceeding 10%.

Figure 4 shows that the five paths' travel time variance estimation errors derived by the three methods. As shown in Figure 4, most of the cases using Method 2.1 have the minimum estimation error of path travel time variance. For path 18-17-16 (Chang'an Street, signalized arterials) and path 136-135-134 (Jing-Kai Expressway), Methods 2.2 and 2.3 have similar errors for most intervals. For the other three paths (all of which are expressways), no evident regular patterns are found. The MARE of three methods for path travel time variance estimation of the five paths is shown in Table 3.

The following observations are drawn from Table 3. (1) For the entire day, the MARE of paths 18-17-16, 70-69-68, and 136-135-134 show that Method 2.1 has the lowest error, whereas the MARE of paths 133-134-135 and 140-139-138 show that Method 2.3 has the lowest error. (2) For the a.m. peak hours, the MARE of four paths (i.e., 18-17-16, 70-69-68, 140-139-138, and 136-135-134) show that Method 2.1 has the lowest error. For path 133-134-135, Methods 2.1 and 2.3 have similar precision. (3) For the p.m. peak hours, the MARE of the two paths with data show that Method 2.1 has the lowest error. For nonpeak hours, a similar regular pattern throughout the entire day is found.

These results are relatively different from those described in [16], which utilized field data and simulated data. In this study, Method 2.2 is shown to have the highest estimation errors, whereas Method 2.1 is shown to have the best estimation performance under most situations.

These findings suggest that the path average travel time can be estimated by summing the travel time along the set of links constituting the path. On the basis of the performance of Method 2.1, the assumption that travel times on all links along a path are generated by statistically independent distributions is reasonable under most traffic conditions, particularly for congested periods. Nevertheless, for a number of paths, Method 2.3 is more suitable than Method 2.1 for calculating the path travel time variance.

5. Conclusion

This paper discussed three issues concerning travel time variability and reliability estimation on the basis of the

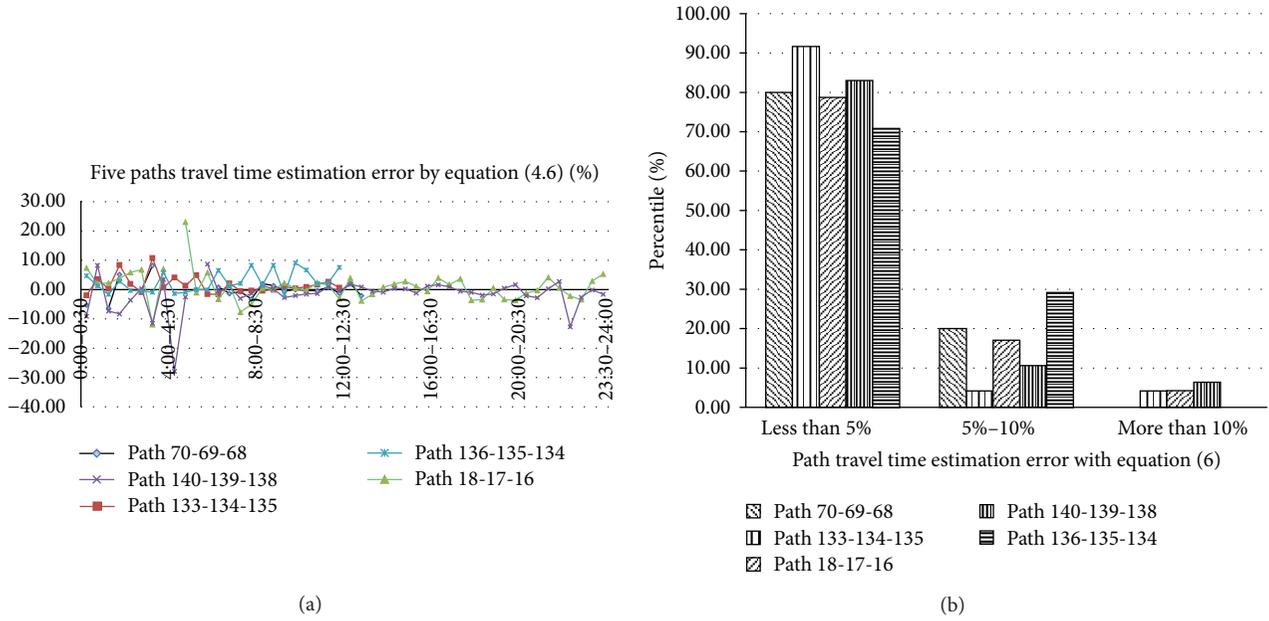


FIGURE 3: (a) Path travel time estimation error by (6). (b) Statistical graph of path travel time estimation error by (6). *Note.* Path 70-69-68 indicates the unidirectional path from intersection 70 to intersection 68 via intersection 69. All numbers following the word “path” are the ID numbers of intersections.

TABLE 3: MARE of the three methods for path travel time variance estimation of five paths.

MARE	Method	Path 18-17-16	Path 70-69-68	Path 133-134-135	Path 140-139-138	Path 136-135-134
Whole day	Method 2.1	18.96	9.40	24.49	17.88	28.54
	Method 2.2	57.04	26.70	22.85	29.67	33.72
	Method 2.3	53.97	16.47	16.12	17.15	36.96
a.m. (7:00–9:00)	Method 2.1	6.83	12.94	12.23	5.37	31.17
	Method 2.2	47.60	22.03	24.11	24.18	35.78
	Method 2.3	43.77	19.79	11.49	6.02	35.88
p.m. (17:00–19:00)	Method 2.1	10.36	#	#	4.39	#
	Method 2.2	40.33	#	#	21.88	#
	Method 2.3	39.46	#	#	4.84	#
Nonpeak hours	Method 2.1	21.57	8.22	26.95	20.69	27.99
	Method 2.2	61.18	28.25	22.60	31.10	38.12
	Method 2.3	57.89	15.36	17.05	19.68	37.18

indicates that, during this period, there are no data for the paths.

ANPR data in the Beijing road network. The first observation involves link travel time distribution estimation through goodness-of-fit tests within short-term time intervals of 15 min. The results show that for congested periods, approximately 40% of the 15 min travel time distribution followed Weibull distribution, with 43.2% for signalized arterials and 37.8% for urban expressways. The increase in average travel speed caused the 15 min travel time distribution percentile following log-normal distribution to gradually increase until the values become 45.5% for signalized arterials and 56.9% for urban expressways (uncongested traffic conditions).

The second observation involves the quick estimation process of the average travel time for 15 min intervals. The comparison between estimation results based on normal distribution and those based on other more fitting distributions shows that the estimation results based on normal distribution for the average travel time estimation are acceptable, with errors of no more than 2% and less than 1% for more than 90% of the estimations. For travel time variance estimation, although the MARE is 6.9%, approximately 1/3 of the absolute relative error is more than 50%. Thus, under most traffic conditions, normal distribution can be used to estimate

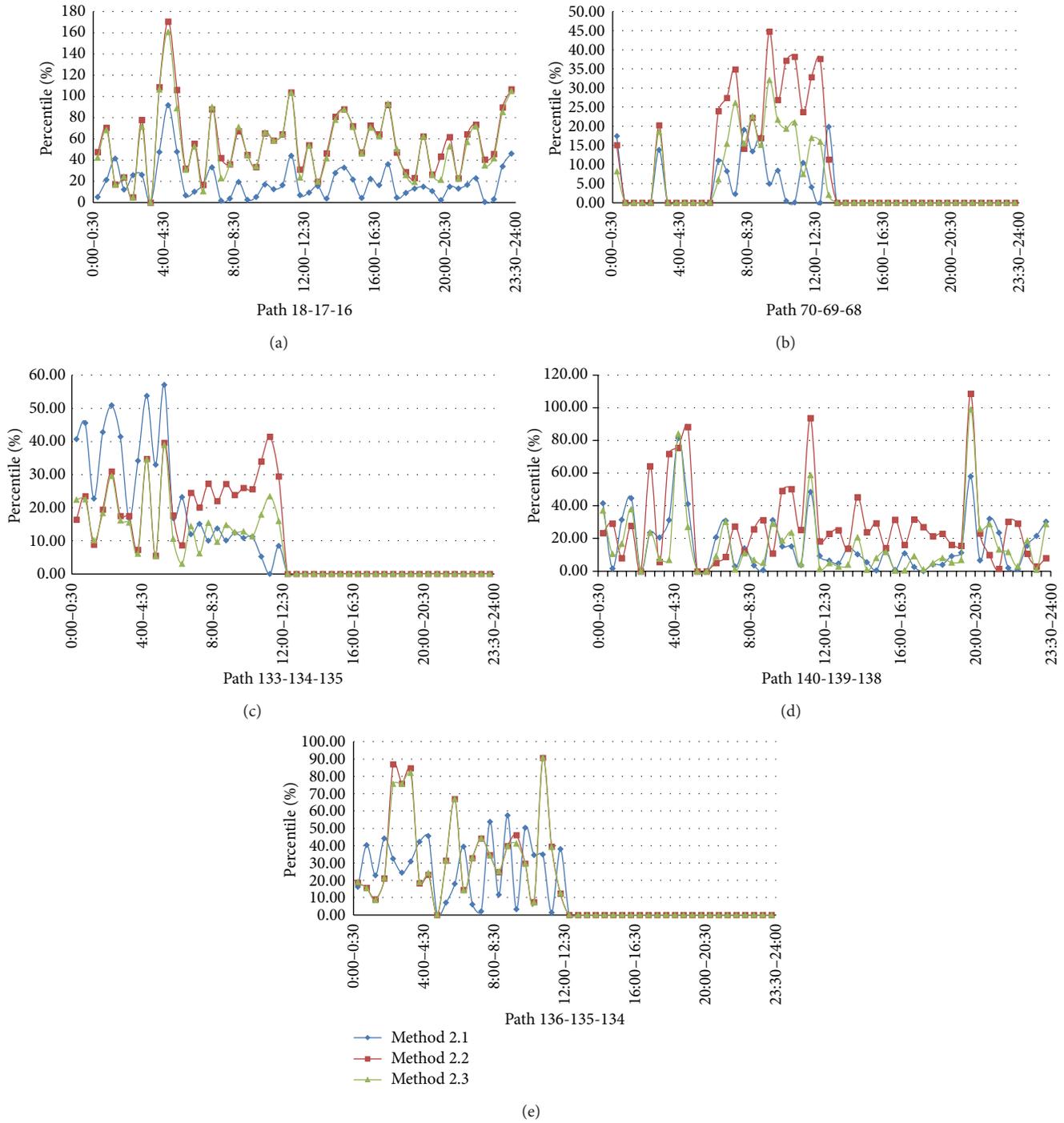


FIGURE 4: Standard deviation of travel time estimation. *Note.* Path 18-17-16 is from intersection 18 to 16 via 17; the value 0 of error means that, in this time interval, no vehicle travel time data along the entire path is found.

the average travel time, whereas the most fitting distribution can be used to estimate the travel time variance.

The third observation involves the estimation process of the path travel time and variance on the basis of the travel time of the links constituting the path. The results show that the path travel time can be computed by summing the average travel time of links constituting the path. In addition, Method

2.1 produces the best results for path travel time variance estimation under most traffic conditions, particularly during congested periods.

In the future, more travel time data should be collected to conduct similar studies based on large-scale data sets and to further understand the travel time distribution pattern and relationship between path travel time and travel time of links

constituting a given path, particularly for long distance paths or trips with more than three links.

Conflict of Interests

The authors declare the lack of conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 71361130015). The authors are personally grateful to Beijing Traffic Management Bureau for its support.

References

- [1] J. Bates, J. Polak, P. Jones, and A. Cook, "The valuation of reliability for personal travel," *Transportation Research Part E*, vol. 37, no. 2-3, pp. 191–229, 2001.
- [2] R. E. Turochy and B. L. Smith, "Measuring variability in traffic conditions by using archived traffic data," *Transportation Research Record*, no. 1804, pp. 168–172, 2002.
- [3] M. Martchouk, F. Mannering, and D. Bullock, "Analysis of freeway travel time variability using bluetooth detection," *Journal of Transportation Engineering*, vol. 137, no. 10, pp. 697–704, 2011.
- [4] N.-E. E. Faouzi and M. Maurin, "Reliability metrics for path travel time under log-normal distribution," in *Proceedings of the 3rd International Symposium on Transportation Network Reliability*, The Hague, The Netherlands, 2007.
- [5] J. W. C. Van Lint and H. J. Van Zuylen, "Monitoring and predicting freeway travel time reliability: using width and skew of day-to-day travel time distribution," *Transportation Research Record*, no. 1917, pp. 54–62, 2005.
- [6] S. Susilawati, M. A. P. Taylor, and S. V. C. Somenahalli, "Travel time reliability measurement for selected corridors in the Adelaide Metropolitan area," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 8, pp. 86–102, 2010.
- [7] J. G. Wardrop, "Some theoretical aspects of road traffic research," *ICE Proceedings: Engineering Divisions*, vol. 1, no. 3, pp. 362–378, 1952.
- [8] R. Herman and T. Lam, "Trip time characteristics of journeys to and from work," in *Proceedings 6th International Symposium on Transportation and Traffic Theory*, Sydney, Australia, 1974.
- [9] A. Polus, "A study of travel time and reliability on arterial routes," *Transportation*, vol. 8, no. 2, pp. 141–151, 1979.
- [10] A. J. Richardson and M. A. P. Taylor, "Travel time variability on commuter journeys," *High Speed Ground Transportation Journal*, vol. 12, no. 1, pp. 77–99, 1978.
- [11] G. C. Dandy and E. A. McBean, "Variability of individual travel time components," *Journal of Transportation Engineering*, vol. 110, no. 3, pp. 340–356, 1984.
- [12] M. Mogridge and S. Fry, "The variability of car journeys on a particular route in central London," *Traffic Engineering and Control*, vol. 25, no. 10, pp. 510–511, 1984.
- [13] S. Arroyo and A. L. Kornhauser, "Modeling travel time distributions on a road network," in *Proceedings of the 11th World Conference on Transport Research*, Berkeley, Calif, USA, 2007.
- [14] F. O. Montgomery and A. D. May, "Factors affecting travel times on urban radial routes," *Traffic Engineering and Control*, vol. 28, no. 9, pp. 452–458, 1987.
- [15] H. Rakha, I. El-Shawarby, M. Arafef, and F. Dion, "Estimating path travel-time reliability," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC '06)*, pp. 236–241, September 2006.
- [16] H. Rakha, I. El-Shawarby, and M. Arafef, "Trip travel-time reliability: issues and proposed solutions," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 14, no. 4, pp. 232–250, 2010.
- [17] K. Chen, L. Yu, J. F. Guo, and H. M. Wen, "Characteristics analysis of road network reliability in Beijing based on the data logs from taxis," in *Proceedings of the 79th Transportation Research Board Annual Meeting*, Washington, DC, United States, 2007.
- [18] S. Susilawati, M. A. P. Taylor, and S. Somenahalli, "Distribution of variability in travel times on urban roads—a longitudinal study," in *Proceedings of the 12th World Conference of Transport Research*, Lisbon, Portugal, 2010.
- [19] S. Susilawati, M. A. P. Taylor, and S. Somenahalli, "Travel time reliability and the bimodal travel time distribution for an arterial road," *Road and Transport Research*, vol. 19, no. 4, pp. 37–50, 2010.
- [20] M. Fosgerau and D. Fukuda, "Valuing travel time variability: characteristics of the travel time distribution on an urban road," *Transportation Research Part C*, vol. 24, no. 5-6, pp. 83–101, 2012.
- [21] M. Ramezani and N. Geroliminis, "On the estimation of arterial route travel time distribution with Markov chains," *Transportation Research Part B*, vol. 46, no. 10, pp. 1576–1590, 2012.
- [22] L. Du, S. Peeta, and Y. H. Kim, "An adaptive information fusion model to predict the short-term link travel time distribution in dynamic traffic networks," *Transportation Research Part B*, vol. 46, no. 1, pp. 235–252, 2012.
- [23] H. D. Sherali, J. Desai, and H. Rakha, "A discrete optimization approach for locating automatic vehicle Identification readers for the provision of roadway travel times," *Transportation Research Part B*, vol. 40, no. 10, pp. 857–871, 2006.

Research Article

Motor Vehicle Emission Modeling and Software Simulation Computing for Roundabout in Urban City

Haiwei Wang,¹ Huiying Wen,¹ Feng You,¹ Jianmin Xu,¹ and Hailin Kui²

¹ School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China

² College of Transportation, Jilin University, Changchun 130025, China

Correspondence should be addressed to Feng You; youfeng77@126.com

Received 16 July 2013; Revised 11 October 2013; Accepted 27 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Haiwei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In urban road traffic systems, roundabout is considered as one of the core traffic bottlenecks, which are also a core impact of vehicle emission and city environment. In this paper, we proposed a transport control and management method for solving traffic jam and reducing emission in roundabout. The platform of motor vehicle testing system and VSP-based emission model was established firstly. By using the topology chart of the roundabout and microsimulation software, we calculated the instantaneous emission rates of different vehicle and total vehicle emissions. We argued that Integration-Model, combining traffic simulation and vehicle emission, can be performed to calculate the instantaneous emission rates of different vehicle and total vehicle emissions at the roundabout. By contrasting the exhaust emissions result between no signal control and signal control in this area at the rush hour, it draws a conclusion that setting the optimizing signal control can effectively reduce the regional vehicle emission. The proposed approach has been submitted to a simulation and experiment that involved an environmental assessment in Satellite Square, a roundabout in medium city located in China. It has been verified that setting signal control with knowledge engineering and Integration-Model is a practical way for solving the traffic jams and environmental pollution.

1. Introduction

In the traditional traffic environmental management research field, in order to improve fuel efficiency and reduce emissions, varieties of static conservation and emissions reduction technologies have been widely studied and used, such as the static control testing technology of the vehicle fuel consumption and exhaust emissions on cars or engines [1]. Recently, researchers from different countries have begun to explore the relationship between the traffic activities and the vehicle fuel consumption and emissions. They focused on the study of reducing the vehicle fuel consumption and vehicle emissions by intelligent transportation information control. Compared to the conventional static control technology, such a new control strategy stands out for its significant characteristics on dynamic energy-saving, with focus on the dynamic management control of the vehicles and traffic flow. Practices have shown that this dynamic control strategy can

serve the purpose of energy conservation and emissions reduction in traffic system. Moreover, it is especially suitable for the economic construction in China at present.

Generally, vehicle emission models can be divided into the macroregional emission model, the medium emission model, and the microscopic emission model according to the applicable scale and function. Because of the difference in applicable scale, the emission model structures are also different [2]; these structures of the emission model are shown in Table 1.

As an important part in traffic networks, an intersection is a place where multiple traffic flows converge. Therefore, the intersection has always become the bottleneck in the traffic network. The vehicles which run through the intersection often have to slow down, stop, and then speed up and during this process more gas emissions will be produced in the intersection. Thus, it is significant to take an insight into the characteristics of vehicle exhaust emissions at intersections

TABLE 1: Classification of vehicle emission models.

Classification of emission models	Transportation model /data	Model parameter	Emission model pattern
Macroscopic emission model	Regional transportation model	Average speed (macroscopic parameter)	Regional emission (speed correction factor method)
Medium Emission model	The transport model based on equipment; behavior distribution pattern	Acceleration, speed, and v/c that are based on the traffic facilities (medium parameter)	The emissions of entire network based on the types of equipment
Microscopic emission model	Microscopic transport mode; Driving cycle mode	Vehicle operation data per second (microscopic parameter)	Vehicle emissions per second

over the entire traffic network. Moreover, there are also some relevant practical researches about reducing vehicle emission by signal control in foreign countries.

In this field, Tzeng and Chen [3] are among the foregoers who take in account the environmental factors in traditional traffic assignment, and they established the multiobjective traffic assignment model. Rilett and Benedek [4] used a simple two node two path network to evaluate the environmental improvement quantitatively by using the IVHS technology. Hallmark et al. [5] studied the effects on exhaust emissions by the method of intersection signal timing. Under the condition of different speed limits, Jackson et al. [6] analyzed the motor emissions on the highway. For the single point control of the intersections, Yoon et al. [7] calculated the optimal signal formula to reduce the queuing delay and exhaust emission effect. Pandian et al. [8] calculated and compared the traffic characteristics and the vehicle exhaust emissions between the ETC system and the conventional toll stations. Zhangb and Zietsmana [9] made a comparative analysis of the emission factor on these two strategies of vehicles idling in front of red light and the motor restart after flameout in the intersection.

In addition, some scholars used the existing emission models to calculate vehicle emissions directly and combined it with the city traffic signal control optimization. Hallmark and Guensler [10] used Measure model to prove the effect of the signal timing on the emissions of CO. Li et al. [11] used the computation formula of Mobile model and proposed an overall optimization scheme that uses the signal timing to reduce the emissions. In China, the optimization methods and the related researches are also conducted by Southeast University, Wuhan University of Technology, Beijing Jiaotong University, and Jilin University [12–18].

According to the study of vehicle emission test on the road and at the intersection both at home and abroad, it is evident that the researches on vehicle emission are studied extensively in recent years. At the same time, more attention is paid to the subject of how to reduce the vehicle exhaust pollution by traffic management and traffic control methods, especially by the method of traffic signal control. However, when it comes to the city road vehicle exhaust pollution problems, there seems to be a long way to go in theoretical analysis and methodology.

This paper is organized as follows. Following the introduction, the problem description of roundabout is presented in Section 2. Section 3 illustrates emission measure test of

city road and Section 4 illustrates the microscopic emission model for motor vehicle based on VSP. In the Section 5, we introduce the optimizing transportation management and control laws for roundabout intersection with low carbon. Finally, Section 6 concludes the paper.

2. The Problem Description of Roundabout: Congestion and Polluted

2.1. Traffic Characteristics of Roundabout. Changchun (China) is a city which owns a smaller number of overpasses, and most of the road segments are connected with various intersections. At present, with the improvement of people's living standard in modern society and the remarkable increase of motor vehicle volume, all kinds of road segments which are connected with intersections are overloaded. In particular some roundabouts (such as Xinmin square, People Square, and Satellite Square) have failed to satisfy the traffic demands, and serious traffic congestions often occurred in rush hours. And furthermore, these intersections also become the traffic accident black-spots. Once a traffic accident or traffic blocks occurs at such spots, it would turn into a large obstacle to the whole traffic system and cause serious traffic environment pollution, which cuts down the traffic capacity of road network continuously.

In this paper, a typical roundabout, Satellite Square Roundabout, as is shown in Figure 1, is selected as the study site. It is an important transport hub throughout the northsouth and eastwest areas. With a circle radius of 70 m, the roundabout is a landmark building in Changchun. And it is a large intersection of no traffic signal control, which means that traffic flows running through mingle together freely. The traffic characteristics in this intersection are manifested as follows.

- (1) The mixed traffic flow of pedestrians and vehicles is unbalanced in time and space distributions. Thus, traffic congestion occurs frequently on the intersection, especially at the two-way junction on the People Road.
- (2) According to a traffic survey of each intersection, the traffic flow through the satellite square intersection has far exceeded the threshold of the intersection capacity in rush hours. It implies that at the present the original design capacity of the intersection has



FIGURE 1: Satellite square roundabout aerial in Changchun.

failed to meet the growth of traffic demand in Changchun city.

- (3) Taxi is a very important means of transportation in Changchun and occupies a large share of traffic flows. Furthermore, the operation and management mode of taxi enhances the probability of illegal driving activities. Also, the centralized distribution of bus stops near the Satellite Roundabout leads to large passenger flow and stagnation of the bus on the bus stop, which also exacerbates traffic conditions there.
- (4) In the area of Satellite Roundabout, the coordination of intersection traffic flow is so poor, and it lacks drive auxiliary facilities, both of which contribute to the anomalies traffic flows on a certain crossroad. Moreover, unharmonious traffic flow of the intersection traffic eventually leads to serious traffic congestion and traffic pollution at the roundabout during the rush hour.

Therefore, this typical intersection (suffering from big saturation, serious traffic congestion, and serious motor exhaust pollution) serves the purpose of traffic pollution reduction research perfectly.

2.2. Traffic Control Scheme on the Roundabout. The key point of the intersection signal control is to assign a right-of-way for vehicles by signal light. Compared to intersections without signal control, stop lines are added on the imports of different lanes at the roundabout and signal lights with applicable signal timing are set up for the purpose of controlling vehicles running through. In this way, the traffic flows from different directions are separated in time dimension, which makes it possible to eliminate traffic congestion and problems of traffic pollution existing in the no signal control roundabout. The signal set analysis is shown in Figure 2. In a word, by optimizing the allocation of time and space resource, both the intersection traffic capacity and traffic order will be improved. Correspondingly, the traffic congestion and jams will decrease and the traffic efficiency will also be promoted.

3. Emission Measure Test of City Road

3.1. On-Board Emission Test System. In this paper, the on-board emission testing platform is chosen as the method for

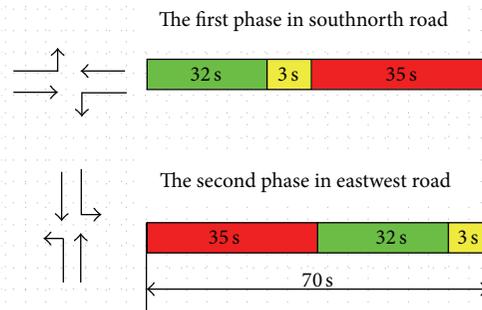


FIGURE 2: Signal set analysis.

test and Changchun city roads with different road grades are selected as the study object. Then, we can get the driving state of test vehicle going on the actual road and the corresponding emission test data per second, so as to pave the way for traffic area emissions analysis on the different traffic situations. In the experiment, we chose the portable car emissions measurement instrument OEM 2100 as the instrument for emission measurement. OEM 2100 is produced by Clean Air Technologies Inc of America and has passed the New York environmental protection department test in EPA's nation fuel and vehicle emission laboratory. For speed measurement, we chose TN-200GPS receiver (produced by Rayming and notebook computer) as real-time road speed measurement system. Its precision can reach 1 meter, and it can be placed at any position in the car, so it is a suitable subsidiary measuring instrument that measures vehicle driving condition on road motor experiment. Besides, we used GPS mapping software made by Fugawi Company to record vehicle road running parameters gathered by GPS. Actual on-road vehicle speed parameter, distance parameter, and other parameters are measured, and these data will be recorded in the computer. Figure 3 is the operation panel and the installation position in experiment of OEM-2100. Figure 4 is the structure diagram of the vehicle emission testing system.

3.2. Experimental Vehicle Selection. In accordance with “the measurement method of air pollution from motor vehicle



FIGURE 3: The operation panel and the installation position in experiment of OEM 2100.

TABLE 2: The test vehicle parameter statistics.

Vehicle type	Model	Productive life	Mileage	Engine displacement	Fuel supply
Light vehicle	Jetta Gix	2001	42000 km	1.6 L	Electric injection
	Red flag CA7201	2004	35000 km	2.0 L	Electric injection
Midsize car	Sea lion Ruichi	2002	28000 km	2.4 L	Electric injection
	Haf H3	2008	18000 km	2.4 L	Electric injection

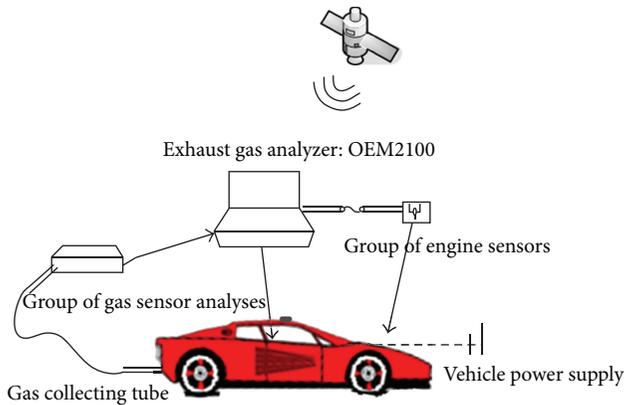


FIGURE 4: The structure of on-board emission test system.

emissions in urban” which is formulated by the State Environmental Protection Administration, research group conducted the emissions experiments of the on-road vehicle and the emissions calculation [19]. All test vehicles participated in the annual inspection on time per year and passed the test of pollution emissions. So the requirements of the road emission experiment are satisfied and the data we obtain are receivable. Besides, bus transportation is also included in the emission calculation. According to the measurement method, bus transportation (diesel vehicle) belongs to heavy-duty vehicle. However, our test system can only measure emission data of light-duty gasoline vehicles, so there is no heavy-duty vehicle in the test experiment. As a result, we choose to refer to overseas and domestic research status when we set up emission model to calculate vehicle emissions.

The statistics data of the technology parameter of the test vehicle are as follows in the Table 2.

3.3. *Experimental Road Selection.* The road network pattern in Changchun city is a combination of a radioactive and irregular chessboard type, and commercial areas and administrative areas are both located in the center of city. We choose these typical city roads for experiment, as shown in Figure 5, for the reason that they bear the major traffic flows and are distinguished from each other in some aspects.

Path Line One. People Street (trunk road), Minkang Road (secondary trunk road), Jiefang Road (express way), Anda Street (secondary trunk road), Xi’an Road (express way), Changchun Street (express way), Yatai Street (Trunk Road), Nanhu Road (express way).

Path Line Two. North People Street, South People Street, Xi’an Road, Changchun Street, Minkang Road. Each road drives into People’s Square Roundabout and each road drives out of People’s Square Roundabout, and then drive around the People’s Square Roundabout on the inner and outer ring road at normal speed.

3.4. *The Emission Experiment Data Acquisition in Real-Time Road.* By processing and matching experiment data, we can obtain the final database which is established from both the emission experiment data and vehicle driving data.

The experimental data processing flow chart is as shown in Figure 6.

4. Microscopic Emission Model for Motor Vehicle Based on VSP

4.1. *Calculation Method of Vehicle Specific Power (VSP).* In order to describe and calculate accurately the on-road motor vehicle emission pollution, Jose had proposed the concept of Vehicle Specific Power (VSP), he pointed out that driving

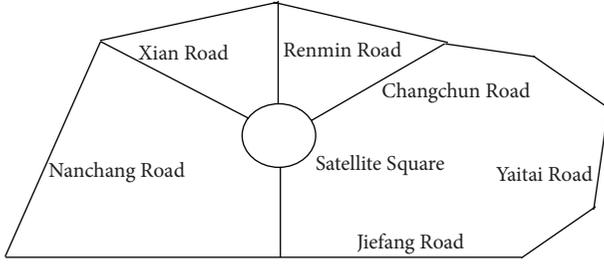


FIGURE 5: The road route of emission experiments.

conditions of motor vehicle have serious impact on these vehicle emission pollution.

According to the literature [20–23], vehicle specific power (VSP), with the unit of kw/t, is also referred to as instantaneous power per quality of motor vehicle. It consists of two parts. One is the output power of engine which is used for overcoming the friction from wheel rotation resistance, aerodynamics resistance and increasing kinetic, and potential energy of motor vehicle. And the other is transmission mechanical power loss caused by internal fingernail friction during transmission. As Figure 7 is the VSP of vehicle.

Mathematical expression (Jose, 1999a) is shown in [24]

$$\text{VSP} = \frac{(d(EK + PE)/dt) + Fv + F_A v + F_{v_{ir}}}{m}; \quad (1)$$

formula expansion is

$$\begin{aligned} & \left(\left(\frac{d}{dt} \right) (0.5 \times m (1 + \varepsilon_i) v^2 + mgh) \right. \\ & \quad \left. + C_R mgv + 0.5 \times \rho_a C_D A (v + v_m)^2 v + C_i mgv \right) \\ & \quad \times (m)^{-1} \\ & = v [a (1 + \varepsilon_i) + g \times \sin \theta + g \times C_R] \\ & \quad + 0.5 \times \rho_a \frac{C_D A}{m} (v + v_m)^2 v + C_i gv, \end{aligned} \quad (2)$$

where v is vehicle speed (in m/s); m is vehicle weight (in kg); a is vehicle acceleration (in m/s^2); ε_i is quality factor (nondimensional); h is vehicle altitude (in m); θ is road grade; g is gravitational acceleration, (9.81 m/s^2); C_D is drag coefficient (nondimensional); C_R is rolling resistance coefficient (nondimensional), which is connected with the type of tire and pavement materials; A is vehicle windshield area (in m^2); ρ_a is ambient air density (1.207 kg/m^3 at 20 centigrade); v_m is wind speed; C_i is internal friction coefficient (nondimensional), which is negligible.

Put the data measured directly into the formula, after simplifying the formula, we will get the following general formula of vehicle specific power:

$$\text{VSP} = v [1.1a + 9.8(as) + 0.132] + 0.000302v^3, \quad (3)$$

where $s = \tan(\sin \theta)$ is the gradient (nondimensional), which will be perceived as quotient of vehicle lift height and slope length when calculated.

TABLE 3: Ranges of VSP.

VSP (kw/t)	Range	Frequency
$\text{VSP} < -10$	1	0.035
$-10 \leq \text{VSP} < 2$	2	0.061
$-2 \leq \text{VSP} < 0$	3	0.074
$0 = \text{VSP}$	4	0.294
$1 \leq \text{VSP} < 3$	5	0.212
$3 \leq \text{VSP} < 5$	6	0.206
$5 \leq \text{VSP} < 9$	7	0.037
$9 \leq \text{VSP} < 13$	8	0.016
$13 \leq \text{VSP} < 17$	9	0.016
$17 \leq \text{VSP} < 20$	10	0.013
$20 \leq \text{VSP}$	11	0.037

In order to show the load conditions of motor vehicles on-road, it is necessary to depict the driving conditions of moving vehicles precisely. Firstly, we figure out the correlation between instantaneous gas emissions of on-road vehicles and VSP and keep a track of driving conditions of on-road vehicle. Then on the basis of statistical probability value of VSP, we divided vehicle specific Power into eleven ranges [25–27], as shown in the Table 3. In the process of dividing, we adhere to the following two principles.

- (1) When there are great fluctuations between different ranges, we will get these two different vehicle specific power frequency values into different ranges in order to calculate the different ranges of VSP.
- (2) Dividing the vehicle specific power ranges as much as possible can help in reducing standard deviation of values of each range, as a result of which the computing result of the emissions will be more accurate.

4.2. The Microemission Model under VSP Model. Traditionally, researchers may use the average emission factors to calculate the actual pollutant emissions of road vehicles. However, such conventional methods generate large errors. Therefore, we now use actual road test to accurately describe and calculate the actual road vehicle emissions of gaseous pollutants situation.

Jose proposed road vehicle specific power for the first time, and he pointed out that emissions are affected by the probability of time throughout average speed and instantaneous speed of the vehicle and acceleration, and also affected by vehicle categories on-road.

According to the characteristics of the actual emissions of vehicles transient analysis, as is shown in Figure 8, we can know that for the actual road running vehicles, the instantaneous rate of discharge is significantly related to vehicle speed and specific power [28]. This correlation is even more evident in certain conditions of the running vehicles. According to the aforementioned specific power ranges, the research group established an actual road vehicles microscopic emissions model based on a large number of experimental data.

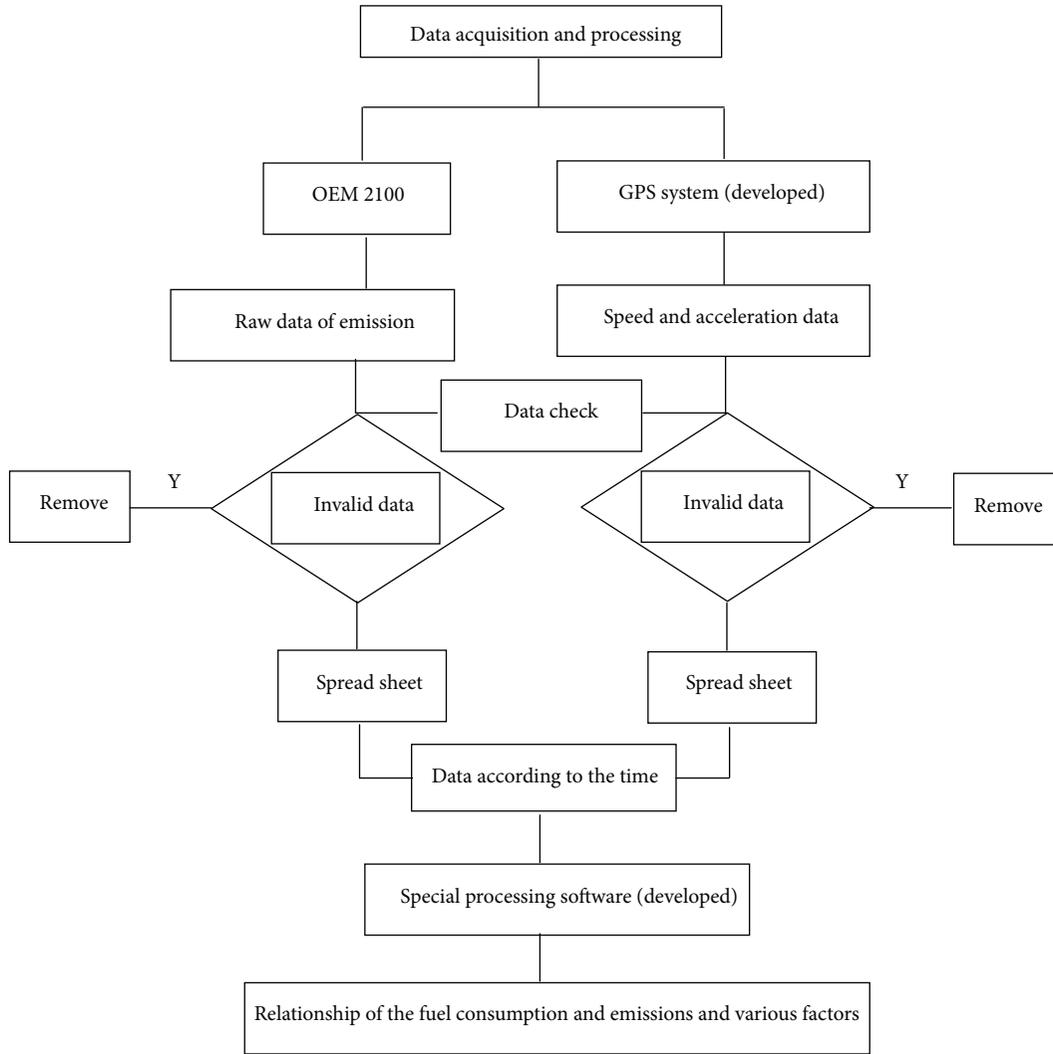


FIGURE 6: The processing flowchart for experimental data.

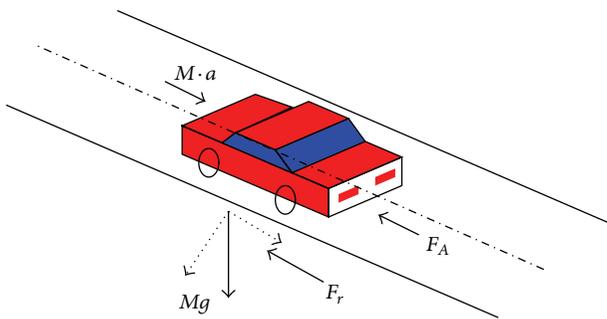


FIGURE 7: VSP of vehicle.

By analyzing driving condition data of on-road vehicle, we can get the required data of vehicle real-time speed and vehicle acceleration. Then, we import the data into VSP formula to generate the second-by-second data of VSP. In the next stage, these VSP values will be divided into different ranges, which then will be analyzed to develop a database.

According to the data characteristics, we could get the related statistical regular pattern in the way of regression analysis. On the basis of different probability distribution of the same VSP range, we calculate these 3 gas instantaneous vehicle emissions quality in each VSP range, as shown in Tables 4, 5, and 6. Thus, the relationships between 11 VSP ranges and the 3 gas instantaneous vehicle emissions quality selected are formed, respectively. According to the analysis, VSP mostly spread over the ranges of $-3 \sim 7$, as shown in Figures 9, and 10.

5. Optimizing Transportation Management and Control for Roundabout Intersection with Low Carbon

In this section, a typical intersection suffering from traffic congestion is chosen as the object of the research. VISSIM (dynamic traffic simulation software) is used to simulate the real-time operation state. Then, the combining of VISSIM

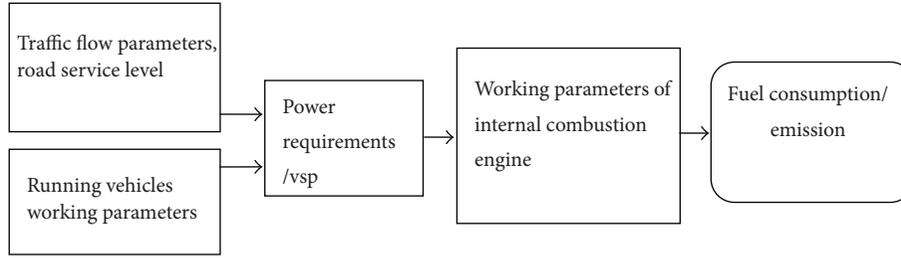


FIGURE 8: Microscopic emissions model flow of vehicles.

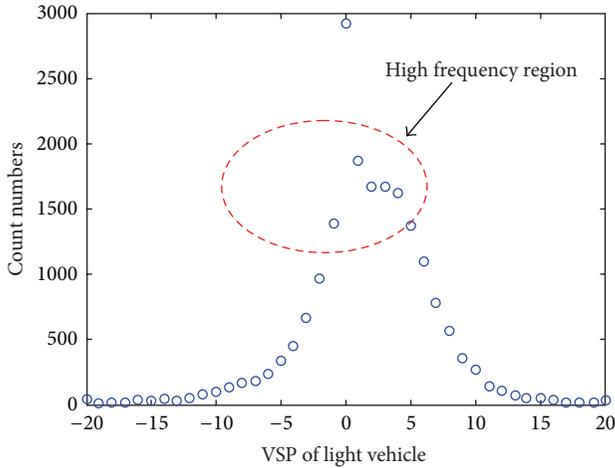


FIGURE 9: Specific power frequency distribution of light vehicles.

TABLE 4: Rate of instantaneous vehicle emission quality under light vehicles specific power ranges.

Ranges of VSP	Range of values	CO (mg/s)	HC (mg/s)	NOx (mg/s)
1	VSP < -10	1.9037	0.0659	0.3433
2	-10 ≤ VSP < 2	2.0903	0.1032	0.5041
3	-2 ≤ VSP < 0	2.5399	0.1589	0.5559
4	0 = VSP	1.9546	0.2456	0.5889
5	1 ≤ VSP < 3	2.3780	0.1938	0.7026
6	3 ≤ VSP < 5	2.2679	0.3001	0.8429
7	5 ≤ VSP < 9	2.9534	0.3329	1.0572
8	9 ≤ VSP < 13	4.0731	0.4380	1.1759
9	13 ≤ VSP < 17	3.9956	0.5469	1.3584
10	17 ≤ VSP < 20	4.5140	0.5177	1.4520
11	20 ≤ VSP	4.2339	0.5065	1.4531

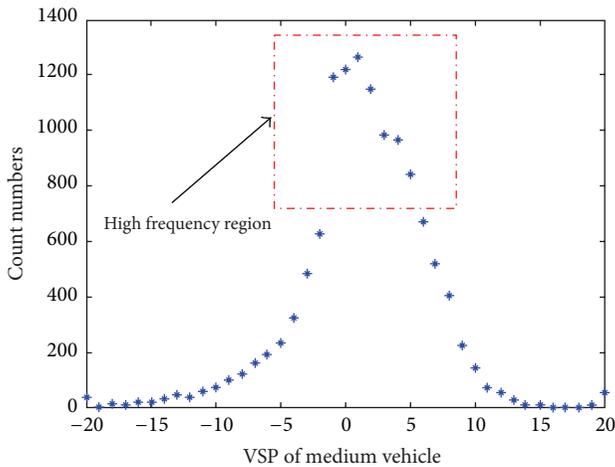


FIGURE 10: Specific power frequency distribution of middle vehicles.

TABLE 5: Rate of instantaneous vehicle emission quality under middle-sized car vehicle specific power ranges.

Ranges of VSP	Range of values	CO (mg/s)	HC (mg/s)	NOx (mg/s)
1	VSP < -10	1.7339	0.2288	0.1237
2	-10 ≤ VSP < 2	2.3420	0.2641	0.0899
3	-2 ≤ VSP < 0	2.5245	0.2679	0.0767
4	0 = VSP	3.1301	0.3124	0.1695
5	1 ≤ VSP < 3	3.4744	0.2699	0.1799
6	3 ≤ VSP < 5	3.9401	0.3187	0.2536
7	5 ≤ VSP < 9	5.0478	0.4341	0.3924
8	9 ≤ VSP < 13	5.7706	0.5351	0.5249
9	13 ≤ VSP < 17	5.5012	0.6296	0.5877
10	17 ≤ VSP < 20	4.4231	0.6101	0.7433
11	20 ≤ VSP	4.0965	0.5921	0.7249

with microscopic emission model can provide the necessary support for the emission of a certain traffic area. By calculating the vehicle driving emission of the chosen traffic area, we can find out approaches to improve the traffic situation. By the way, it also lays a foundation for energy conservation and emissions reduction, which can be achieved by optimizing the existing traffic control scheme.

5.1. The Establishment of the Roundabout Simulation Model.

To establish the roundabout simulation model, it is necessary to know the characteristics of actual traffic area road data, traffic flow, and a specific time period. As shown in Figure 11 and Table 7. Then, according to the basic steps of simulation model, the aforementioned data are added to the model as parameters.

TABLE 6: Rate of instantaneous vehicle emission quality under bus specific power ranges.

Ranges of VSP	Range of values	CO (mg/s)	HC (mg/s)	NOx (mg/s)
1	VSP < -10	0.3428	0.8936	2.3350
2	-10 ≤ VSP < 2	0.8659	1.2138	3.8250
3	-2 ≤ VSP < 0	3.3380	1.7685	12.293
4	0 = VSP	4.6788	1.9398	19.076
5	1 ≤ VSP < 3	6.9784	1.9929	22.7980
6	3 ≤ VSP < 5	7.7868	1.9834	23.8749
7	5 ≤ VSP < 9	8.5160	1.9482	24.209
8	9 ≤ VSP < 13	7.9499	2.0052	26.620
9	13 ≤ VSP < 17	6.5076	2.1758	27.345
10	17 ≤ VSP < 20	6.2015	2.1810	30.827
11	20 ≤ VSP	5.2770	2.2033	32.970

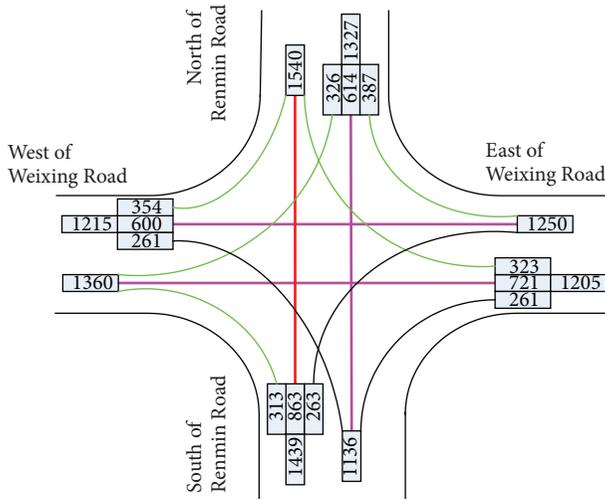


FIGURE 11: The traffic flow graph of intersection at peak hour.

On the basis of the traffic survey data, the motor vehicle flow at the intersection is about 5000 pcu/h. Additionally, the intersection is located at the confluence of two trunk roads and pedestrians flow is large. As a result, the capability of entrance roads without intersection controls cannot satisfy the increasing traffic demand, which becomes the traffic bottlenecks.

5.2. Vehicle Emissions Compute and Simulation Model for Roundabout. After establishing the typical intersection traffic area by microscopic traffic simulation software, we can obtain the vehicle's instantaneous speed and acceleration in any single area of simulation. Then, by combining with micro emissions model based on VSP, we can calculate the motor vehicle emissions of intersection area and car emissions share rate, which could serve the purpose of the evaluation of traffic environment.

We get the integer data based on VSP and put them into different 11 VSP ranges for statistical compute. With the

emission model, we get the average of the transient emissions rate from different types of vehicles in traffic simulation area, which is shown in Table 8. The following table shows the average rate of instantaneous vehicle emission quality.

Depending on the average rate of instantaneous vehicle emission quality as well as the total running time of the driving vehicle in the simulation, we can calculate the three kinds of different vehicle pollutant emissions in the morning rush for one hour.

For a particular intersection traffic simulation area, the calculation formula of the total toxic emission is as follows:

$$G_{Ej} = \sum_{i=1}^m \sum_{j=1}^n G(p_{ij}) M(p_{ij}), \quad (4)$$

where G_{Ej} is total vehicle emissions of a certain vehicle; J represents three vehicle models (light, medium, and heavy); i is for different VSP ranges; p_{ij} is for the j models in VSP range of the i ; $G(p_{ij})$ means the cumulative time s of the j models in VSP range i ; $M(p_{ij})$ is for quality of emissions of the j models in VSP range i , g/s.

The total emissions in the whole area can be calculated by the following formula:

$$A = \sum_{j=1}^n G_{Ej}, \quad (5)$$

A is the total regional emissions. Table 9 is total gas pollutants emissions in an hour in simulation area.

5.3. Gas Emission Calculate Result of the Roundabout under Traffic Signal Control. Roundabout is an important part of city roads, and its capacity largely depends on how well an intersection traffic flow runs. Traffic capacity of Roundabout is limited by the weaving section capacity. In the case of a traffic increase in particular, when the traffic flow gets close to or exceeds the traffic capacity of roundabout, it can cause ring road traffic congestion. And then the vehicle emissions pollution exceeds amount in the intersection area.

To achieve a reasonable allocation on time and space of traffic flow in traffic areas by setting up traffic signal light and signal timing optimization at the entrance of the roundabout, we can not only balance the entrance traffic flow but also eliminate the traffic congestion at the entrance, which reduces the vehicle delay in the traffic area. As shown in Figure 12.

By adding traffic signal lights and optimizing signal timing, we get the instantaneous speed and acceleration of each single vehicle from simulation software and calculate the single vehicle's VSP. Then we put VSP data into model to calculate so that we can get instantaneous emission rate and total quality of these three kinds of gas emissions in one hour. We put these data to calculate instantaneous gas emissions at the roundabout with traffic signals for different vehicle models and calculate the whole intersection of gas emissions. Results are shown in Tables 10 and 11.

5.4. Comparison of Environmental Protection Quota for Roundabout under Signal Control or under No Signal Control. By setting up traffic signal light and optimizing the signal timing at the roundabout, it can be found that the

TABLE 7: Intersection road structure and traffic data.

Entrance name	Lane number * width (m)	Steering vehicle volume percentage (%)			Traffic flow (Pcu/h)	Midsize cars
		Left	Straight	Right		
People's avenue north	3 × 3.2	21	56	23	1540	13%
People's avenue south	3 × 3.2	23	54	23	1136	10%
Satellite Road left	3 × 2.6	24	53	23	1360	20%
Satellite Road east	3 × 2.6	21	48	31	1250	13%

TABLE 8: Average of rate of instantaneous vehicle emission quality for different models.

Gas (discharge rate)	Models			
	Light vehicle	Midsize vehicle	Bus	Sum (mg/s)
NOx (mg/s)	0.6066	0.2325	21.029	21.8681
CO (mg/s)	2.3011	3.3060	4.0993	9.7064
HC (mg/s)	0.2400	0.3706	1.9075	2.5181

TABLE 9: Total gas pollutants emissions in an hour in simulation area.

Gas pollutants (kg/h)	Models			
	Light vehicle	Midsize vehicle	Bus	Sum (kg/h)
NOx	9.1717	0.837	2.2711	12.2798
CO	34.7926	11.9016	0.4427	47.1369
HC	3.6288	1.3341	0.2060	5.1689
Vehicle emissions	47.5931	14.0727	2.9198	64.5856

TABLE 10: Instantaneous gas emissions at the roundabout under traffic signals for different vehicle models.

Gas (mg/s)	Models		
	Light vehicle	Midsize vehicle	Bus
CO	2.1123	3.0012	3.5688
HC	0.2036	0.3298	1.6970
NOx	0.5798	0.2175	19.1548

TABLE 11: Gas emissions at the roundabout under traffic signals.

Gas pollutant	Emissions from all sources (kg/h)	Percentage (%)
CO	42.3289	73.12
HC	10.7791	18.62
NOx	4.7843	8.26

vehicle under the traffic control is in a relatively steady state, as shown in the simulation output of road vehicle running status, we can get the comparisons of different vehicle's instantaneous gas emission rate in Tables 12, 13 and 14. During the rush hour, intersection congestion rarely appears at the entrance, which means that the traffic flow is in smooth running state, and the times of acceleration and deceleration of a single vehicle reduce which makes the transient emissions rate of gas pollutant decline obviously.

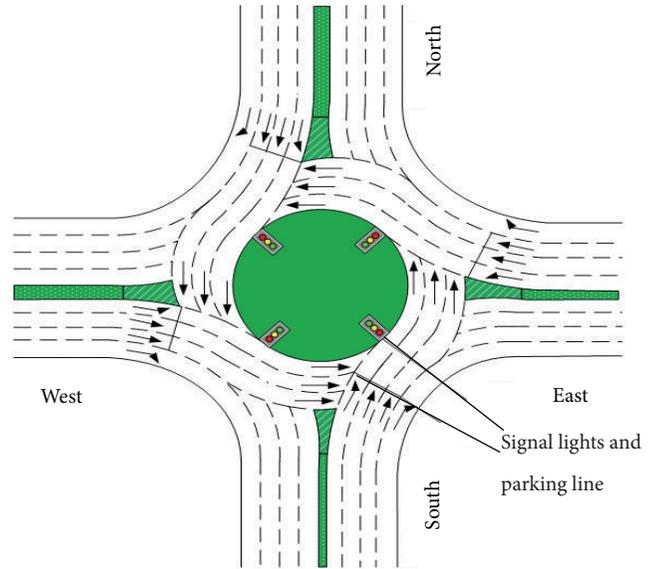


FIGURE 12: Setting signals at the roundabout.

TABLE 12: Comparison of light vehicle's instantaneous gas emissions rate.

Control	Gas (mg/s)		
	CO	HC	NOx
In control	2.3011	0.2400	0.6066
No control	2.1123	0.2036	0.5798
Improved percentage	8.20%	15.16%	4.42%

5.5. Contrastive Analysis of the Traffic Flow Indicators at the Roundabout Based on Simulation. We get some important evaluation indexes by VISSIM 3.60. Because at the beginning of the simulation the traffic needs a certain amount of time to stabilize, this period of time must be excluded from the simulation time of 3600s when we make the evaluation for a variety of traffic parameters [17, 18, 29]. That means we make the data statistics with the output data at the time of a relatively stable traffic flow. To get a better understanding of the test, we compare two kinds of states at the roundabout and the evaluation indexes are shown in the following Tables 15 and 16, and Figure 13.

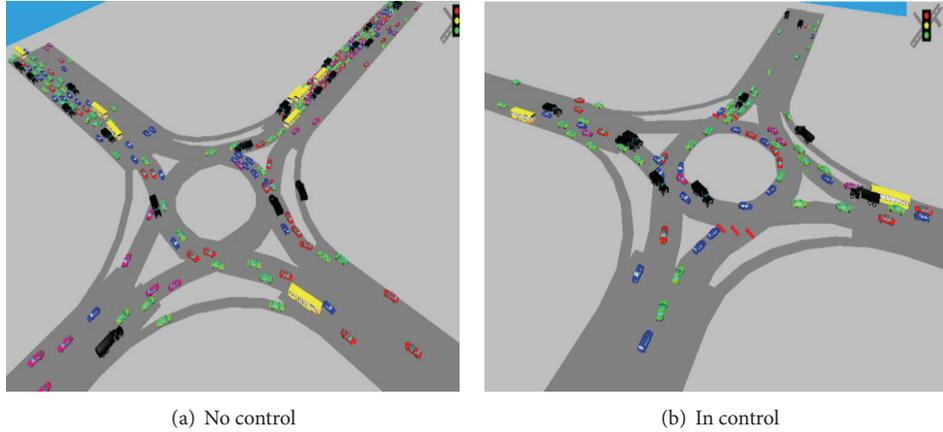


FIGURE 13: Roundabout with control VS roundabout without control.

TABLE 13: Comparison of midsize vehicle's instantaneous gas emissions rate.

Control	Gas (mg/s)		
	CO	HC	NOx
No control	3.306	0.3706	0.2325
In control	3.0012	0.3298	0.2175
Improved percentage	9.22%	11.00%	6.45%

TABLE 14: Comparison of bus's instantaneous gas emissions rate.

Control	Gas (mg/s)		
	CO	HC	NOx
No control	4.0993	1.9075	21.0290
In control	3.5688	1.6970	19.1548
Improved percentage	12.94%	11.04%	8.91%

6. Conclusion

Choose the car exhaust gas detection equipment (MONTATA OEM 2100) certificated by the U.S. environmental protection agency (EPA) and Vehicle driving condition monitor system based on GPS to construct the motor-mounted data acquisition system for emission test and make data acquisition for on-road vehicle driving condition and vehicle exhaust second-by second test. Through theoretical analysis of VSP and a large number of theoretical researches both at home and abroad, we use VSP which includes on-road vehicle instantaneous velocity, acceleration, and specific power of road slope as a modeling basis.

We use the database from the experiments and basic principle of vehicles specific power and the reference standard of the on-road driving situation to get the motor vehicle actual road micro emissions model based on three different vehicle models. At the same time, we get rate of instantaneous vehicle emission quality in the different VSP ranges. Thus, combined with the microsimulation model we can evaluate the traffic environment in the certain traffic region based on the Microscopic emission model.

TABLE 15: Comparison of traffic average volume.

Control	Models		
	Light vehicle (km/h)	Midsize vehicle (km/h)	Bus (km/h)
No control	12.36	12.77	8.90
In control	14.11	14.94	9.83
Percentage (%)	14.16	16.99	10.45

TABLE 16: Comparison of traffic operation.

Control	Indicators		
	Average delay (s/veh)	Average number of parking (times/veh)	Saturation
No control	91.4	4.69	1.42
In control	39.7	2.57	0.83
Improved percentage	56.5%	45.2%	41.5%

The results of the experiment show that under the traffic signal control the total gas emissions of CO, HC, and NOX for the same flow all reduced but in different degrees, with CO by 10.20%, HC by 12.22%, and NOX by 7.44%. When it comes to pollutant instantaneous gas emissions rate of different vehicle models, we can find that HC gas emissions show the biggest drop in light vehicle, and the NOX gas emissions reduce most in bus. In summary, it shows that with the increment of traffic flow, setting traffic signal control at the roundabout can reduce traffic pollution emissions of the region effectively. In addition, when putting optimized timing parameter into the simulation software VISSIM, we got the evaluation parameters of traffic flow, including the average delay and average number of parking, saturation, and so forth. By comparing the simulation results (which are mainly related to traffic flow and traffic environment evaluation) under both conditions (with signal control and without signal control at the roundabout), we may deduce that under the traffic signal control, whether there is a rush hour or not, both the instantaneous vehicle emission quality and total traffic gas pollutant emissions are reduced to a certain extent.

At the same time, corresponding traffic flow status is also improved and regional traffic congestion is alleviated.

In summary, this paper provided a new idea and technology investigation about traffic environment evaluation for a particular area under traffic control.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (Granted nos. 51378222, 51108192 and 51208500), Chinese Postdoctoral Science Foundation (the Granted no. 2012M521824 and no. 2013T60904), and the Fundamental Research Funds for the Central University of China (Granted no.: 2012ZZ0100). The first author would like to appreciate Dr. Feng You and Professor Wen Huiying, for the valuable discussions to improve the quality and presentation of the paper.

References

- [1] H. Liu and M. Barth, "Identifying the effect of vehicle operating history on vehicle running emissions," *Atmospheric Environment*, vol. 2, no. 5, pp. 873–879, 2012.
- [2] S. H. Cousins, J. Garcia Bueno, and O. Palomares Coronado, "Powering or de-powering future vehicles to reach low carbon outcomes: the long term view 1930–2020," *Journal of Cleaner Production*, vol. 15, no. 11–12, pp. 1022–1031, 2007.
- [3] G.-H. Tzeng and C.-H. Chen, "Multiobjective decision making for traffic assignment," *IEEE Transactions on Engineering Management*, vol. 40, no. 2, pp. 180–187, 1993.
- [4] L. R. Rilett and C. M. Benedek, "Traffic assignment under environmental and equity objectives," *Transportation Research Record*, no. 1443, pp. 92–99, 1994.
- [5] S. L. Hallmark, I. Fomunung, R. Guensler, and W. Bachman, "Assessing impacts of improved signal timing as a transportation control measure using an activity-specific modeling approach," *Transportation Research Record*, vol. 1738, pp. 49–55, 2000.
- [6] E. Jackson, Y. Qu, B. A. Holmén, and L. Aultman-Hall, "Driver and road type effects on light-duty gas and particulate emissions," *Transportation Research Record*, vol. 1987, pp. 118–127, 2006.
- [7] S. H. Yoon, J. P. Cha, and C. S. Lee, "An investigation of the effects of spray angle and injection strategy on dimethyl ether (DME) combustion and exhaust emission characteristics in a common-rail diesel engine," *Fuel Processing Technology*, vol. 91, no. 11, pp. 1364–1372, 2010.
- [8] S. Pandian, S. Gokhale, and A. K. Ghoshal, "Evaluating effects of traffic and vehicle characteristics on vehicular emissions near traffic intersections," *Transportation Research D*, vol. 14, no. 3, pp. 180–196, 2009.
- [9] J. Lv, Y. Zhangb, and J. Zietsmana, "Investigating emission reduction benefit from intersection signal optimization," *Journal of Intelligent Transportation Systems*, vol. 17, no. 3, pp. 200–209, 2013.
- [10] S. L. Hallmark and R. Guensler, "Comparison of speed-acceleration profiles from field data with NETSIM output for modal air quality analysis of signalized intersections," *Transportation Research Record*, no. 1664, pp. 40–46, 1999.
- [11] X. Li, G. Li, S.-S. Pang, X. Yang, and J. Tian, "Signal timing of intersections using integrated optimization of traffic quality, emissions and fuel consumption: a note," *Transportation Research D*, vol. 9, no. 5, pp. 401–407, 2004.
- [12] K. Chen and L. Yu, "Microscopic traffic-emission simulation and case study for evaluation of traffic control strategies," *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, no. 1, pp. 93–99, 2007.
- [13] P. Jiao and H. Lu, "Study on evaluation of national economic benefit of urban traffic manager," *Journal of Highway and Transportation Research and Development*, vol. 9, article 32, 2005.
- [14] F. Mo, L. Yu, and G. H. Song, "An overview of portable emissions measurement system technologies and relevant research," *Vehicle & Power Technology*, vol. 4, article 13, 2006.
- [15] S.-X. Guo, L. Yu, and G. H. Song, "A comparative analysis of real-world emission factors and MOBILE6 predictive value for a heavy-duty diesel vehicle," *Safety and Environmental Engineering*, vol. 14, no. 2, pp. 17–21, 2007.
- [16] K. Chen and L. Yu, "Microscopic traffic-emission simulation and case study for evaluation of traffic control strategies," *Journal of Transportation Systems Engineering and Information Technology*, vol. 7, no. 1, pp. 93–99, 2007.
- [17] W. Wang, X. Jiang, S. Xia, and Q. Cao, "Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation," *Safety Science*, vol. 48, no. 10, pp. 1248–1262, 2010.
- [18] C. Ding, W. Wang, X. Wang, and M. Baumann, "A neural network model for driver's lane-changing trajectory prediction in urban traffic flow," *Mathematical Problems in Engineering*, vol. 2013, Article ID 967358, 9 pages, 2013.
- [19] D. Wan, J. Xu, J. Zhang, X. Tong, and J. Hu, "Historical and projected emissions of major halocarbons in China," *Atmospheric Environment*, vol. 43, no. 36, pp. 5822–5829, 2009.
- [20] G. Song, L. Yu, and Z. Tu, "Distribution characteristics of vehicle-specific power on urban restricted-access roadways," *Journal of Transportation Engineering*, vol. 138, no. 2, pp. 202–209, 2011.
- [21] M. C. Coelho, H. C. Frey, N. M. Roupail, H. Zhai, and L. Pelkmans, "Assessing methods for comparing emissions from gasoline and diesel light-duty vehicles based on microscale measurements," *Transportation Research D*, vol. 14, no. 2, pp. 91–99, 2009.
- [22] P. Zavadinka, "Simulation of vehicle transport duty cycle with using of hydrostatic units control algorithm," in *Mechatronics*, pp. 395–402, Springer, 2012.
- [23] E. Demir, T. Bektaş, and G. Laporte, "A comparative analysis of several vehicle emission models for road freight transportation," *Transportation Research D*, vol. 16, no. 5, pp. 347–357, 2011.
- [24] R. Casati, V. Scheer, R. Vogt, and T. Benter, "Measurement of nucleation and soot mode particle emission from a diesel passenger car in real world and laboratory in situ dilution," *Atmospheric Environment*, vol. 41, no. 10, pp. 2125–2135, 2007.
- [25] Y. Wang, D. Guo, S. Li et al., "Research on urban intersection vehicle emission reduction based on microcosmic simulation," in *Proceedings of the 9th International Conference of Chinese Transportation Professionals*, pp. 2068–2074, Harbin, China, August 2009.

- [26] H. Huo, Y. Wu, and M. Wang, "Total versus urban: well-to-wheels assessment of criteria pollutant emissions from various vehicle/fuel systems," *Atmospheric Environment*, vol. 43, no. 10, pp. 1796–1804, 2009.
- [27] Y. Wang, D. Guo, S. Li, Y. Wang, and Z. Li, "Signal timing optimization simulation on urban road intersection based on vehicle emissions," in *Proceedings of the 8th International Conference of Chinese Logistics and Transportation Professionals—Logistics*, pp. 4762–4767, Chengdu, China, August 2008.
- [28] W. Wuhong, Z. Wei, and B. Herner, "Car-following safety algorithms based on adaptive cruise control strategies," in *Proceedings of the 5th International Symposium on Intelligent Systems and Informatics*, pp. 135–140, Subotica, Serbia, August 2007.
- [29] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.

Research Article

Deformation of Overlong Isolated Buildings Caused by Thermal and Concrete Shrinkage

Yu Dang and Ying-ke Liu

School of Civil Engineering, Lanzhou University of Technology, Lanzhou, Gansu Province 730050, China

Correspondence should be addressed to Yu Dang; 601363791@qq.com

Received 8 July 2013; Revised 3 September 2013; Accepted 21 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Y. Dang and Y.-k. Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Temperature variations and concrete shrinkage influence structural behavior by reducing the strength of materials and changing their thermal strain contributions. This problem is particularly important for isolated buildings that are characterized by large horizontal dimensions and are sensitive to thermal action and shrinkage. In this study, the measurement of an overlong isolated building shows that the deformations of some isolators exceed the allowed deviation during the construction phase because the building is completely exposed. These deformations are induced by climatic thermal changes and shrinkage effects and cause the complex dynamic behavior and instability of the structure. To ensure the safety of overlong isolated buildings, the structural stress and deformation caused by temperature variations and shrinkage effects are studied. A three-story frame model is developed, and the rule of deformation within isolated frame buildings is analyzed by the deformation distribution method. The theoretical predictions are consistent with the experimental measurements. Therefore, the theoretical model is used to predict the deformation of isolated buildings caused by temperature variations and shrinkage effects. For reinforced concrete frame isolated buildings, expansion joint distances are proposed according to different thermal design regions and heating design conditions.

1. Introduction

In a concrete structure, temperature variations and concrete shrinkage influence stress and strain distribution within the elements of the structure. Studies on thermal action and concrete shrinkage are important for isolated buildings with large horizontal dimensions. Given that large floor slabs have no expansion joints, temperature variation and concrete shrinkage generate stresses that must be considered by a designer. Moreover, large displacements of isolated layers are generally incompatible with the deformations allowed by the isolation system during the construction phase when the building is unprotected by thermal insulation. Thus, the influences of thermal actions and shrinkage effects are important in the first phase of a structure's life.

Temperature varies because of different reasons: seasonal changes in air temperature, daily solar radiation on the element surface, heat dissipation in the building, and hydration heat. In this study, the magnitude of thermal effects mainly depends on seasonal changes in air temperature because seasonal temperature variations cause overall

structural deformation, which cannot be reduced by thermal insulation.

The issue of thermal actions caused by temperature variation or concrete shrinkage in buildings, bridges, and other structures has been extensively studied by many researchers. These studies have mainly focused on the thermal response of buildings with long longitudinal axes [1, 2] or precast continuous-beam bridges with hollow cross sections [3]. Studies have also investigated the effect of concrete shrinkage [4]. However, the literature on isolated buildings because of temperature variations and concrete shrinkages is limited. Han and Lytton [4] studied a model of natural and lead rubber bearings on the seismic responses of a highway bridge. The model parameters for the isolation bearings were evaluated from experimental results at a low temperature (-20°C). The effect of modeling isolation bearings at low temperature was significantly observed in the responses. Razzaq et al. [5] investigated eight building seismic rubber bearings in fire tests that follow the ISO834 standard heating process. The effects of vertical load and fire duration on the failure mode and residual mechanical properties of the rubber bearings

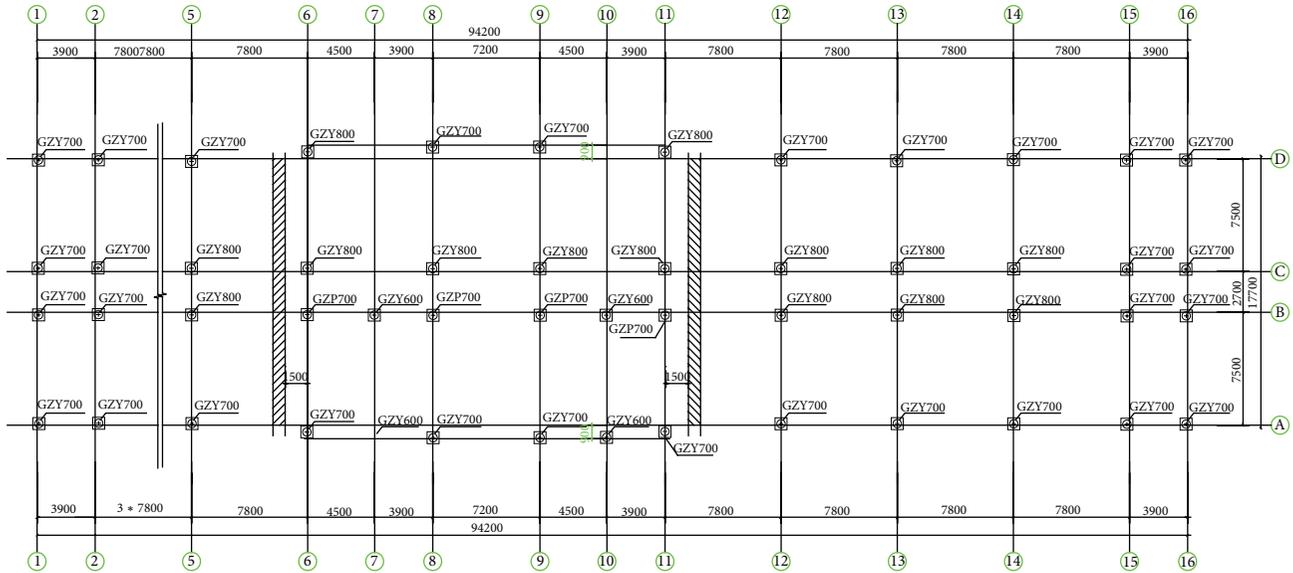


FIGURE 1: Plan of the isolated building with isolation bearing locations.

TABLE 1: Deformation of isolated bearings (mm).

Axis number	①-©	⑤-©	⑥-©	⑪-©	⑫-©	⑮-©
Type of isolated bearing	GZY700	GZY700	GZY800	GZY800	GZY700	GZY700
Deformation (mm)						
2°C	10	0	3	-2	-2	-10
4°C	9	0	0	-2	-1	-10

were analyzed. Wu et al. [6] and Kalpakidis and Constantinou [7] discussed the effects of heating on the behavior of lead rubber bearings. Kalpakidis and Constantinou [8] and Yakut and Yura [9] investigated the performance of isolated bearings at low temperatures. Huang et al. [10] created a mechanical model of isolated bearings according to temperature variations. These studies focused on the mechanical behavior of isolation bearings with thermal effects. However, few studies have investigated isolated building deformation caused by temperature variations and concrete shrinkages.

In this study, an experimental investigation is conducted to characterize the effects of temperature variations and concrete shrinkages in an overlong isolated building. To evaluate properly the deformation of an isolated building because of temperature variations and shrinkage, a three-story frame model is developed and the deformation distribution method is used for numerical analysis. Based on the test and analysis results, the distances of the expansion joints and poststrips of isolated buildings are proposed.

2. Measured Deformation of an Overlong Isolated Building

The isolated building measured is a seven-story Chinese hospital with a reinforced concrete frame structure. The building has a 94.2 m-long floor slab. The plan of the isolated building is shown in Figure 1. To avoid problems in dynamic



FIGURE 2: Deformation of the isolated bearing.

behavior, expansion joints are not installed. The isolated layer has two postcast strips.

The isolated bearings were installed in August 2010, when the average air temperature was 22°C. In March 2011, we found some isolation bearings with large deformations. Figure 2 shows an example of this case.

Six isolated bearings were selected to measure the extent of deformation (Table 1). We measured these bearings at separate times on April 1 and 3, 2011, when the air temperatures were 2 and 4°C, respectively.

Table 1 shows the measured deformation of the isolated bearings. A positive sign indicates that the deformation occurs along the longitudinal axis x , whereas a negative sign indicates that the deformation is opposite the axis x . The deformations of the ①-© and ⑮-© bearings are positive and

negative, respectively. We observe that the isolated building shortens. The deformations of the ①-© and ⑬-© bearings, which are located at the end of the building, are almost 10 mm. The deformations of the ⑥-© and ⑪-© bearings, which are located in the middle of the building, are almost 3 mm. The isolated bearings at the end of the building are more deformed than the isolated bearings in the middle of the building. Moreover, the deformations of the bearings at the end of the building exceed the maximum allowed deviation of the bearings; that is, the deformations exceeded the serviceability limit states during the construction period.

Isolated bearings are deformed by thermal action and concrete shrinkage because such bearings are installed at an average air temperature of 22°C but are measured at 2 and 4°C. Moreover, the postcast strips are placed in the superstructure and the influence of concrete shrinkage is considered.

3. Deformation of Isolated Bearings Caused by Thermal Changes and Concrete Shrinkage

3.1. Calculation of Temperature Difference. Isolated reinforced concrete structures are deformed by temperature variation and concrete shrinkage. The coefficient of heat conduction is low, and most buildings have thermal insulation. The magnitude of the thermal effects depends on seasonal climatic changes. Therefore, in this study, the temperature difference is defined as the seasonal temperature difference. To calculate the deformation, concrete shrinkage is considered equivalent to the temperature variation [11]. Temperature differences include seasonal climate thermal changes and equivalent shrinkage temperature variations:

$$t = t_1 + t_2, \quad (1)$$

where t_1 , t_2 are the seasonal thermal changes and equivalent shrinkage temperature variations, respectively. t_2 is expressed as follows:

$$t_2 = \frac{\varepsilon}{\alpha}, \quad (2)$$

where α is the coefficient of linear expansion according to concrete properties and is equal to $1 \times 10^{-5}/^\circ\text{C}$ [12]. ε is the deformation caused by concrete shrinkage and is mainly a function of concrete use, for example, the physical properties of the material used, the condition of concrete curing, and the local climatic condition. ε can be obtained as follows:

$$\varepsilon = 3.24 \times 10^{-4} (1 - e^{-0.01T}) M_1 M_2 \cdots M_n, \quad (3)$$

where T is the time of concrete curing (i.e., 180 d). M_1, M_2, \dots, M_n are the modification coefficients according to non-standard conditions such as cement type, water cement ratio, concrete curing time, environment humidity, reinforcement ratio, and wind speed [13].

3.2. Structural Model and Deformation Analysis. Isolated bearings are fixed between the base and superstructure. Therefore, isolation bearings are approximated by columns

with stiffness being equal to the horizontal stiffness of the isolated bearing. Structural strains and any stress caused by thermal action depend on the geometry and boundary conditions of the element considered. Given that column stiffness is inversely proportional to the cube of column height, structural strains and stresses decrease with increasing building height. At floors above the second floor, the resulting structural stresses are approximately equal to zero. Therefore, an isolated building can be simplified by a three-story building. When the superstructure is a frame, the model of an isolated building affected by thermal action is created by a three-floor frame, where the frame on the first floor stands for the isolated layer and the second and third floors are the first and second stories of the superstructure, respectively.

According to the deformation distribution method, beam deformation can be expressed as follows:

$$\delta = \alpha t L, \quad (4)$$

where α is the coefficient of the concrete linear expansion and t , L are the temperature difference and beam length, respectively.

The left and right ends of the deformation distribution coefficient of the beam are written as follows:

$$c_L = \frac{k_R}{k_R + k_L}, \quad c_R = \frac{k_L}{k_R + k_L}, \quad (5)$$

where k_R is the sum of the stiffness of the right columns and k_L is the sum of the stiffness of the left columns.

Thus, the right and left ends of the beam deformation can be obtained as follows:

$$\delta_R = c_R \delta, \quad \delta_L = c_L \delta. \quad (6)$$

The i th point deformation of the structure is expressed as follows:

$$\delta_i = \sum_{j=i}^n \delta_{Lj} - \sum_{r=1}^i \delta_{Rr}, \quad (7)$$

where $\sum_{j=1}^n \delta_{Lj}$ is the sum of the left ends of the beam deformation at the i th point and $\sum_{r=1}^i \delta_{Rr}$ is the sum of the right ends of the beam deformation at the i th point.

3.3. Numerical Example and Analysis. The measured isolated building is selected as the example for numerical analysis. For the isolated building, each story has a height of 3.3 m, except the first story, which is 4.2 m high. The sectional dimension of the beam in the isolated layer is 400 × 800 mm; the sectional dimension of the frame columns is 600 × 600 mm; the types of the isolated bearings are GZY700 and GZY800; the yield displacements of the isolated bearings are 11 mm to 12 mm; the reinforcement ratio is 1.5%; the grade of the concrete strength is C30; the grade of rebar is HRB335; the common cement grade is 425; the water cement ratio is 0.5; the ratio of cement paste to total weight is 25%; the initial curing time of concrete is 14 d; the local environmental humidity is 40%; and the temperature is 22.2°C to 2°C. The building is

TABLE 2: Column stiffness.

Column	GZY700	1~9	2~10	3~11, 4~12, 5~13	6~14	7~15
K (kN/m)	1034	7074	15580	18243	17812	18313
Column	GZY800	9~17	10~18	11~19, 12~20, 13~21	14~22	15~23
K (kN/m)	1238	11800	24346	27204	26470	27325

TABLE 3: Node displacements of the isolated layer.

Nodes	1	2	3	4	5	6	7
Analyzed displacement	20.0	18.1	14.8	11.5	8.2	4.8	1.5
Measured displacement	10	—	—	—	0	3	—

exposed without thermal insulation during construction. The building, except the isolated bearings, is protected by thermal insulation and heating measures during the service period.

Given that the isolated building is symmetric, the building can be simplified in half. The model of the isolated building affected by thermal action and concrete shrinkage is shown in Figure 3, where the equivalent stiffness of the isolated bearing is defined as the yield stiffness. Table 2 shows the column stiffness, which consists of the equivalent stiffness of the isolated bearings and the actual stiffness of the columns.

Seasonal temperature change can be represented as follows:

$$t_1 = 22.2^\circ\text{C} - (2^\circ\text{C}) = 20.2^\circ\text{C}. \quad (8)$$

Consider the following link parameters: time, cement type, cement fineness ratio, aggregate condition, water cement ratio, initial curing time, environment humidity, reinforcement ratio, and beam sections. The equivalent shrinkage temperature variation is expressed as follows:

$$t_2 = \frac{\varepsilon}{\alpha} = \frac{1.9371 \times 10^{-4}}{1 \times 10^{-5}} = 19.4^\circ\text{C}. \quad (9)$$

Thus, we obtain the temperature difference as follows:

$$t = t_1 + t_2 = 20.2 + 19.4 = 39.6^\circ\text{C}. \quad (10)$$

The deformation distribution method is applied to obtain the deformation of the building affected by thermal changes and concrete shrinkage. The deformation of the half building is shown in Figure 4. The deformation in the middle of the structure is zero, whereas the deformations increase with the increasing distance between elements. An increasing symmetry axis maximizes the deformations at the end of the building. Moreover, deformation is predominant on the first floor, that is, the isolated layer, and the deformations of the second floor are almost equal to that of the third floor; that is, the deformation of the superstructure is nearly uniform.

Table 3 shows the displacement of the isolated bearing with measured data and theoretical results. The theoretical results are higher than the measured data because the postcast strips are disregarded in the numerical analysis. The numerical result shows a safe value to ensure that the deformation of the isolated building is caused by thermal changes and

concrete shrinkage. Thus, the theoretical results can be represented as the deformation rule of isolated buildings affected by thermal change and concrete shrinkage.

Based on the analysis, several key issues are presented as follows.

- (1) The drift of isolated buildings affected by thermal changes and shrinkage is concentrated on the isolated layer, whereas the deformation of the superstructure is nearly equal.
- (2) The isolated bearings at the end of the building have maximum deformation. The deformations of the isolated bearings are reduced by shortening the distance between the bearings and the symmetric axis. If the temperature difference is large and the concrete is poorly cured, the deformations of the bearings at the end of the building will exceed the maximum allowed deviations of the bearings. Thus, thermal changes and concrete shrinkage significantly affect overlong isolated buildings and cannot be neglected.
- (3) After the construction of a building, the building can be protected by thermal insulation and heating measures. Moreover, most of shrinkage has been finished. Therefore, the influence of thermal changes and concrete shrinkage mainly manifests during construction period.

4. Expansion Joint Distance

Temperature variations and concrete shrinkage increase deformations in overlong isolated structures. To avoid exceeding the allowed deviations of the isolated bearings, expansion joints in floor slabs are present. Although considerable literature is available on expansion joint distances in fixed-base buildings, few studies have investigated expansion joint distances in isolated buildings. Based on the analysis, we obtain the distance of expansion joints in isolated buildings. The theoretical model is used to predict the deformations of isolated buildings affected by temperature variations and shrinkage effects. Given that the magnitude of temperature difference depends on local climatic conditions and building finishes, the distances of expansion joints are proposed according to the thermal design regions and heating design conditions of buildings.

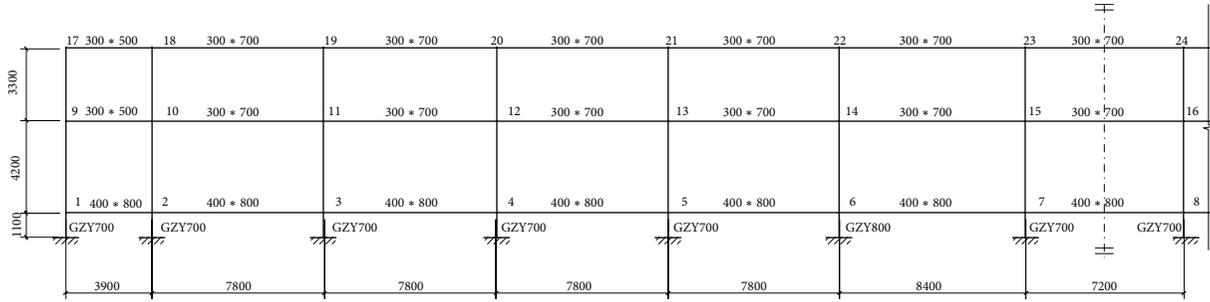


FIGURE 3: Model of the isolated building affected by thermal action and concrete shrinkage.

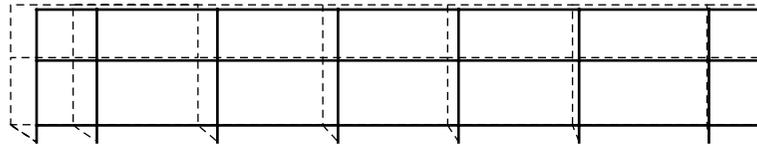


FIGURE 4: Deformation of the isolated building affected by temperature variations.

TABLE 4: Distance of postcast strips for isolated buildings.

Type of thermal design region	Distance of postcast strips (m)
Cold area and severely cold area	20
Other areas	30

The deformation of an isolated bearing because of thermal action and concrete shrinkage can be expressed as follows:

$$\delta = \alpha t l, \tag{11}$$

where l is the distance between an isolated bearing and a fixed point where deformation is zero. The stiffness center of the isolated layer is generally defined as the fixed point. Thus, the isolated bearing with the maximum deformation is the farthest from the building's center. To ensure the safety of the isolated building, no deformations of the isolated bearings must exceed the allowed deviation. Thus, we obtain the following:

$$\alpha t x_c \leq [\Delta], \tag{12}$$

where $[\Delta]$ is the allowed deviation of the bearing $[\Delta] = 5 \text{ mm}$ [14] and x_c is the distance between the end isolated bearing and the stiffness center of the isolated layer.

Equation (12) takes the following form:

$$x_c \leq \frac{[\Delta]}{\alpha t}. \tag{13}$$

Suppose that the stiffness center is close to the center of the building,

$$x_c = \frac{X}{2}. \tag{14}$$

Substituting (14) and $[\Delta]$ into (13) yields the equation for X :

$$X \leq \frac{0.01}{\alpha t}, \tag{15}$$

where X is the maximum length of the floor slab; that is, when the length of the isolated buildings is limited by (13), the effects of thermal action can be neglected. If the isolated building is overlong, X is defined as the distance of expansion joints.

If the material of the building is concrete, substituting α solves X :

$$X \leq \frac{1000}{t}, \tag{16}$$

where t is the temperature difference that is considering thermal changes and concrete shrinkage depending on local climatic conditions, building finishes, building heating, and thermal insulation. During construction period, the building is exposed and the influence of concrete shrinkage is important; thus, t is large. During service period, concrete shrinkage is nearly finished and the building is protected by heating and insulation measures; thus, t is small. Therefore, the expansion joint distance can be proposed according to these two conditions.

4.1. Construction Period. The building is exposed without heating and thermal insulation during construction to ensure the safety of the building. Temperature variation is defined as the maximum seasonal temperature difference, which is equal to the average temperature in the hottest month minus the average temperature in the coldest month. A direct correlation exists between the seasonal temperature differences and the thermal design regions of the building. In China, the different thermal design regions consist of cold areas, severely cold areas, warm areas, hot summer and

TABLE 5: Distance of expansion joints for isolated building.

Building's heating and thermal insulation condition	Temperature difference ($^{\circ}\text{C}$)	Distance of expansion joints (m)
Indoor heating	7~12	80~140
Indoor without heating but wall with exterior insulation	20~25	40~50
Open air	25~35	30~40

cold winter areas, and hot summer and warm winter areas [15]. The resulting seasonal temperature differences are 30, 28, 15, 20, and 15°C , respectively. Furthermore, the influence of concrete shrinkage is important during construction. Suppose that concrete curing is in the standard condition. Thus, the equivalent variation in shrinkage temperature is 20°C . We obtain the temperature differences of 50, 48, 35, and 40°C in the different thermal design regions of the building. By substituting 50, 48, 35, and 40°C into (16), the expansion joint distances in the different thermal design regions of the building are obtained and simplified (Table 4). Considering that expansion joints are used during construction, such joints also represent the postcast strips.

4.2. During Service Period. The buildings are generally protected with heating and thermal insulation during the service period. Thus, temperature variation is directly correlated with the heating and thermal insulation conditions of the building. In China, the heating and thermal insulation conditions of buildings have three types, that is, indoor heating, indoor without heating but walls with exterior thermal insulation, and open air [15]. The temperature variations are shown in Table 5. Given that concrete has mostly shrunk after construction, the equivalent variation in shrinkage temperature can be neglected during the service period. Thus, the temperature differences are equal to the temperature variations. Substituting these values into (16) yields the expansion joint distances (Table 5).

5. Conclusion

In a concrete structure, temperature variations arise from different causes such as solar radiation on the element surface, fire, and hydration heat. In concrete structures, temperature variation and concrete shrinkage generate stresses and strains that must be considered by the designer. The problem of thermal actions and shrinkage effects should be addressed for isolated buildings with large horizontal dimensions. The deformations of some isolated bearings exceed their allowed deviation because of temperature variation and shrinkage, which affect the safety of isolated buildings. In this study, we measure an actual overlong isolated building to assess the deformations caused by temperature variations and shrinkage. To evaluate properly the deformation of isolated buildings because of temperature variations and shrinkage, a three-story frame model is developed and the deformation distribution method is used for numerical analysis. The conclusions of this paper can be summarized as follows.

- (1) The drift of isolated buildings affected by thermal changes and shrinkage is concentrated on the isolated

layer, whereas the deformation of the superstructure is nearly equal. Thus, for the superstructure of isolated buildings, temperature variations and shrinkage effects can be neglected.

- (2) Isolated bearings at the end of the building have maximum deformation. The deformations of isolated bearings are reduced by shortening the distance between the bearings and symmetric axis. For an overlong isolated building, the problem of thermal actions and shrinkage effects should be considered. If the temperature difference is large and the concrete is poorly cured, the deformations of the bearings at the end of the overlong building will exceed the maximum allowed deviation of the building. To ensure the safety of the building, construction measures such as reduction of water cement ratio should be employed to reduce concrete shrinkage deformation.
- (3) For overlong isolated buildings, we use expansion joints to reduce the effect of thermal and concrete shrinkage. After the construction of the building, the building is protected by heating and thermal insulation while most of the concrete shrinkage finishes. The influences of thermal changes and shrinkage mainly manifest during construction. Therefore, for isolated buildings, the distance of the postcast strip is proposed during construction, whereas the distance of the expansion joints is confirmed during the service period.

Acknowledgments

This project was funded by the National Natural Science Foundation of China (Grant no. 61262016), the National Natural Science Foundation of China (Grant no. 51368039), and the Lanzhou University of Technology Doctoral Grant.

References

- [1] J. Ímal, "Charles bridge in Prague—measurement of temperature fields," *International Journal for Restoration of Buildings and Monuments*, vol. 9, no. 2, pp. 585–602, 2003.
- [2] J. M. Fisher, "Expansion joints: where, when, and how," in *Proceedings of the NASCC*, April 2005.
- [3] R. Barsotti and M. Froli, "Statistical analysis of thermal actions on a concrete segmental box-girder bridge," *Structural Engineering International*, vol. 10, no. 2, pp. 111–115, 2000.
- [4] M. Y. Han and R. L. Lytton, "Theoretical prediction of drying shrinkage of concrete," *Journal of Materials in Civil Engineering*, vol. 7, no. 4, pp. 204–207, 1995.
- [5] M. Razaq, Y. Okui, H. Mitamura, and T. Imai, "Seismic performance of a highway bridge with different modeling techniques

- for laminated rubber bearings at a low temperature,” *Cold Regions Engineering*, vol. 120, no. 4, pp. 467–477, 2012.
- [6] B. Wu, L. Han, F. Zhou, C. Shen, and P. Tan, “Experimental study on fire resistance of building seismic rubber bearings,” *Journal of Structural Engineering*, vol. 137, no. 12, pp. 1593–1602, 2011.
- [7] I. V. Kalpakidis and M. C. Constantinou, “Effects of heating on the behavior of lead-rubber bearings. I: theory,” *Journal of Structural Engineering*, vol. 135, no. 12, pp. 1440–1449, 2009.
- [8] I. V. Kalpakidis and M. C. Constantinou, “Effects of heating on the behavior of lead-rubber bearings. II: verification of theory,” *Journal of Structural Engineering*, vol. 135, no. 12, pp. 1450–1461, 2009.
- [9] A. Yakut and J. A. Yura, “Parameters influencing performance of elastomeric bearings at low temperatures,” *Journal of Structural Engineering*, vol. 128, no. 8, pp. 986–994, 2002.
- [10] Y. N. Huang, A. S. Whittaker, and N. Luco, “Seismic performance assessment of base-isolated safety-related nuclear structures,” *Earthquake Engineering and Structural Dynamics*, vol. 39, no. 13, pp. 1421–1442, 2010.
- [11] A. Escandar, M. O. Moroni, M. Sarrazin, and P. N. Roschke, “Mechanical properties and fuzzy modeling of high-damping rubber with thermal effects,” *Journal of Materials in Civil Engineering*, vol. 19, no. 5, pp. 428–436, 2007.
- [12] Code for Design of Concrete Structures GB, 50010-2002, 2002.
- [13] A. Salazar, “On thermal diffusivity,” *European Journal of Physics*, vol. 24, no. 3, pp. 351–358, 2003.
- [14] Rubber bearings—part 3: elastomeric seismic-protection isolators for buildings GB, 20688.3-2006, 2006.
- [15] “Thermal design for civil building,” GB50176-93, 1993.

Research Article

Research on Large-Scale Road Network Partition and Route Search Method Combined with Traveler Preferences

De-Xin Yu,¹ Zhao-Sheng Yang,¹ Yao Yu,² and Xiu-Rong Jiang²

¹ State Key Laboratory of Automobile Simulation and Control, School of Traffic, Jilin University, Changchun 130025, China

² School of Traffic, Jilin University, Changchun 130025, China

Correspondence should be addressed to Yao Yu; yuy-jlu@foxmail.com

Received 5 September 2013; Accepted 24 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 De-Xin Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Combined with improved Pallottino parallel algorithm, this paper proposes a large-scale route search method, which considers travelers' route choice preferences. And urban road network is decomposed into multilayers effectively. Utilizing generalized travel time as road impedance function, the method builds a new multilayer and multitasking road network data storage structure with object-oriented class definition. Then, the proposed path search algorithm is verified by using the real road network of Guangzhou city as an example. By the sensitive experiments, we make a comparative analysis of the proposed path search method with the current advanced optimal path algorithms. The results demonstrate that the proposed method can increase the road network search efficiency by more than 16% under different search proportion requests, node numbers, and computing process numbers, respectively. Therefore, this method is a great breakthrough in the guidance field of urban road network.

1. Introduction

Path optimization is one of the most key issues in traffic flow guidance filed, and it is mainly used for vehicle positioning and system navigation. As we know, travelers need some route information to assist them to get to their destinations. So path guidance technology has been developed which not only can provide travel routes but also relieve the traffic congestion.

The optimal path search theory [1] is the core of the path optimization. There are a lot of relevant researches on the path search algorithms and classic calculation methods around the world. For example, Dijkstra algorithm, A* algorithm, and ant colony algorithm are mature algorithms. But all the above algorithms are generally applied to small-scale road network. Currently, the advanced search path algorithms were improved via parallel computing [2] in order to solve the large network path problem (LNPP). So some optimized algorithms such as Dijkstra algorithm with approximation barrel structure, Pallottino algorithm, and Dijkstra algorithm [3] with quad heap structure have been proposed. These mentioned improved path search algorithms can meet the needs of search speed of large-scale road network path

optimization. However, it rarely considers the travelers' drive habits and preferences, which leads to the path optimization results which are often not accepted by travelers. That is, the travelers' obey rate will drop, and guidance effects cannot be presented.

In view of above-mentioned insufficiency, we consider the drive habits and preferences in order to ensure implementation of the real-time path optimization of large-scale road network. And we construct a traffic network simulation system and a multilayer and multitasking network data storage structure in which parallel data is shared according to the traffic network data [4, 5]. We also adequately analyze the interactional relationship between road network decomposition and shortest path calculation. And then a Hierarchical Pallottino parallel search algorithm model has been proposed in accordance with the data storage and parallel computing technology. The proposed model can dynamically adjust the search level in terms of travel distance. And the multilayer Pallottino search algorithm is run separately in subnets, achieving real-time interaction within boundary nodes via information transmission among the adjacent subnets. Consequently, a global optimal solution can be obtained.

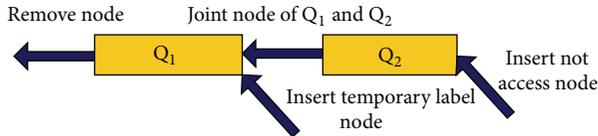


FIGURE 1: Schematic diagram of deque data structure.

2. A Hierarchical Pallottino Parallel Search Algorithm Model Based on the Data Storage

Pallottino algorithm adopts priority deque structure to calculate the optimal path. Deque algorithm [6–8] is a combination of two first-in first-out (FIFO) queues, Q_1 and Q_2 . Q_2 has a higher priority than Q_1 . As for a node, if its state numeral is not accessed, then this indicates it has a higher priority and it will enter the candidate node set from the end of Q_2 ; if its state numeral is temporary, then this indicates it has a lower priority and the candidate node will enter from the end of Q_1 . During the process of route search, the nodes of Q_2 are always firstly searched for achieving the two-queue search mechanism and the fast iterative correction of node labeling [9]. When removing nodes from the candidate node set, the deque needs to be detected whether it is empty or not. If the queue Q_1 is not empty, then it will delete nodes from the front of Q_1 ; if the queue Q_1 is empty, it will continue to detect whether the queue Q_2 is empty or not. If the queue Q_2 is not empty, then it will delete nodes from the front of Q_2 . Operate the above cycle after deleting a node. The operation does not end until the queue Q_2 is empty. The schematic diagram of deque data structure is shown in Figure 1.

In daily life, a sizable number of drivers prefer to select the main sections to travel. However, when driving for a near distance, they may not select the main road [10]. Therefore, a multilayer road network data storage means is utilized to improve the Pallottino search algorithm, and a Hierarchical Pallottino parallel search algorithm is put forward. When the driver travels for a far distance, the proposed algorithm should be used to search the main road node layer. When the driver travels for a near distance (when the number of the cells between the origin and destination is less than or equal to 1, it is regarded as “the travel distance is near”), it needs to search both the main road node layer and secondary road node layer.

2.1. Urban Road Network Partition and Optimization

2.1.1. Urban Road Parallel Data Structure. The data structure of urban road network [11] is a key point of solving path problem, which directly determines the difficulty and timeliness of path algorithm procedures. When organizing network data, it not only invokes conveniently and fast, but also reduces the occupation of memory resources and expresses clearly in order to facilitate data calls and program checks. In terms of urban road traffic network, its corresponding data structure has the following three characteristics: (a) the data is so large, and the program needs to call and access it frequently; (b) the topological relations of road network are clear and the

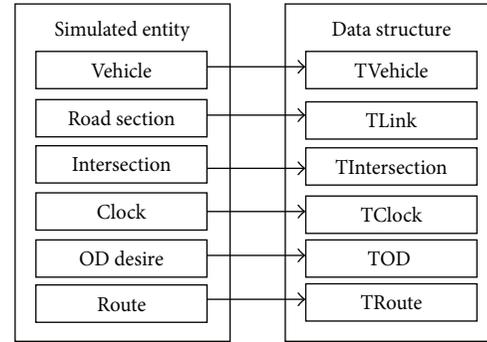


FIGURE 2: Traffic network simulation systems and shared parallel data structure.

relationship between the nodes and the sections is fixed; (c) there is concurrency when the data structure is accessed by the program.

The basic element of traffic network data system is composed of car, section, route, intersection, clock, and so forth. Furthermore, every traffic network element and road network structure all can be considered as a single object. This paper will adopt C++ to set up traffic network simulation system and shared parallel data structure through object-oriented class definition (shown in Figure 2), to meet the demand of the following Hierarchical Pallottino parallel search algorithm.

The above different types of member variables and member properties are shown in Algorithm 1.

After defining various types of entities in network structure, for every type of entity, a single object is just one unit of traffic network objects set. Therefore, multitasking concurrent accessing data could be implemented through establishing shared data structures so as to improve the speed of accessing data, satisfying the timeliness requirement of large-scale road network guidance.

2.1.2. Network Structure Optimization Considering Drivers' Characteristics. As we know, drivers often choose the road they are familiar with or the urban trunk roads as their priority travel routes. In view of the characteristics, we should give priority to search and calculate this layer of the road network to ensure the speed of calculation, which is particularly important for the large-scale road network.

In this paper, urban road network is divided into multiple levels [12, 13]. The roads preferred by drivers are taken as main roads to constitute the main road network which divides the urban road network into small zones. In addition, a series of minor road networks are constituted by the road network of every zone (like zone K, M, and L in Figure 3). With regard to doing route optimization, the final route optimization network is made up of the minor road network within the scope of drivers' origin zone or destination zone and the main road network. Thus searching road network will be simplified and the speed of searching improved.

As for the road impedance function, there are five popular items including traveling time, driving distance, congestion degree, road quality, and comprehensive cost [14–16]. In view

```

Deque<TVehicle>vehiclist // Define the vehicle queues;
Deque<TLink>linklist // Define the road section queues;
Deque<TIntersection>intersectionlist // Define the intersection queues;
Deque<TOD>ODlist // Define the OD queues;
Deque<TRoute>Routelist // Define the Route queues;
    
```

ALGORITHM 1: Definitions of road network structure and queue structure.

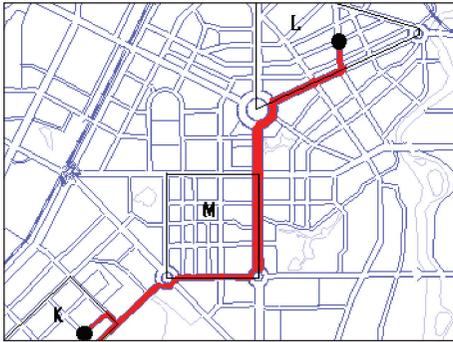


FIGURE 3: Multilevels of urban road network.

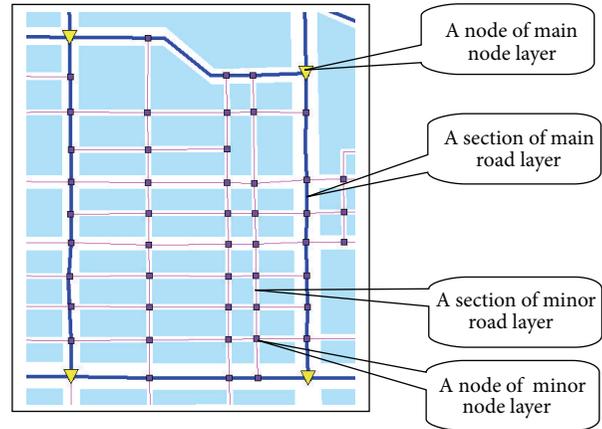


FIGURE 4: Layered search thought diagram.

of above ideas, real-time dynamic generalized travel time could be defined as road impedance (1) in this paper. There are several ways to determine the road impedance data, which reflects different user requirements and system control strategies. Considering the circumstances, the final route is not the absolute optimal one in accordance with generalized road impedance in the whole road network, but the result of quasi-drivers changing directions based on optimization principle.

Consider the following:

$$C_i(k) = t_i(k). \tag{1}$$

The formula above is the most common dynamic road impedance form, which indicates the average time that vehicles take on the k th road section, containing travel time and parking delay time (including intersection delay).

In order to avoid common problems caused by the road network partition, this paper adopts the thought of layered optimization. Firstly, the main roads and the minor roads are separated into two layers. The intersections of the main roads are defined as main nodes and the intersections of the minor roads are defined as minor nodes (Figure 4). The triangle symbol indicates the node of the main nodes layer and its number is equal to the number of intersections of the main roads in the whole map. According to the network layered perspective, removing the intersections of the main roads and the minor roads will greatly reduce the time complexity in the process of route calculation, thereby the calculation speed will be improved.

In order to solve the problem that drivers may not choose the main roads when the driving distance is relatively short, this paper adopts the calculating method of neighboring zones union. If the origin and the destination are in the same

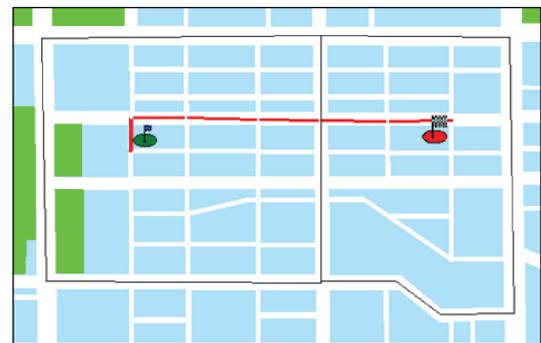


FIGURE 5: Neighboring zones incorporated principle diagram.

or the neighboring zones, this paper will take the nodes of the same or the neighboring zones as a union and search the union during the process of calculation. In principle, if there are excessive nodes in zone, it will affect the speed of route calculation. Thus in the process of zone division, the scope of zones should be reduced as much as possible in order to reduce the number of nodes. Figure 5 shows the route calculation results incorporated into the principle.

2.2. The Hierarchical Pallottino Parallel Search Algorithm Model. According to the above mentioned road network data storage and network decomposition method, a large-scale network will be decomposed into several subnetworks [17] whose number is the same as that of the processor. We use the Hierarchical Pallottino search algorithm in every processor. At this moment, it cannot obtain the information of adjacent

nodes of boundary nodes, so the global optimal solution cannot be obtained. Therefore, while using the Hierarchical Pallottino search algorithm for parallel computing, the paper uses the way of information transfer to send the boundary node's real-time information to the processor. And the processor receives the information sent from the adjacent subnet so as to complete the boundary node label iterative correction and obtain global optimal solution.

The specific calculation steps are as follows.

Step 1. According to the starting point O and the terminal point D position that user inputs, processor could analyze the road node layer and generate a road barrier matrix W . Define the node collection of cell i as $P(i)$, the node collection of the main road as PZ , and the road network node collection as P . Identify the starting point O and the terminal point D that user inputs and determine their cells k , and l , respectively. If the starting point and the terminal point are located in the same cell, then set $P = P(k)$; if they are not in the same cell, then set $P = PZ$ and incorporate the nodes of the cells k and l into P . If there is only one cell m between the starting point and the terminal point, then the nodes of cell m will be incorporated into P .

Step 2. Label each node i in the search network node collection. There are three numerals which are length numeral d_i , predecessor node numeral p_i , and current node status label s_i . Length numeral d_i indicates the accumulated weights from the starting point to the point i . Predecessor node numeral p_i is the previous direct point when the shortest distance between the starting point q and the point i obtains the maximum value. Current node status label s_i contains unvisited temporary labels and permanent labels.

Step 3. Define a message queue about the boundary node information for each subnetwork: $M_x[h]$ ($x = 1, 2, \dots, k$, $h < n/k$).

Step 4. Initialize the queues Q_1 and Q_2 when calculating the shortest path from arbitrary starting point q to all other nodes i , and we can set

$$d_q = 0 \quad d_i = +\infty \quad i \neq q. \quad (2)$$

Step 5. Judging the condition of queue Q , if it is empty, then go to Step 7, and if not, go to Step 6.

Step 6. Judge whether the formula (3) is established or not for each adjacent node j of node i in the nonempty queue Q

$$d_j > d_i + \omega_{i,j}, \quad (3)$$

where $\omega_{i,j}$ is the section weight between nodes i and j . If the formula (3) is established, the following formula will be set:

$$d_j = d_i + \omega_{i,j}. \quad (4)$$

If node j is not in the Q and s_j is not accessed, j will be inserted into the end of the Q_1 and s_j will be updated to a temporary label. If node j is not in the Q and s_j is temporary mark state, j will be inserted into the end of the Q_2 .

Step 7. In current subnet, observing each adjacent node r of node l and judging whether it is the shortest path, we can write

$$d_r > d_l + \omega_{l,r}. \quad (5)$$

If (5) is established, the node r will be added to its subnet message queue, and message queue will be sent to all adjacent subnetworks.

Step 8. Judge the condition of message queue. If it is empty, go to Step 10, and otherwise, go to Step 9.

Step 9. Each subnet sends a message queue to adjacent subnets and accepts feedback from other subnets, and each received adjacent node a which is adjacent to node t in the current process is calculated as follows:

$$d_a > d_t + \omega_{t,a}. \quad (6)$$

If (6) is established, then the following formula will be set:

$$d_a = d_t + \omega_{t,a}. \quad (7)$$

If node a is not in the Q and s_a is not accessed, then a will be inserted into the end of the Q_1 and s_a will be updated to a temporary label. If node a is not in the Q and s_a is temporary mark state, a will be inserted into the end of the Q_2 .

Step 10. Each processor sends the results of shortest path calculation to the host processor, and the algorithm ends.

3. Experimental Analysis

In order to verify the optimization results of the proposed Hierarchical Pallottino parallel search algorithm based on data storage, the paper uses C++ and OpenMP parallel programming method to design and develop parallel computing program of large-scale road network, utilizing client/server/(C/S) mode for setting up parallel computing platform. We test the proposed method based on real road network data of Guangzhou. This traffic network includes 11,489 nodes and 18,364 paths.

Experimental platform: Cluster of Workstations; CPU: Intel(R) Core(TM)2 Duo CPU E8600@3.33 GHZ Dual-Core Processor; Memory: 2 GB, each node Gigabit Switch: 4 sets.

3.1. Road Network Decomposition. Combined with the characteristics of drivers route choice, Guangzhou road network (central urban area as shown in Figure 6(a)) is divided into the following areas, and the dynamic topological relations of each subinterval (the corresponding dual connectivity graph) are shown in Figure 6(b).

According to the corresponding dual connectivity characteristics, it can be illustrated that the selected network, to some extent, possesses the small-world property, and road network has both smaller characteristic path length and larger clustering coefficient. Furthermore, fewer obstacles of road network are reflected from the small-world property, which has a good accessibility. According to the calculation

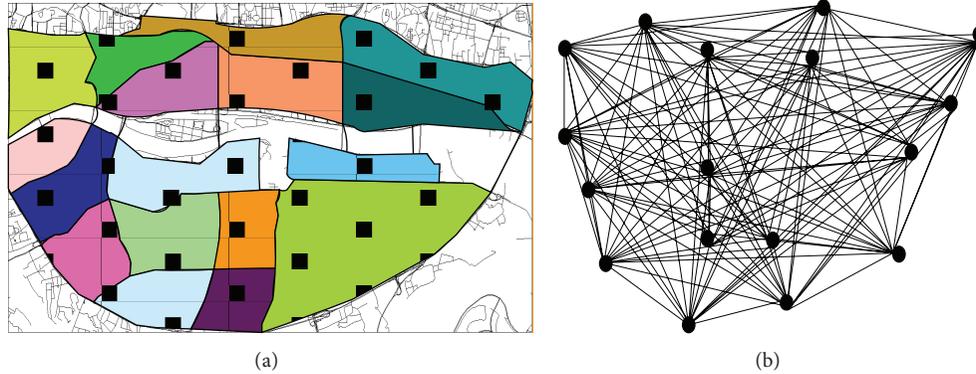


FIGURE 6: (a) Subarea of road network; (b) the dynamic topological relations.

results of network characteristic value and characteristic judgment conditions of small-world network, combined with the characteristics of the network and the connectivity of the intersection, we are convinced that it is the basis for evaluation of road network structure and the optimization strategy. Due to limited space, this paper is not to research this part. Decomposition of Guangzhou actual road network is shown in Figure 7.

3.2. Shortest Path Parallel Computing Experiment. On the traffic network, the paper separately applies the proposed Hierarchical Pallottino parallel search algorithm (HTWO_Q), Dijkstra algorithm based on approximate barrel structure (DIKBA), Pallottino algorithm (TWO_Q), and Dijkstra algorithm based on improved quad-stack structure (DIKQH) to calculate shortest path search time within all pairs based on the weights of generalized travel time.

In order to ensure the executive effectiveness of algorithms, the paper adopts both undecomposed and decomposed networks to analyze algorithms. In the undecomposed network conditions, we separately compare them in different conditions (i.e., 1:1, 1:N, and N:N) and search results are shown in Table 1.

Table 1 shows that the search speeds of the proposed HTWO_Q algorithm and TWO_Q algorithm are basically equivalent in the standard network. In 1:1 search, the computation times of DIKQH and DIKBA algorithms are superior to the proposed algorithm, but their calculation times are inferior to the HTWO_Q algorithm and TWO_Q algorithm in N:N search. Due to the advantages of the deque data extraction and storage, the path search times of HTWO_Q and TWO_Q algorithms are superior to the other two search algorithms by more than 5% in 1:N search.

Under the decomposed network conditions, we use the parallelization algorithms to calculate the search time by computing cluster under the processes 8 and 10. The calculation results are shown in Table 2.

From Table 2 we can see that the search time of the proposed HTWO_Q algorithm is obviously superior to those of the other three algorithms in 1:N and N:N search, but it is inferior to DIKQH and DIKBA algorithms in 1:1 search. In the actual urban guidance system, 1:1 path search has

limitations, while 1:N and N:N search could satisfy the dynamic requirements of traveler and have more extensive application value.

During the path search, it is obvious that different numbers of network nodes will affect executive time of algorithm. In order to better illustrate the calculation time of these algorithms under different vertices, the algorithms' executive efficiency is compared and analyzed, respectively, in 4000, 8000, and the whole road network nodes. Results are shown in Figure 8. When road network data is stored in the way we have mentioned, the time of multiprocessor parallel computing optimal path (or global optimal solution) is closely related to the number of processes. Then, we compare the algorithms, respectively, using 6, 8, and 10 processes, and the results are shown in Figure 9.

Figure 8 demonstrates that the search speeds of HTWO_Q and TWO_Q are similar within 4000 nodes' network, which are lower than those of DIKBA and DIKQH. However, when searching in 8000 nodes and in the whole network, the proposed HTWO_Q algorithm is superior to the other three algorithms, and its operating efficiency is 16.32% faster than that of the other three ones.

As shown in Figure 9, with the increasing of processes, the path-searching efficiency is not upward trend. By contrast, it firstly increases and then decreases, which is concerned with consumption of communication time when parallel algorithms are running. The result indicates that when process number is 8, the search efficiency of road network can be greatly enhanced.

4. Conclusions

Considering travelers' preferences, the purpose of this paper is to solve large-scale road network path optimization problem. With the way of object-oriented classified definition, a simulated traffic network system and parallel data structures have been constructed. To analyze the relationship between data storage and network decomposition, a multilayer Pallottino parallel search algorithm is proposed. Then, series of sensitive experiments are conducted to verify the algorithm under actual road network environment. Finally, the results show that the search time of the proposed optimization

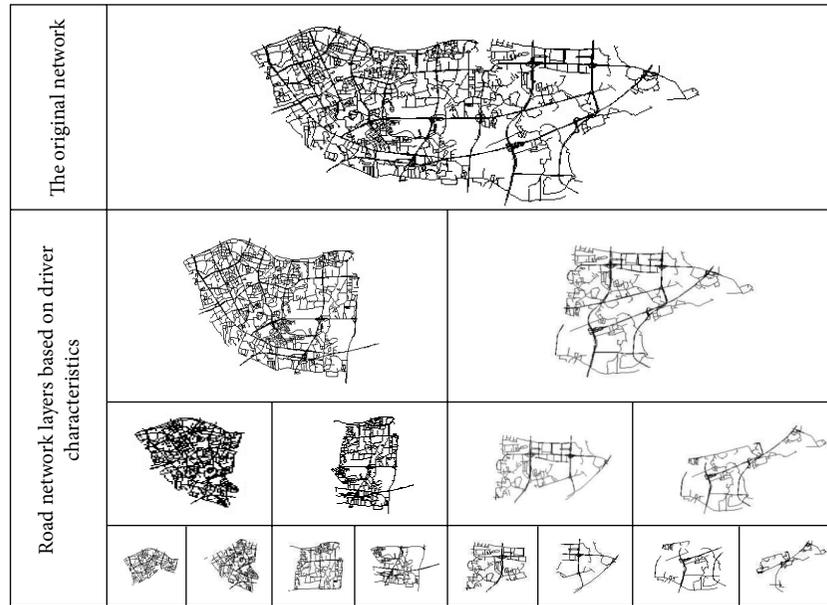


FIGURE 7: Decomposition of actual road network of Guangzhou.

TABLE 1: Comparison of path search times under initial network environment.

Computing time (s)	1:1 route searching	1:N route searching	N:N route searching
HTWO_Q	0.0097331000	0.0097331000	94.2922860000
DIKBA	0.0051053340	0.0103312830	100.4570000000
TWO_Q	0.0098270450	0.0098270450	95.3031400000
DIKQH	0.0050473020	0.0110842745	103.4130000000

TABLE 2: Comparison of path search times under Hierarchical network environment.

Computing time (10^{-3} s)	1:1		1:N		N:N	
	8	10	8	10	8	10
HTWO_Q	0.94271	1.00325	0.94263	1.00399	80.3311	83.6300
DIKBA	0.71433	0.91125	1.30764	1.38241	153.296	168.117
TWO_Q	1.15031	1.20301	1.14998	1.21334	100.775	119.068
DIKQH	0.807979	0.91966	1.58446	1.78677	174.210	195.429

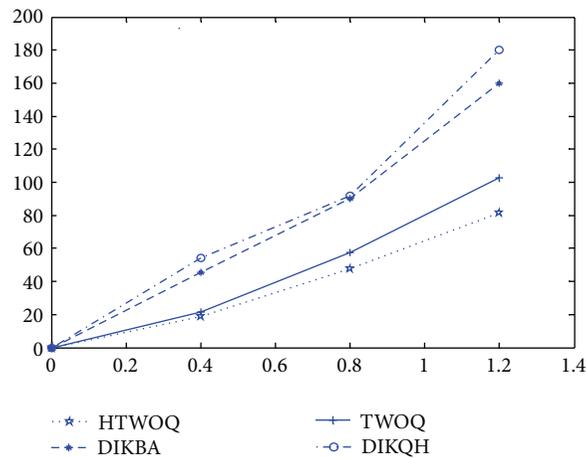


FIGURE 8: Computation times comparison of the four algorithms under N:N search pattern. The x-axis represents numbers of vertices (unit: thousand) and the y-axis represents computed times (unit: ms).

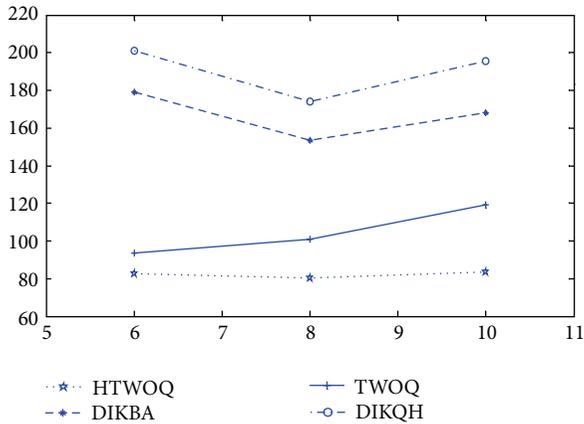


FIGURE 9: Computation times comparison of the four algorithms under multiple processes. The x -axis represents numbers of processes and the y -axis represents computed times (unit: ms).

method is significantly better than those of DIKBA, TWO_Q, and DIKQH during the time of running 1:N and N:N patterns. Calculated results can not only meet the real-time demand but also satisfy the preferences of travelers, which can effectively improve the obey rate of guidance service and have positive impacts on development of intelligent guidance systems.

Conflict of Interests

The authors claim that there is no conflict of interests in their research.

Acknowledgments

The authors express their sincere appreciation to the Chinese National High Technology Research and Development Program Committee for the financial supports for this research under Grants no. 2012AA112307 and National Natural Science Foundation no. 61074137.

References

- [1] M. Mesbah, M. Sarvi, I. Ouveysi, and G. Currie, "Optimization of transit priority in the transportation network using a decomposition methodology," *Transportation Research C*, vol. 19, no. 2, pp. 363–373, 2011.
- [2] S. Cicerone, G. D'Angelo, G. di Stefano, D. Frigioni, and V. Maurizio, "Engineering a new algorithm for distributed shortest paths on dynamic networks," *Algorithmica*, vol. 66, no. 1, pp. 51–86, 2013.
- [3] M. R. Khouadjia, E. Talbi, L. Jourdan, B. Sarasola, and E. Alba, "Multi-environmental cooperative parallel metaheuristics for solving dynamic optimization problems," *The Journal of Supercomputing*, vol. 63, no. 3, pp. 836–853, 2013.
- [4] C. Artigues, M.-J. Huguet, F. Gueye, F. Schettini, and L. Dezou, "State-based accelerations and bidirectional search for bi-objective multi-modal shortest paths," *Transportation Research C*, vol. 27, pp. 233–259, 2013.
- [5] L. Fu, "An adaptive routing algorithm for in-vehicle route guidance systems with real-time information," *Transportation Research B*, vol. 35, no. 8, pp. 749–765, 2001.
- [6] Y. Nagasaka, K. Okada, A. Nakamura et al., "Method combining focus-shift and angle multiplexing method for holographic data storage," *Japanese Journal of Applied Physics*, vol. 47, no. 7, pp. 5904–5908, 2008.
- [7] M. A. A. Martínez, J. Cordeau, M. Dell'Amico, and M. Iori, "A branch-and-cut algorithm for the double traveling salesman problem with multiple stacks," *INFORMS Journal on Computing*, vol. 25, no. 1, pp. 41–55, 2013.
- [8] H. Sundell and P. Tsigas, "Lock-free deques and doubly linked lists," *Journal of Parallel and Distributed Computing*, vol. 68, no. 7, pp. 1008–1020, 2008.
- [9] A. Arvind and C. P. Rangan, "Symmetric Min-Max heap: a simpler data structure for double-ended priority queue," *Information Processing Letters*, vol. 69, no. 4, pp. 197–199, 1999.
- [10] D. Luxen and P. Sanders, "Hierarchy decomposition for faster user equilibria on road networks," in *Experimental Algorithms*, vol. 6630 of *Lecture Notes in Computer Science*, pp. 242–253, Springer, Berlin, Germany, 2011.
- [11] J. Gilles and Y. Meyer, "Properties of BV-G structures + textures decomposition models. Application to road detection in satellite images," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2793–2800, 2010.
- [12] D. Rosas, J. Castro, and L. Montero, "Using ACCPM in a simplicial decomposition algorithm for the traffic assignment problem," *Computational Optimization and Applications*, vol. 44, no. 2, pp. 289–313, 2009.
- [13] M. Holzer, F. Schulz, and D. Wagner, "Engineering multilevel overlay graphs for shortest-path queries," *Journal of Experimental Algorithmics*, vol. 13, article 5, pp. 1–26, 2008.
- [14] M. Yoshikawa and H. Terai, "Hardware-oriented ant colony optimization and its application to dynamic route guidance problem," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 6, no. 3, pp. 221–228, 2011.
- [15] C. Ren, "Study on improved tabu search algorithm for Min-Max vehicle routing problem," *Applied Mechanics and Materials*, vol. 87, pp. 178–181, 2011.
- [16] L. Roditty, "On the k shortest simple paths problem in weighted directed graphs," *SIAM Journal on Computing*, vol. 39, no. 6, pp. 2363–2376, 2010.
- [17] H. Meyerhenke, B. Monien, and T. Sauerwald, "A new diffusion-based multilevel algorithm for computing graph partitions," *Journal of Parallel and Distributed Computing*, vol. 69, no. 9, pp. 750–761, 2009.

Research Article

LQR-Based Power Train Control Method Design for Fuel Cell Hybrid Vehicle

Yun Haitao, Zhao Yulan, Liu Zunian, and Hao Kui

School of Automobile and Traffic, Qingdao Technological University, Qingdao 266033, China

Correspondence should be addressed to Yun Haitao; yunht@163.com

Received 9 September 2013; Revised 13 October 2013; Accepted 24 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Yun Haitao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the mathematical model of fuel cell hybrid vehicle (FCHV) proposed in our previous study, a multistate feedback control strategy of the hybrid power train is designed based on the linear quadratic regulator (LQR) algorithm. A Kalman Filter (KF) observer is introduced to estimate state of charge (SOC) of the battery firstly, and then a linear quadratic regulator is constructed to compute the state feedback gain matrix of the closed-loop control system. At last, simulation and actual test are utilized to demonstrate this new approach.

1. Introduction

Fuel cell vehicle (FCV) is one of the most important trends of new energy vehicles. Since the dynamic response of fuel cell has a property of time lag, which fails to meet the demands of the fast-changing vehicle load, a hybrid power system equipped with auxiliary energy storage device is indispensable. This is so-called fuel cell hybrid vehicle. For such a complex system which has multiple power sources, it is very important for energy management strategy to maintain the power balance of the power train.

There have been many academic papers discussing the energy management of the FCHV in recent years. Hyun et al. [1] discussed the feature of hydrogen FCHV power train based on experimental study. Pede et al. [2] discussed hybrid degree of fuel cell vehicle power train. Mohammadian et al. [3] utilized neural network and genetic algorithm to solve the control problem of the fuel cell power system. Kim and Peng [4] designed a power allocation algorithm based on Markov decision programming (MDP). Eren et al. [5] proposed a control strategy for fuel cell hybrid power system based on fuzzy logic. Ryu et al. [6] designed different energy distribution strategies according to the driving modes of the motor. Torreglosa et al. [7] put forward a power control method based on optimal theory. Recently, there are several

patents [8–10] with the focus on configuration and control of FCHV power train. In addition, several research teams pay attention to model and simulation of FCHV [11–14].

2. Simulation Model of the FCHV Power Train

The fuel cell hybrid vehicle (FCHV) studied in this research is *SHANGHAI* prototype vehicle from Shanghai Motor Company, whose power train configuration is shown in Figure 1. The mathematic model of *SHANGHAI* power train has been constructed based on test data and the equivalent circuit model in our previous research [15].

The following equation is the state-space equation of the mathematic model:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u},\end{aligned}\tag{1}$$

where the state vector $\mathbf{x} = [I_{dcf} \ I_m \ U_E \ U_1 \ U_2]^T$, the input vector $\mathbf{u} = [I_{dcfset} \ I_{mset}]^T$, the output vector $\mathbf{y} = U_{bus}$, and \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} are the state matrix, as defined in (2), respectively:

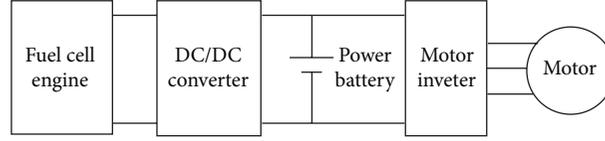
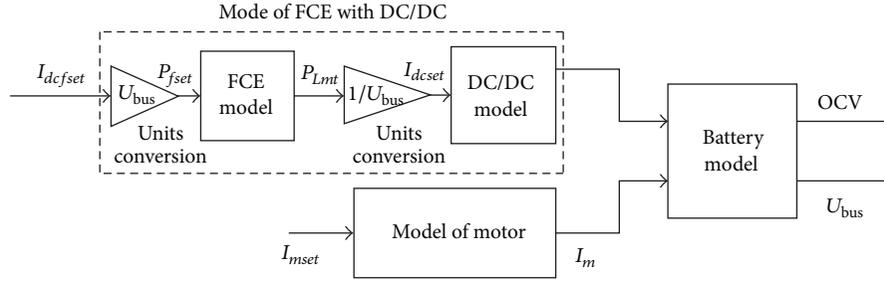


FIGURE 1: Power train configuration of SHANGHAI FCHV.



I_{dcfset} : current setting (FCE with DC/DC); I_{mset} : motor current setting

P_{fset} : FCE power setting; P_{Lmt} : upper limit of FCE power

I_m : motor output current; I_{dcf} : DC/DC output current

OCV: battery open circuit voltage; U_{bus} : bus voltage

I_{dcset} : DC/DC current setting

FIGURE 2: The block diagram of the simulation model.

$$A = \begin{bmatrix} -\frac{1}{T_f} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{T_{mc}} & 0 & 0 & 0 \\ \frac{1}{C_E} & -\frac{1}{C_E} & 0 & 0 & 0 \\ \frac{1}{C_1} & -\frac{1}{C_1} & 0 & -\frac{1}{R_1 C_1} & 0 \\ \frac{1}{C_2} & \frac{1}{C_2} & 0 & 0 & -\frac{1}{R_2 C_2} \end{bmatrix}, \quad (2)$$

$$B = \begin{bmatrix} \frac{K_f}{T_f} & 0 \\ 0 & \frac{K_{mc}}{T_{mc}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$C = [R_0 \quad -R_0 \quad 1 \quad 1 \quad 1], \quad D = [0 \quad 0].$$

Herein, K_f , T_f , K_{mc} , T_{mc} , C_E , C_1 , C_2 , R_0 , R_1 , and R_2 are constants of the model, which can be estimated by the least-squares parameter estimation method.

Based on the mathematic model, a simulation model is constructed. Figure 2 is the overall system model block diagram, which indicates the input and output signals of all component models and how they are interconnected.

To verify the mathematic model and simulation model of the power train, the exemplified comparison between simulation results with test data is presented in Figure 3, confirming that the FCHV model can meet the demand of model-based FCHV control design.

3. Multistate Feedback Control Strategy Designing Based on LQR

According to state equation of the FCHV power train, the power train is marginally stable and little random interference may cause system instability. Therefore, it is very important for energy management system to maintain the power balance of the power train.

In this study, a multistate feedback control strategy is designed based on linear quadratic regulator (LQR) algorithm. A multistate feedback controller is designed using LQR technique and a Kalman filter (KF) observer is designed for battery open circuit voltage (OCV) estimation. Figure 4 shows the schematic diagram of this new control strategy, where u_f is feedforward control input, u_b is feedback control input, u is actual input, y is actual output, x_{target} is target setting of the controller, and \hat{x} is the state observed by the KF observer.

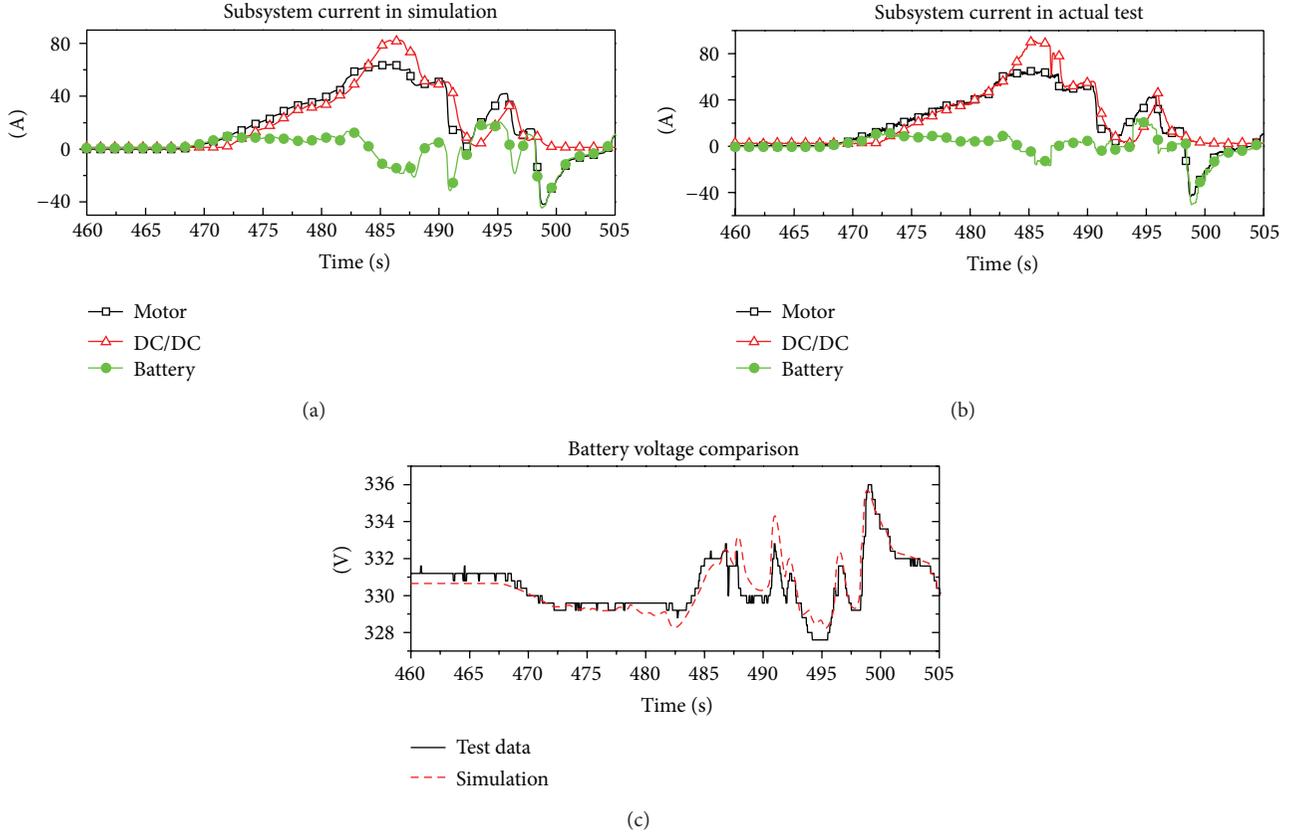


FIGURE 3: Comparison between simulation results and test data (part of J1015 cycle).

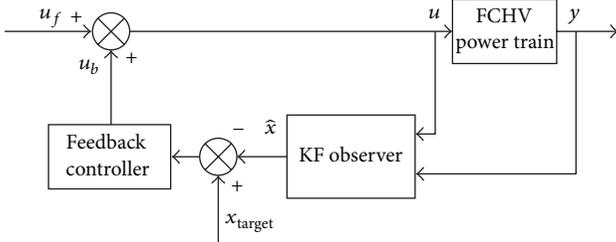


FIGURE 4: The schematic diagram of the multistate feedback control strategy.

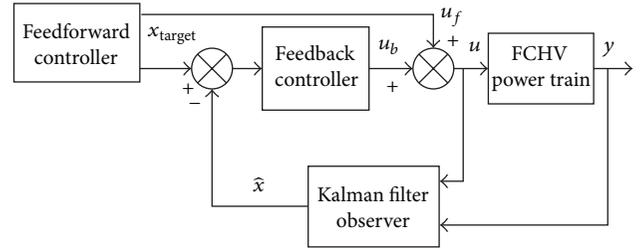


FIGURE 5: The schematic diagram of the multistate feedback strategy (reformed).

To make the control framework more suitable for the actual control input, the control framework shown in Figure 4 is reformed, as shown in Figure 5. This control strategy can be described by

$$u = u_f + u_b$$

$$u_f = [0 \ 0 \ I_{mset_{ff}}]^T$$

$$u_b = \begin{bmatrix} P_{fset_{fb}} \\ I_{dcfset_{fb}} \\ I_{mset_{fb}} \end{bmatrix} = K (x_{target} - \hat{x})$$

$$= \begin{bmatrix} K'_{11} & K'_{12} & K'_{13} & K'_{14} & K'_{15} \\ K_{21} & K_{22} & K_{23} & K_{24} & K_{25} \\ K_{31} & K_{32} & K_{33} & K_{34} & K_{35} \end{bmatrix}$$

$$\times \left(\begin{bmatrix} 0 \\ 0 \\ OCV_{obj} \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \hat{I}_{dcf} \\ \hat{I}_m \\ \hat{U}_E \\ \hat{U}_1 \\ \hat{U}_2 \end{bmatrix} \right)$$

$$\times [K'_{11} \ K'_{12} \ K'_{13} \ K'_{14} \ K'_{15}]$$

$$= U_{bus} [K_{11} \ K_{12} \ K_{13} \ K_{14} \ K_{15}] \cdot$$

(3)

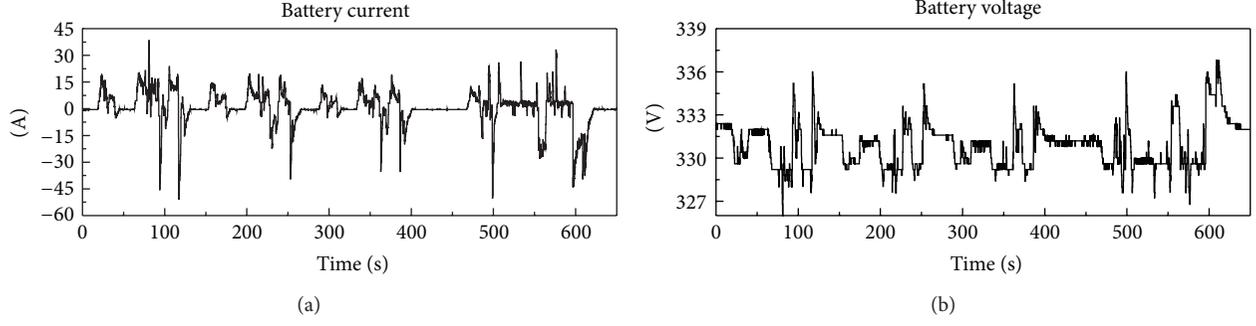


FIGURE 6: Current and voltage of the battery (J1015 cycle).

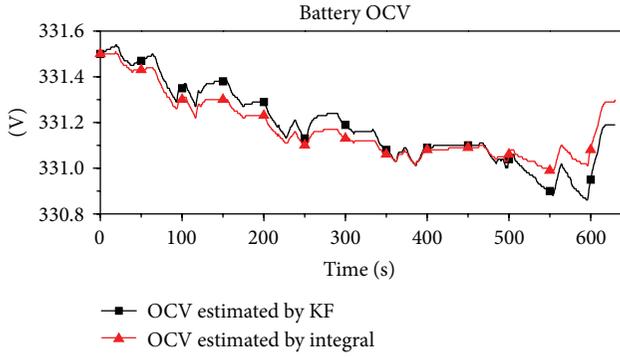


FIGURE 7: Battery OCV estimated by the KF observer and current integral method.

Herein, OCV_{target} represents the target value of the battery OCV, and K is the state feedback matrix. The variables shown in (3) have been explained in Figure 2. In the actual control system, u_f is computed by power demand of the power train and limited by some boundary condition u_b is computed by control deviation and the state feedback matrix K .

3.1. Observer Designing Based on KF. Although battery OCV is not a linear function of battery SOC, according to [16], when battery SOC is in 20% ~ 90% (its work range), battery OCV has a corresponding relationship with battery SOC. Therefore, the fluctuated ranges of the SOC can be regulated by changing the target value of the battery OCV OCV_{target} , namely, U_E of the state-space model. However, the state U_E is not a measurable variable in the power train. To estimate the state U_E , a KF observer is designed. The discrete state-space model of the battery can be obtained from the state-space model of the battery equation as follows:

$$\begin{aligned} x(k) &= Gx(k-1) + Hu(k) + \Gamma w(k), \\ z(k) &= Mx(k) + Nu(k) + v(k), \end{aligned} \quad (4)$$

where state vector $\mathbf{x}(k) = [U_E(k) \ U_1(k) \ U_2(k)]^T$, input vector $\mathbf{u}(k) = [I_b(k)]$, output vector $\mathbf{y}(k) = [U_{\text{bus}}(k)]$, Γ is the noise transfer matrix, $w(k)$ and $v(k)$ are zero-mean white

Gaussian stochastic process, and G , H , M , and N are discrete state matrices defined as

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{R_1 C_1}{R_1 C_1 + T} & 0 \\ 0 & 0 & \frac{R_2 C_2}{R_2 C_2 + T} \end{bmatrix}, \quad H = \begin{bmatrix} -\frac{T}{C_E} \\ -\frac{TR_1 C_1}{R_1 C_1 + T} \\ -\frac{TR_2 C_2}{R_2 C_2 + T} \end{bmatrix},$$

$$M = [1 \ 1 \ 1], \quad N = [-R_0]. \quad (5)$$

Herein, T is the sampling period. The filter is initialized with the best information available on the state and error covariance, and the recursive equation of the KF is summarized as follows:

Initialization:

$$\begin{aligned} \widehat{X}(0) &= E[X(0)], \quad P(0) = E[X(0)X(0)^T] \\ Q &= E[w \times w^T], \quad R = E[v \times v^T] \end{aligned}$$

Recursive Computation:

$$\begin{aligned} P(k|k-1) &= GP(k-1)G^T + \Gamma Q \Gamma^T, \\ K_k &= P(k|k-1)M^T [MP(k|k-1)M^T + R]^{-1}, \\ \widehat{X}(k) &= [I - K_k H] [G\widehat{X}(k-1) + Hu(k)] \\ &\quad + K_k z(k) - Nu(k), \\ P(k) &= [I - K_k M] P(k|k-1). \end{aligned} \quad (6)$$

Although the hypothesis about the noise process cannot conform exactly to the fact, the result obtained from the test shows that the KF observer still performs well.

To validate the KF observer designed here, a J1015 dynamic test was performed (sample period was 50 ms) As the input data of the KF observer, the current and voltage of the battery were collected in the test, as shown in Figure 6. The estimated results of the battery OCV by the KF observer and current integral calculation method shown in Figure 7

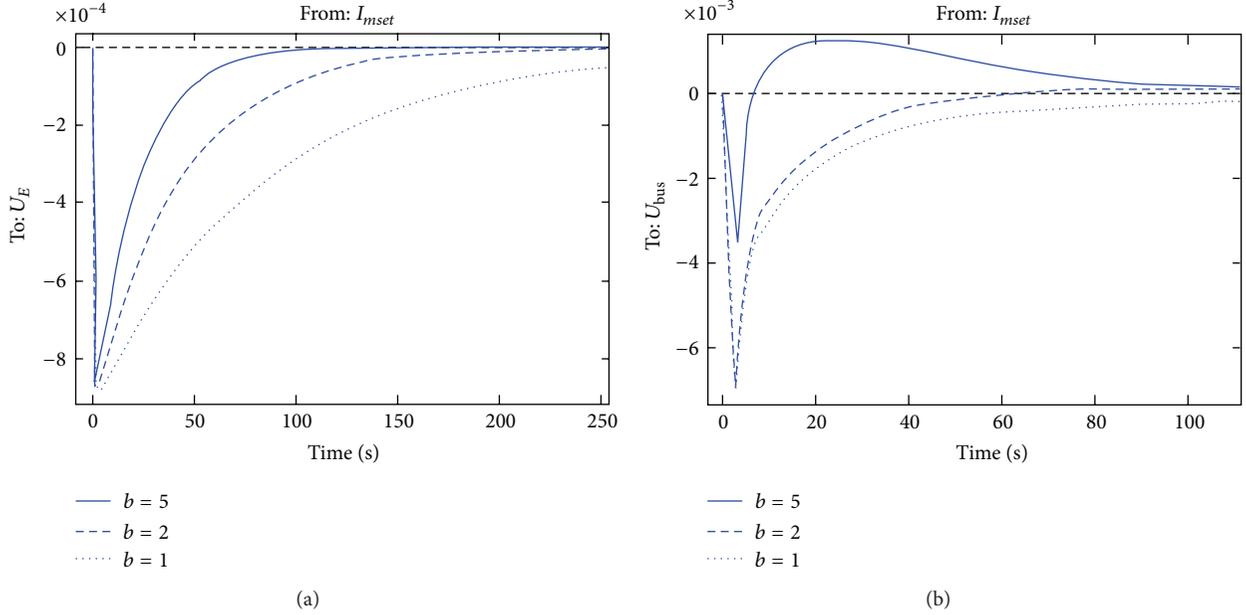


FIGURE 8: Step responding of the model.

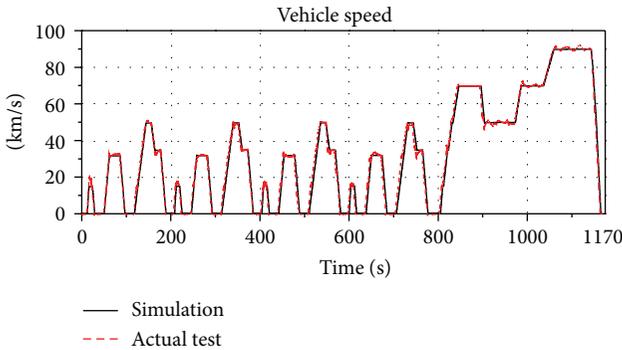


FIGURE 9: Changing process of vehicle speed (ECE_EUDC cycle).

indicate that, if the initial deviation is excluded from consideration, the maximum error is not more than 1% in the cycle test. Though the current test data has its inevitable error, in a relatively short time OCV estimated by current integral calculation approaches the actual value. Therefore, the test result indicates that the KF designed for battery OCV estimation is valid.

3.2. Multistate Feedback Controller. The state feedback matrix K is very important for the multistate feedback controller. In actual control strategy, elements of the matrix K need to be restrained by some limitation and boundary condition, as shown in the following:

$$\begin{aligned} K_{11} < m, \quad K_{13} < n, \quad K_{14} < o, \quad K_{15} < p, \\ K_{21} = K_{22} = K_{23} = K_{24} = K_{25} = 0, \end{aligned} \quad (7)$$

where m , n , o , and p are the limitations determined by configuration parameters of the power train and practical

experience. K_{11} is selected according to the test, K_{13} , K_{14} , and K_{15} can be selected properly by analyzing the step responding of the model. To compute state feedback gain in an optimal way, LQR is constructed based on the linear mathematic model of the power train, and the quadratic characteristic index function of the regulator is as follows:

$$\begin{aligned} J &= \frac{1}{2} \int_{t_0}^{\infty} \left[a(I_{dcf} - I_m)^2 + bU_E^2 + cU_1^2 + dU_2^2 \right. \\ &\quad \left. + eI_{dcfset}^2 + fI_{mset}^2 \right] dt \\ &= \frac{1}{2} \int_{t_0}^{\infty} \left[x^T(t) Q x(t) + u^T(t) R u(t) \right] dt. \end{aligned} \quad (8)$$

In (8), a , b , c , d , e , and f are the weight of each item, Q is the state weight matrix, and R is the control weight matrix. Using Hamilton minimum principle, the state feedback matrix can be obtained by solving (8) for analytic solution. Note that the selection of proper weight matrix is very important for solving the equation. There are several helpful rules on the selection of Q and R : (1) Q and R are symmetrical matrixes, in which Q is a half positive definite matrix and R is a positive definite matrix; (2) because Q and R are interrelated, once R is determined firstly, Q can be determined by simulation and test. In addition, state feedback matrix is limited by the boundary conditions, namely, (7) in this paper.

In state weight matrix Q , b is very important, which affects the stability of the close-loop control system directly. To make control system stable, b must be properly selected to make sure that all poles of the close-loop control system are in left quadrant. In addition, it can be seen that the larger the b , the more stable the control system.

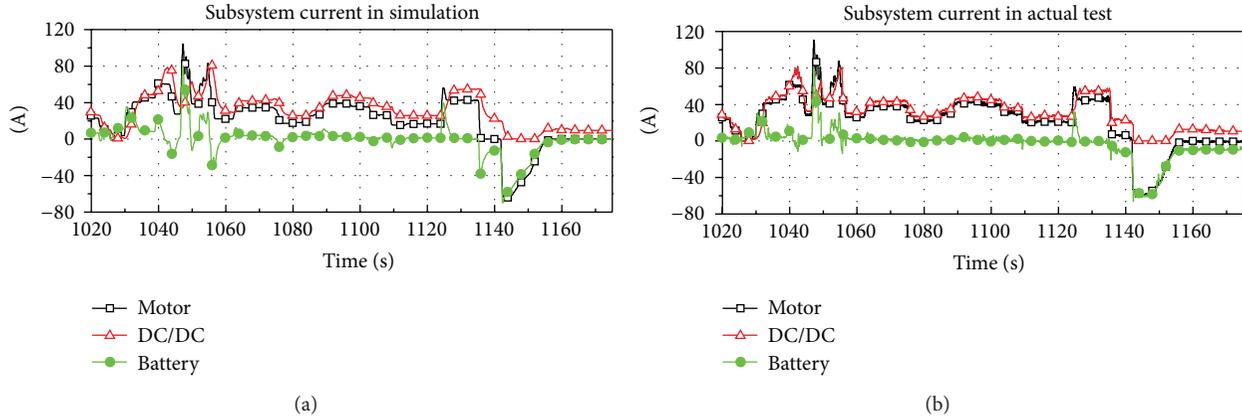


FIGURE 10: Changing process of subsystem current (part of the ECE_EUDC cycle).

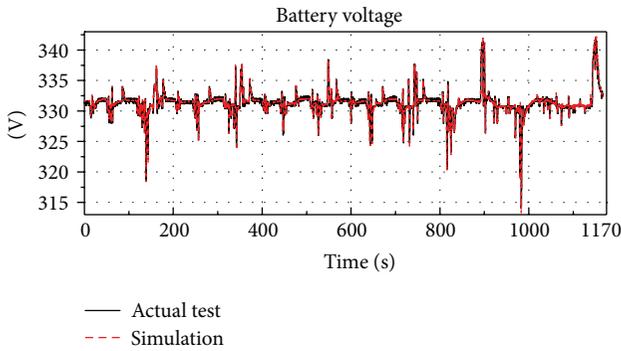


FIGURE 11: Changing process of battery voltage (ECE_EUDC cycle).

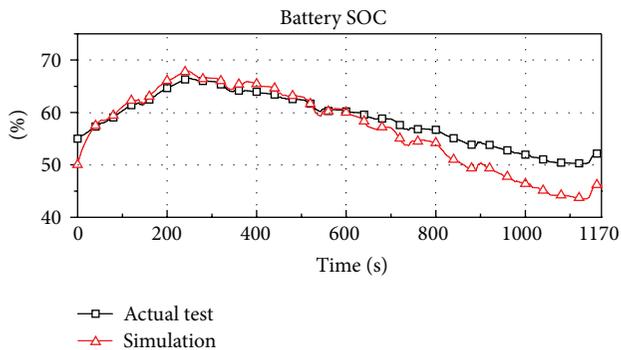


FIGURE 12: Changing process of battery SOC (ECE_EUDC cycle).

Figure 8 presents the exemplified system step responding curves. It can be seen that the larger b leads to the faster converging behavior of U_E . However, if b is over large, U_{bus} will become vibration. When b equals 2, U_E and U_{bus} tend to balance position with moderate speed.

4. Verification of the Control Strategy

Simulation and practical driving test on the designed control strategy with three driving cycles are conducted. The driving

cycle includes J1015, UDDS, and NEDC. Exemplified simulation and test data from NEDC driving cycle are presented in Figures 9–12. Figure 9 shows the changing process of vehicle velocity in the simulation and practical test during the NEDC cycle, indicating that the vehicle's power performance can be ensured using the designed strategy. Figure 10 shows the changing process of the subsystem current. Figure 11 shows the changing processes of the battery voltage and SOC during the NEDC driving cycle, where the battery voltage and SOC fluctuate within the prescriptive range, indicating that the proposed control strategy can guarantee the stable control of the FCHV power train.

5. Conclusions

In summary, a LQR-based control strategy designing method for FCHV power train is proposed in this study. A multistate feedback control strategy is designed and applied in the power train of SHANGHAI FCHV. In this control strategy, a KF observer is introduced to estimate battery OCV and the state feedback gain matrix of the multistate feedback control system is computed by the LQR algorithm. To verify the designed control strategy, simulation and practical driving test with three types of driving cycles are conducted. The results of the simulation and practical testing show that the designed LQR-based power control strategy can ensure power performance of the vehicle and guarantee the stable control of the hybrid power train, which proves that the designed control strategy works excellently.

The LQR-based control design method presented in this research well suits to control development for FCHV power train. And it provides valuable reference to the model-based control strategy development for other types of hybrid power train.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The paper is supported by the National Natural Science Foundation of China (no. 51205215).

the IEEE International Conference on Vehicular Electronics and Safety (ICVES '06), pp. 342–347, Shanghai, China, December 2006.

References

- [1] H. S. Lee, K. S. Jeong, and B. S. Oh, "An experimental study of controlling strategies and drive forces for hydrogen fuel cell hybrid vehicles," *International Journal of Hydrogen Energy*, vol. 28, no. 2, pp. 215–222, 2003.
- [2] G. Pede, A. Iacobazzi, S. Passerini, A. Bobbio, and G. Botto, "FC vehicle hybridisation: an affordable solution for an energy-efficient FC powered drive train," *Journal of Power Sources*, vol. 125, no. 2, pp. 280–291, 2004.
- [3] M. Mohammadian, S. M. T. Bathaee, and S. M. M. Ansarey, "Neuro-genetic energy management for hybrid fuel cell power train," in *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, pp. 1043–1048, Singapore, December 2004.
- [4] M.-J. Kim and H. Peng, "Power management and design optimization of fuel cell/battery hybrid vehicles," *Journal of Power Sources*, vol. 165, no. 2, pp. 819–832, 2007.
- [5] Y. Eren, O. Erdinc, H. Gorgun, M. Uzunoglu, and B. Vural, "A fuzzy logic based supervisory controller for an FC/UC hybrid vehicular power system," *International Journal of Hydrogen Energy*, vol. 34, no. 20, pp. 8681–8694, 2009.
- [6] J. Ryu, Y. Park, and M. Sunwoo, "Electric powertrain modeling of a fuel cell hybrid electric vehicle and development of a power distribution algorithm based on driving mode recognition," *Journal of Power Sources*, vol. 195, no. 17, pp. 5735–5748, 2010.
- [7] J. P. Torreglosa, F. Jurado, P. García, and L. M. Fernández, "Hybrid fuel cell and battery tramway control based on an equivalent consumption minimization strategy," *Control Engineering Practice*, vol. 19, no. 10, pp. 1182–1194, 2011.
- [8] K. Isamu, "Electric power control of fuel cell vehicle," European Patent WO2008004564, January 2008.
- [9] U. S. Kwon, S. I. Jeon et al., "Method for controlling fuel cell vehicle," US Patent Application 20090222157, September 2009.
- [10] U. Limbeck, S. Schmalzriedt et al., "Method and control unit for automatic selection of an operating mode for a vehicle with fuel cells," US Patent Application 20110196554, August 2011.
- [11] J. T. Pukrushpan, H. Peng, and A. G. Stefanopoulou, "Simulation and analysis of transient fuel cell system performance based on a dynamic reactant flow model," in *Proceedings of the ASME International Mechanical Engineering Congress and Exposition (IMECE '02)*, pp. 637–648, November 2002.
- [12] S. Jemei, D. Hissel, M. C. Péra, and J. M. Kauffmann, "On-board fuel cell power supply modeling on the basis of neural network methodology," *Journal of Power Sources*, vol. 124, no. 2, pp. 479–486, 2003.
- [13] C. N. Maxoulis, D. N. Tsinoglou, and G. C. Koltsakis, "Modeling of automotive fuel cell operation in driving cycles," *Energy Conversion and Management*, vol. 45, no. 4, pp. 559–573, 2004.
- [14] M. Joong Kim, H. Peng, C. Chiao et al., "Testing, modeling, and control of a fuel cell Hybrid Vehicle," in *Proceedings of the American Control Conference*, pp. 3859–3864, 2005.
- [15] H. Yun, Y. Zhao, and J. Wang, "Modeling and simulation of fuel cell hybrid vehicles," *International Journal of Automotive Technology*, vol. 11, no. 2, pp. 223–228, 2010.
- [16] H. Dai, Z. Sun, and X. Wei, "Online SOC estimation of high-power lithium-ion batteries used on HEVs," in *Proceedings of*

Research Article

Particle Filter for Estimating Freeway Traffic State in Beijing

Jun Bi,¹ Can Chang,¹ and Yang Fan²

¹ School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

² Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA

Correspondence should be addressed to Jun Bi; blinghc@163.com

Received 15 July 2013; Revised 10 October 2013; Accepted 24 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Jun Bi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Freeway traffic state estimation is useful for intelligent traffic guidance, control, and management. The freeway traffic state is featured with rapid and dramatic fluctuations, which presents a strong nonlinear feature. In theory, a particle filter has good performance in solving nonlinear problems. This paper proposes a particle filter based approach to estimate freeway traffic state. The freeway link between the west of Peace Bridge and the west of San Yuan Bridge of the third ring in Beijing is used as the experimental object. According to the traffic characteristics and measurement mode of the link, the second-order validated macroscopic traffic flow model is adopted to set up the link model. The implementation steps of the particle filter for freeway traffic state estimation are described in detail. The estimation error analysis for the experiments proves that the proposed approach has an encouraging estimation performance.

1. Introduction

The increasing traffic flow is resulting in serious congestion of urban road network, which can decrease flow rate, delay travel time, increase fuel consumption and travel cost, and make negative environmental effect. Measures should be taken to alleviate traffic congestion. Traffic state estimation of freeway network is useful for traffic management, which involves estimating the traffic variables of the network based on available real-time traffic measurements [1]. A real-time Lagrangian traffic state estimator for freeway state has been proposed, which is considered to be more accurate and more computationally efficient than the Eulerian approach [2, 3]. However, the Lagrangian method is only appropriate for the sensing data obtained via GPS technology or other tracking devices providing position and velocity of individual vehicles. Chao Deng et al. present an approach using cluster analysis and multiclass support vector machine to estimate freeway traffic state. Historical traffic flow data are divided into clusters with different traffic states and the multiclass support vector machine is applied to identify the real-time traffic states [4]. Many previous researches in estimating aggregated traffic variables are based on applications of Kalman filter or extended Kalman filter [5–7]. Furthermore, most of them assume that noises obey Gaussian distribution and use

linear models for state functions and observation functions. However, traffic state can fluctuate rapidly and dramatically in a short time, which indicates the strong nonlinear features of the freeway traffic state. Therefore, using Kalman filter may cause inaccurate estimation results and even divergent results. Particle filter is a powerful approximate solution to a general nonlinear problem or a non-Gaussian filtering problem [8, 9]. The basic idea of particle filter is that a posterior probability density function (PDF) of state can be represented by a set of particles with associated weights, and the estimation can be computed as the expected value of the discrete PDF [10, 11]. Currently only limited number of papers have discussed freeway traffic state estimation using particle filter algorithm. Mihaylova and Boel [12] proposed a particle filter (PF) to estimate freeway traffic based on the model of aggregated states and observations, and the investigations are conducted using the real traffic data from a Belgian freeway. Nicolae Marinica developed a particle filter (PF) state estimator using a platoon based model for urban traffic networks [13]. This paper uses the second-order validated macroscopic traffic flow model to evaluate a freeway link in Beijing according to its characteristics and proposes a particle filter method for estimating the freeway speed and density.

The rest of the paper is organized as follows. Section 2 describes a freeway traffic flow model in Beijing. Section 3

presents the design of the traffic state estimator based on the particle filter algorithm. In Section 4, several experiments are conducted to evaluate the particle filter estimation performance. Finally, the conclusions are summarized in Section 5.

2. Traffic Flow Modeling of a Freeway

2.1. Description of a Freeway. A freeway network can be represented as a directed graph. Bifurcations, junctions, on-ramps, and off-ramps of a freeway are represented as nodes, and a freeway stretch between two nodes is represented as a link in the graph [12]. A bidirectional freeway stretch is modeled as two opposite directional links. As shown in Figure 1, the freeway is divided into several links and each link is composed of several sections. In the paper, each link is assumed to have homogeneous geometric characteristics such as the number of lanes, slopes, and curvatures. Detectors are available only at some boundaries between sections. In Figure 1, q_i is the average number of vehicles at the boundary between sections i and $i + 1$, d_i is the average density in section i , v_i is the average speed in section i , s_i is the traffic flow into ramp i , and r_i is the traffic flow off ramp i . The whole state of the link m at time k is described by the vector $x_k = (x_{1,k}^T, x_{2,k}^T, \dots, x_{i,k}^T)^T$, and $x_{i,k} = (v_{i,k}, d_{i,k})^T$ is a local state vector of the i th section.

The progression from one sampling time to the next sampling time is described by the update equation

$$x_{k+1} = f(x_k, p_k, q_k^{\text{in}}, v_k^{\text{in}}, q_k^{\text{out}}, v_k^{\text{out}}, \xi_k), \quad (1)$$

where p_k denotes the vector of all time-varying parameters such as road conditions and the number of available lanes, q_k^{in} is the number of vehicles entering Section 1 during the interval $[t_k, t_{k+1})$, v_k^{in} is the average speed of these vehicles, q_k^{out} specifies the outflow at the speed v_k^{out} from section N , and ξ_k is a disturbance vector reflecting random fluctuation in the traffic states and the model error.

Noisy measurements of the average speed of vehicles crossing the boundary between section i and section $i + 1$ during the time interval $[t_k, t_{k+1})$ together with noisy measurements of the mean density of the vehicles are collected as the measurement data. The observation equation is given as

$$y_{k+1} = h[x_s, \eta_k, s \in [t_k, t_{k+1})], \quad (2)$$

where x_s is the state at time s and η_k is the measurement noise.

2.2. Macroscopic Traffic Flow Model of a Freeway. The second-order validated macroscopic traffic flow model is employed to describe the dynamic behavior of traffic flow along a freeway stretch in terms of appropriate aggregated traffic variables such as traffic density, space average speed, and traffic flow [14]. Generally, a freeway stretch is divided into a number of N sections. Assume that the length of each section is L_i , $i = 1, 2, \dots, N$, the time discretization is based on a time step T , and k ($k = 0, 1, 2, \dots$) is a discrete time index. For the section i , the stochastic nonlinear difference equations

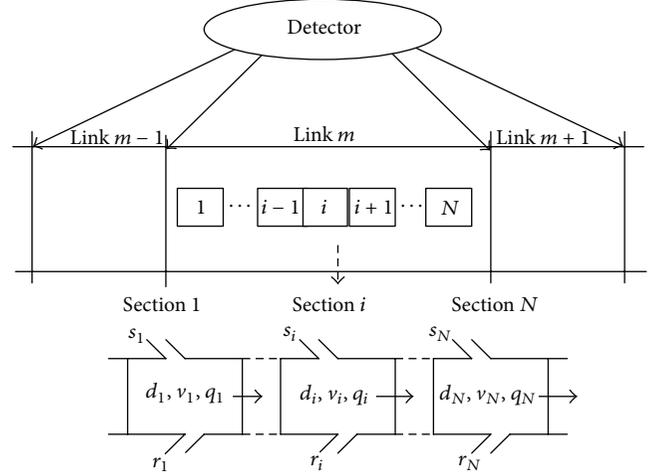


FIGURE 1: Freeway links and measurement points.

based on the second-order macroscopic traffic flow model are described as follows [14]:

$$\begin{aligned} d_i(k+1) &= d_i(k) + \frac{T}{\lambda_i L_i} \\ &\quad \times [q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)], \\ s_i(k) &= \beta_i(k) q_{i-1}(k), \\ v_i(k+1) &= v_i(k) + \frac{T}{\tau} [v(d_i(k)) - v_i(k)] \\ &\quad + \frac{T}{L_i} v_i(k) [v_{i-1}(k) - v_i(k)] \\ &\quad - \frac{nT}{\tau L_i} \frac{[d_{i+1}(k) - d_i(k)]}{d_i(k) + \kappa} \\ &\quad - \frac{\delta T}{\lambda_i L_i} \frac{r_i(k) v_i(k)}{d_i(k) + \kappa} + \xi_i^v(k), \\ v(d) &= v_f \exp \left[-\frac{1}{a} \left(\frac{d}{d_{cr}} \right)^a \right], \\ q_i(k) &= d_i(k) v_i(k) \lambda_i + \xi_i^q(k). \end{aligned} \quad (3)$$

In (3), the variables and parameters are denoted as follows:

- (1) traffic density $d_i(k)$ (in veh/km/lane) is the number of vehicles in section i at time kT divided by the section length L_i and the number of lanes λ_i ,
- (2) space average speed $v_i(k)$ (in km/h) is the average speed of the vehicles in section i at time kT ,
- (3) traffic flow $q_i(k)$ (in veh/h) is the number of vehicles leaving section i during the interval $[kT, (k+1)T)$, divided by T ,
- (4) $r_i(k)$ is the on-ramp inflow and $s_i(k)$ is the off-ramp outflow (both in veh/h) in section i .

- (5) τ , n , κ , a , and λ are the model parameters which are the same given values for all sections,
- (6) v_f denotes the free flow speed, d_{cr} is the critical density, and a is the exponent of the stationary speed equation,
- (7) $\xi_i^v(k)$ and $\xi_i^q(k)$ reflect the model error, $\xi_i^v(k)$ denotes the zero-mean Gaussian white noise acting on the empirical speed equation, and $\xi_i^q(k)$ denotes the zero-mean Gaussian white noise acting on the approximate flow equation.

The detectors in the link entrance or exit can collect the speed data and density data. The speed measurement function is

$$Z_{in\ or\ out}^v(k) = v_{in\ or\ out}(k) + \xi_{in\ or\ out}^v(k), \quad (4)$$

where $Z_{in\ or\ out}^v(k)$ is the measurement value of the average speed through entrance or exit during the interval $[(k-1)T, kT)$ and $\xi_{in\ or\ out}^v(k)$ is the measurement noise of the speed.

The measurement function of the density for the link entrance or exit is

$$Z_{in\ or\ out}^d(k) = d_{in\ or\ out}(k) + \xi_{in\ or\ out}^d(k), \quad (5)$$

where $Z_{in\ or\ out}^d(k)$ is the measurement value of the mean density through entrance or exit during the interval $[(k-1)T, kT)$ and $\xi_{in\ or\ out}^d(k)$ is the measurement noise of the density.

3. Particle Filter for Freeway Traffic State Estimation

Nonlinear characteristics exist in the freeway traffic state, which makes it difficult to estimate the traffic state. Since particle filter is able to solve a general nonlinear problem in theory, particle filter is well studied to estimate the traffic state based on the second-order validated macroscopic traffic flow model.

3.1. Particle Filter Theory. The discrete-time stochastic model of a dynamic system is described as follows:

$$\begin{aligned} x_k &= f_k(x_{k-1}, w_{k-1}), \\ y_k &= h_k(x_k, v_k), \end{aligned} \quad (6)$$

where y_k is the observation vector at time k , x_k represents the state vector at time k , f_k is the nonlinear state transition function, and h_k is the nonlinear observation function. The stochastic processes w_k and v_k represent the state noise process and the measurement noise process, respectively. The available information at time k is the set of measurements $Z_k = \{y_i : i = 1, 2, \dots, k\}$. According to the Bayesian theory, all state information can be obtained from the posterior state distribution. Within the Bayesian framework, particle filter is used to sequentially update a priori knowledge about predetermined state vector x_k by using the measurements

data Z_k . Suppose that the initial probability density function of the state vector is $p(x_0 | Z_0) = p(x_0)$; then we have [9, 10]

$$p(x_k | Z_{k-1}) = \int p(x_k | x_{k-1}) p(x_{k-1} | Z_{k-1}) dx_{k-1}, \quad (7)$$

$$p(x_k | Z_k) = \frac{p(y_k | x_k) p(x_k | Z_{k-1})}{p(y_k | Z_{k-1})}, \quad (8)$$

$$p(y_k | Z_{k-1}) = \int p(y_k | x_k) p(x_k | Z_{k-1}) dx_k. \quad (9)$$

Particle filter uses Monte Carlo techniques and sequence importance sampling (SIS) methods to solve (7). The posterior PDF is computed based on N samples from an important distribution function, as follows,

$$p(x_k | Z_k) = \sum_{i=1}^N w_k^i \delta(x_k - x_k^i), \quad (10)$$

$$w_k^i = \frac{\bar{w}_k^i}{\sum_{i=1}^N \bar{w}_k^i}, \quad (11)$$

$$\bar{w}_k^i = \bar{w}_{k-1}^i \frac{p(y_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, y_k)}, \quad (12)$$

where w_k^i is the weight of each particle and satisfies $0 \leq w_k^i \leq 1$ and q is an important distribution function.

Particle filter has good effects in solving both nonlinear and non-Gaussian applications, which is suitable for estimating the nonlinear freeway state.

3.2. Application to Freeway Traffic State Estimation. The application steps of freeway traffic state estimation based on particle filter are as follows.

(1) *Initialization.* Set the iteration variable $k = 0$. The particle set $\{x_0^i, w_0^i\}$ ($i = 0, 1, 2, \dots, N$, N is the number of particles) is obtained based on the initial distribution $p(x_0)$.

(2) *Particle Generation.* For $i = 1, 2, \dots, N$, x_k^i is sampled from $p(x_k | x_{k-1}^i, Z_k)$ according to (3).

(3) *Weight Computation.* Equation (12) is complex to compute the weight, and usually the weight is updated by the following equation when the new measurement data are obtained by the detector in the section:

$$\bar{w}_k^i = \bar{w}_{k-1}^i \times p(y_k | x_k^i). \quad (13)$$

The normalized weight w_k^i is computed using (11).

(4) *Resampling.* Drawing a random sample u_i from the uniform distribution over $(0, 1]$. If u_i satisfies $\sum_{j=1}^{m-1} w_k^j < u_i \leq \sum_{j=1}^m w_k^j$, the m th particle is resampled as follows:

$$\begin{aligned} \tilde{x}_k^i &= x_k^m, \\ \tilde{w}_k^i &= w_k^m. \end{aligned} \quad (14)$$

(5) *State Estimation.* The approximate posterior PDF is computed using (10); then the state is estimated by

$$\hat{x}_k = \sum_{i=1}^N \tilde{w}_k^i \tilde{x}_k^i. \quad (15)$$

(6) *k Is increased by Itself.* If the end condition of algorithm is not satisfied, the algorithm goes to step (2) again; otherwise, it ends.

4. Freeway State Estimation Experiments

4.1. Experimental Object. A freeway between the west of Peace Bridge and the west of San Yuan Bridge of third ring in Beijing is used as an experimental link. The link has no ramps. The length of the link is 1476 m. The link is composed of 3 sections with 3 lanes and the length of each section is 556 m, 475 m, and 445 m, respectively. The measurement detectors are located in the boundaries between the first section and the third section. The detectors can collect the traffic flow, speed, and density of the link every 2 min.

One example of the entrance traffic state during 3 hours between 8:00 and 11:00 AM on July 23, 2010, is shown in Figure 2. We know that the traffic flow varied between 40 veh/2 min and 60 veh/2 min, the speed varied between 32 km/h and 50 km/h, and the density varied between 23 veh/km and 58 veh/km. In most of the time, the traffic state varied smoothly except for a rapid change at 8:20 AM.

4.2. Experiments and Analysis. The state vector x_k and measurement vector Z_k are computed, respectively, as follows:

$$x_k = [v_1(k), d_1(k), v_2(k), d_2(k), v_3(k), d_3(k)]^T, \quad (16)$$

$$Z_k = [Z_{in}^v(k), Z_{in}^d(k), Z_{out}^v(k), Z_{out}^d(k)].$$

The parameters of traffic model in Section 2 are set as follows.

According to the experimental object in Section 4.1, each section of the freeway has 3 lanes, and the length of each section is about 500 m. The data detection period is 2 min. In daytime, the free flow speed of the freeway is about 80 km/h and the critical density is about 50 veh/km. Therefore, some of the parameters can be set as $\lambda_i = 3$, $L_i = 500$ m, $T = 120$ s, $v_f = 80$ km/h, and $d_{cr} = 50$ veh/km. Based on the prior knowledge, other parameters can be set as $\tau = 20$ s, $n = 45$ km²/s, and $\kappa = 15$ veh/km/lane.

According to the data analysis in Figure 2, the initial state is set to be

$$x_0 = [42, 40, 50, 35, 48, 35]^T. \quad (17)$$

The number of particles N is set to be 500.

The freeway speed and density are estimated using the particle filter algorithm in the Section 3. The estimation results for the speed and the density of each section are, respectively, shown in Figures 3 and 4. In the horizontal axis of each figure, the step time interval is 2 min which is equal to the data collecting interval of the detectors. According to Figures 3 and 4, the estimation trend is similar to the true trend.

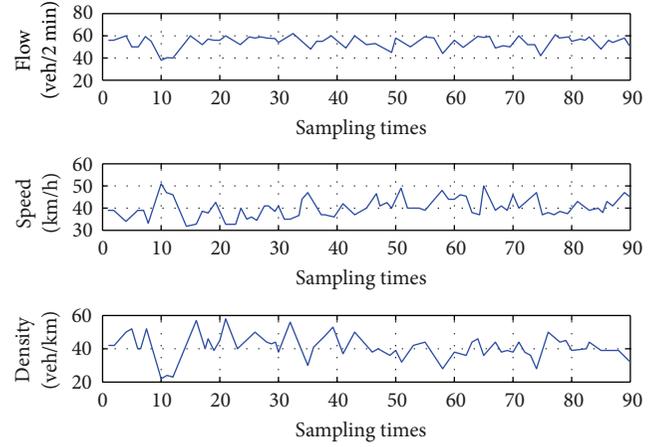


FIGURE 2: Traffic state example of exit in the link.

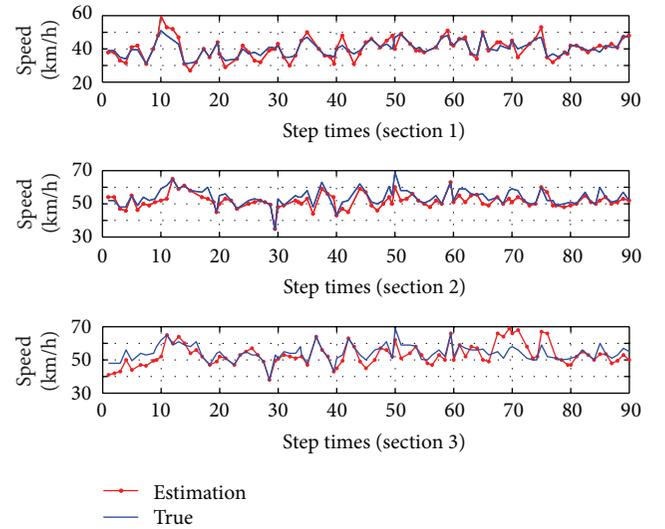


FIGURE 3: Speed estimation for three sections.

The indices of relative error (RE), mean square error (MSE), and root mean square relative error (RMSRE) are utilized to evaluate estimation performance.

The indices of RE, MSE, and RMSE are, respectively, defined as follows:

$$RE = \frac{x_i - \hat{x}_i}{x_i} \times 100\%,$$

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2}, \quad (18)$$

$$RMSRE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \hat{x}_i}{x_i} \right)^2},$$

where x_i represents the true state and \hat{x}_i represents the estimation state.

The REs of speed estimation and density estimation for each section are, respectively, shown in Figures 5 and 6. In the horizontal axis of each figure, each step time interval is 2 min.

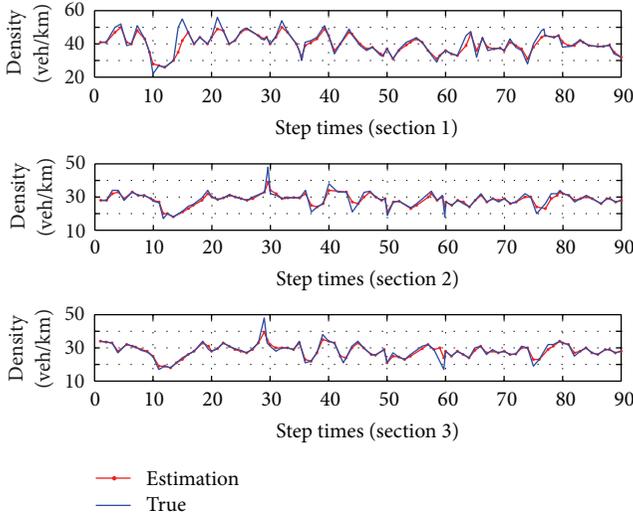


FIGURE 4: Density estimation for three sections.

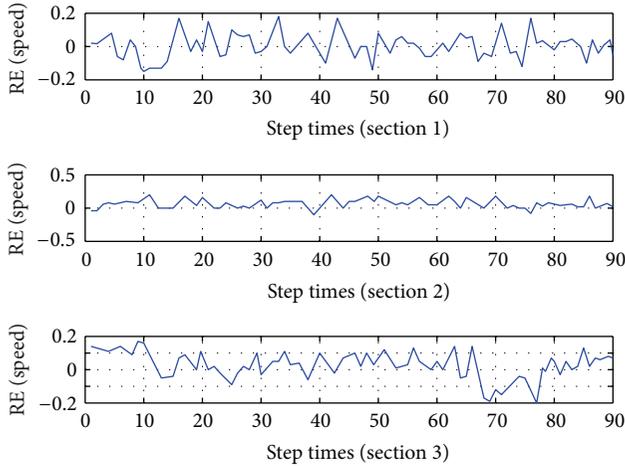


FIGURE 5: REs of speed estimation for three sections.

According to Figures 5 and 6, the RE is in $[-10\%, +10\%]$ and varies smoothly in most of the time. Only at several time intervals, the absolute RE is more than 15%.

The MSE and RMSRE of speed estimation and the MSE and RMSRE of density estimation for each section are shown in Table 1. The maximum MSE difference of speed estimation among the sections is 1.8675 and the maximum RMSRE difference of speed estimation among the sections is 0.0156. The maximum MSE difference of density estimation among the sections is 1.0273 and the maximum RMSRE difference of density estimation among the sections is 0.0062. The mean MSE of the speed estimation is 3.7136, the mean RMSRE of the speed estimation is 0.0750, the mean MSE of the density estimation is 2.6859, and the mean RMSRE of the density estimation is 0.0751. Therefore, the estimation effect is desirable.

In theory, the resampling procedure in particle filter has a complexity of $O(N)$; N is the number of sample variables generated from the uniform distribution [11]. In fact, the computation time can be reduced in the particle filter implementation by using the multithread programming technology.

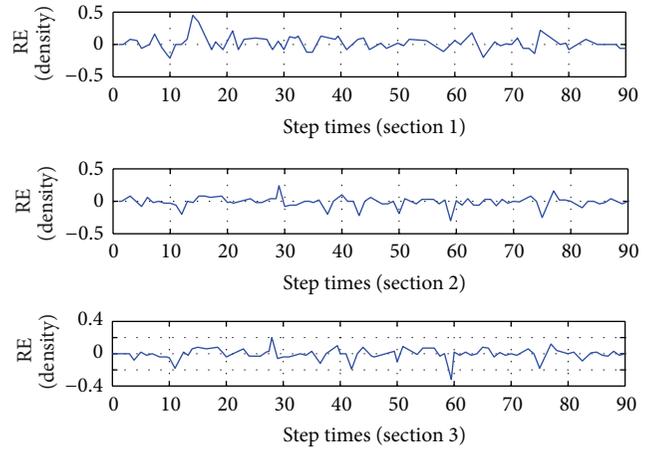


FIGURE 6: REs of density estimation for three sections.

TABLE 1: MSE and RMSRE of estimation of density and speed.

Error index	State variable			
	Speed		Density	
	MSE	RMSRE	MSE	RMSRE
Section 1	2.8269	0.0697	3.0131	0.0786
Section 2	3.6196	0.0700	3.0360	0.0742
Section 3	4.6944	0.0853	2.0087	0.0724
Mean	3.7136	0.0750	2.6859	0.0751

5. Conclusion

This paper proposes an approach to estimate the freeway traffic state based on the particle filter algorithm. The freeway traffic is modeled by the second-order validated macroscopic traffic flow model. A freeway link between the west of Peace Bridge and the west of San Yuan Bridge of third ring in Beijing is used as an experimental object and real traffic data are used in the experiments. The mean square error of the speed estimation is 3.7136 and the mean square error of the density estimation is 2.6859. The results suggest that particle filter is valid and effective in freeway traffic state estimation. In addition, the proposed approach for traffic state estimation is modular, and therefore different traffic models can be used in different sections on a freeway link. In the future research, we will use parallel computation to improve the performance of the particle filter for estimating freeway state.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper is sponsored by the National 863 Program (2011AA110303) and the Fundamental Research Funds for the Central Universities (2013JBM052).

References

- [1] Y. B. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: a case study," *Transportation Science*, vol. 41, no. 2, pp. 167–181, 2007.
- [2] C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transportation Research B*, vol. 44, no. 4, pp. 460–481, 2010.
- [3] Y. F. Yuan, J. W. C. van Lint, R. E. Wilson, and F. van Wageningen-Kessels, "Real-time lagrangian traffic state estimator for freeways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 59–70, 2012.
- [4] C. Deng, F. Wang, and H. M. Shi, "Real-time freeway traffic state estimation based on cluster analysis and multiclass support vector machine," in *Proceedings of the 1st International Workshop Intelligent Systems and Applications*, pp. 1–4, May 2009.
- [5] Y. C. Zhao, Y. C. Liu, L. Q. Shan, and B. Zhou, "Dynamic analysis of Kalman filter for traffic flow forecasting in sensornets," *Information Technology Journal*, vol. 11, no. 10, pp. 1518–1522, 2012.
- [6] Y. B. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: adaptive capabilities and real data testing," *Transportation Research A*, vol. 42, no. 10, pp. 1340–1358, 2008.
- [7] C. P. I. J. van Hinsbergen, T. Schreiter, and F. S. Zuurbier, "Localized extended kalman filter for scalable real-time traffic state estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 385–394, 2012.
- [8] K. Nummiaro, E. Koller-Meier, and L. van Gool, "An adaptive color-based particle filter," *Image and Vision Computing*, vol. 21, no. 1, pp. 99–110, 2003.
- [9] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo in Practice*, Springer, New York, NY, USA, 2001.
- [10] J. Bi, Y. F. Fung, T. K. Ho, and B.H. Mao, "Parallel resampling particle filter algorithm," *Journal of Computational Information Systems*, vol. 7, no. 6, pp. 1838–1845, 2011.
- [11] W. Yi, M. R. Morelande, L. J. Kong, and J. Y. Yang, "A computationally efficient particle filter for multitarget tracking using an independence approximation," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 843–856, 2013.
- [12] L. Mihaylova and R. Boel, "A particle filter for freeway traffic estimation," in *Proceedings of the 43rd IEEE Conference on Decision and Control (CDC '14)*, pp. 2106–2111, December 2004.
- [13] N. Marinica and R. Boel, "Particle filter for platoon based models of urban traffic," in *Proceedings of the 15th WSEAS International Conference on Systems*, pp. 42–47, July 2011.
- [14] M. Papageorgiou, "Some remarks on macroscopic traffic flow modelling," *Transportation Research A*, vol. 32, no. 5, pp. 323–329, 1998.

Research Article

Effects of Transverse Baffle Design on Reducing Liquid Sloshing in Partially Filled Tank Vehicles

Xue-lian Zheng, Xian-sheng Li, Yuan-yuan Ren, Yu-ning Wang, and Jie Ma

College of Traffic, Jilin University, Changchun 130022, China

Correspondence should be addressed to Yuan-yuan Ren; maggie170101@gmail.com

Received 17 July 2013; Revised 12 October 2013; Accepted 20 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Xue-lian Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Longitudinal liquid sloshing in partially filled clear-bore tanks causes extensive degradation of tankers braking performance. To reduce the negative effect of longitudinal liquid sloshing on tankers, three kinds of transverse baffles were designed, namely, the conventional baffle, the circular baffle, and the staggered baffle. Each kind of baffle took several forms to investigate the impact of baffle installation angle, the sizes of holes pierced on the baffle, and their arrangement on the antisloshing effect. FLUENT software was used to simulate liquid sloshing in tanks equipped with different kinds of transverse baffles and subject to constant braking deceleration. A time-series analysis of the forces act on tank walls and transverse baffles was carried out. It was drawn that the baffle shape and its installation angle have great impact on the antisloshing effect of baffles. The study on the antisloshing effect of different transverse baffles is of great significance for tank vehicle driving and braking safety, as well as for the design of optimal transverse baffles.

1. Introduction

Tank vehicles are widely used in the road transportation of liquid cargoes throughout the world. According to the 2007 Commodity Flow Survey, in the United States, about 1200 million tons of hazardous fluid materials were shipped by tankers per year [1]. In China, about 76% of chemical products were delivered by tank trucks, and the total freight load was nearly 1 million tons. Tankers are hugely convenient for transporting liquid cargoes and boosting national economic development. However, many traffic accidents occurred, causing losses in terms of human life and property. In 2011, 416 tanker accidents occurred in China, causing 643 deaths and costing tens of millions of Yuan. On the basis of the reported highway accidents involving heavy vehicles, it has been suggested that tank trucks were 4.8 times more likely to be involved in rollovers [2]. Due to the particularity of transportation cargoes, most of the liquid release tanker accidents made direct contamination to water, air, and land [3, 4].

A lot of work has been done to investigate the characteristics of tanker accidents. Based on the accident cause, researchers divided them into the following accident types,

namely, rollover, rear-end, crash, and over-speeding. Furthermore, the frequencies for each cause were about 45.16%, 25.81%, 17.74%, and 11.29%, respectively [5]. Hence, rollover and rear-end are the main forms of tanker accidents.

There are many reasons for tanker accidents [6], but liquid sloshing in partially filled tanks is the most common one. Due to the difference on cargo density and the limitations on vehicle axle load, tank vehicles tend to be operated in the partially filled condition for the majority of the time [7, 8]. Both the existence of sloshing space and the flow characteristic of liquid cargo make contribution to liquid sloshing when vehicle driving state changes. For moving tanks, a minor change in the vehicle driving state may result in a gigantic surge in the liquid, bringing about liquid sloshing and a large change in the amplitude of the dynamic load [7, 8]. In addition, the free surface oscillation of a low-viscosity liquid will last for a long time. While tank vehicles commonly subject to intense driver operations in practice, which make the liquid sloshing in partially filled tanks quite violent, the vehicle's handling stability and braking performance are degraded greatly.

Through experimental discovery and theoretical analysis, it was found that tank geometry and liquid fill percentage are

the main factors impacting the liquid dynamic load transfer, the oscillation amplitude, and the sloshing duration [9, 10]. To limit liquid sloshing in the longitudinal direction during vehicle braking and acceleration, transverse baffles were designed and have been proved to be quite efficient. To date, the research on transverse baffles can be summarized into the following two aspects.

(1) How do transverse baffle reduce liquid sloshing? To solve this problem, various numerical models have been developed to describe liquid sloshing in tanks with and without transverse baffles [11–15]. Usually, a rectangular tank was chosen as the research object. The fluid fill level was constant while the height of the vertical baffle increased gradually. For vertical baffles of different heights, snap shots of fluid-free surface profiles and elevations, flow patterns, streamlines, and pressure distributions were studied at the same external excitation [11–13]. It has been found that a vortex is originated near the baffle tip and a flow separation is generated before and behind the baffle when liquid sloshing occurs [3]. Moreover, the vortex becomes weaker as the baffle height increases. Compared with liquid sloshing in a clean-bore tank, the gradient of the liquid free surface and the maximum height that the fluid can reach on the tank wall were decreased in tanks with vertical baffles. This leads to smaller movements in the center of gravity of liquid bulk and the enhancement of vehicle braking and acceleration performance.

(2) The influence of the number, appearance, baffle arrangement, and the number and location of holes pierced in the baffle on the antisloshing effect, as in [16–25]. FLUENT simulation and experiment were utilized in this kind of research. The researchers designed kinds of transverse baffle to evaluate their antisloshing effect, including conventional, round, solid-dished, oblique, perforated, and spiral baffles [17–20]. In the simulation process, the liquid fill level was assumed to be constant (50% was the most popular fill level used) while braking deceleration changes. A time-series analysis of the forces on the tank walls was carried out to investigate the effect of the baffle in terms of reducing liquid sloshing. For all the work that has been carried out on this subject, three problems should be mentioned: First of all, most of the baffles were designed arbitrarily, and national standards on baffle design were not considered. Secondly, the baffle effect on reducing liquid sloshing at different fill levels were not considered yet, while in practice different containers have quite different liquid fill levels due to difference on liquid density and the limitations on road axle load. Thirdly, only sloshing forces on the tank walls were paid attention to, and force analysis of the baffles was not mentioned at all.

Therefore, the purpose of this paper is to design different forms of transverse baffle based on the current Chinese standards then investigate which one has the optimal effect on reducing liquid sloshing. To achieve this purpose, three kinds of transverse baffle were designed and several forms of each kind were used to investigate the impact of the installation angle, the sizes of holes, and their arrangement on the antisloshing effect. Then, FLUENT was used to simulate liquid sloshing in partially filled tanks that are subject to constant braking deceleration, and a time-series analysis of the forces

act on the tank walls and the baffles was carried out considering structural damage to both the tank and the baffles.

2. Materials and Methods

Tanks with a circular cross-section have greater volumes but the same surface area, which makes them lighter. Therefore, they are the most popular and were used in this paper. Three kinds of baffle were designed and FLUENT was used to simulate longitudinal liquid sloshing in partially filled tanks with a circular cross-section and equipped with baffles.

2.1. Theory of Similarity. The hydrodynamic equations are solved by FLUENT using a finite-volume method. In this method, the 2- or 3-dimensional tank model is split into thousands of meshes at first and then the hydrodynamic equations are solved in each of them. Therefore, the bigger the tank model is, the more time the simulation or calculation procedure needs. For time saving, the full-sized tank was scaled down to a tank model.

The criteria of geometric similarity, kinematic similarity, and dynamic similarity must be satisfied to ensure that liquid sloshing in the full-sized tank and that in the tank model exhibit similar hydrodynamic behavior. The full-sized tank is given the subscript t and the tank model m . According to the geometric similarity, the length-scale factor can be calculated as follows:

$$\delta_L = \frac{L_t}{L_m}, \quad (1)$$

where L is the geometric size of the tank.

The velocity-scale factor that expresses the relation between the moving speed of the full-sized tank and that of the tank model can be expressed by the following equation:

$$\delta_v = \frac{v_t}{v_m}, \quad (2)$$

where v is the moving velocity of the tank.

In general, speed can be expressed as the differential of distance, which is expressed as follows:

$$v = \frac{dL}{dt}. \quad (3)$$

Therefore, according to (2) and (3), the velocity-scale factor can be reexpressed as follows:

$$\delta_v = \frac{v_t}{v_m} = \frac{dL_t/dt_t}{dL_m/dt_m} = \frac{dL_t}{dL_m} \frac{dt_m}{dt_t} = \delta_L \delta_t. \quad (4)$$

Then, the time-scale factor can be written as follows using (4):

$$\delta_t = \frac{\delta_L}{\delta_v}. \quad (5)$$

The acceleration-scale factor is drawn by differentiating velocity and is presented as follows:

$$\delta_a = \frac{\delta_v^2}{\delta_L}. \quad (6)$$

For longitudinal liquid sloshing in partially filled tank vehicles, forces due to gravity and inertia play a more important role than the viscous and turbulent shear forces. Therefore, the Froude number, which is the function of gravity and inertia, is the most relevant dimensionless parameter for liquid sloshing, which is expressed as follows:

$$F_r = \frac{v_t^2}{L_t g_t} = \frac{v_m^2}{L_m g_m}. \quad (7)$$

For the situation discussed in this paper,

$$g_t = g_m, \quad (8)$$

and the following equations can be deduced:

$$\frac{v_t^2}{v_m^2} = \frac{L_t}{L_m}, \quad (9)$$

$$\delta_v^2 = \delta_L. \quad (10)$$

Therefore, some important conclusions can be derived from (10), (5), and (6) and are presented as follows:

$$\begin{aligned} \delta_v &= \sqrt{\delta_L}, \\ \delta_t &= \sqrt{\delta_L}, \\ \delta_a &= 1. \end{aligned} \quad (11)$$

Equations in (11) are the conditions that must be satisfied to ensure that liquid sloshing in the tank model exhibits similar dynamic behavior to that in the full-sized tank.

While the liquid dynamic behavior in both the full-sized tank and the tank model are similar, the longitudinal sloshing force obtained by liquid sloshing simulation in the tank model is proportionate to that obtained in the full-sized tank. To derive the relationship of the two forces, the momentum for the longitudinal liquid sloshing is needed, which is calculated as follows:

$$\frac{\partial}{\partial t} \int_{\tau} \rho \omega d\tau + \oint_A \rho \omega (\vec{v}\vec{n}) dA = \sum F_z, \quad (12)$$

where τ is the control volume, ω is the velocity of the control volume along the z direction, ρ is liquid's density, A is the control plane, \vec{v} is the velocity of the control volume, and \vec{n} is the normal of the vertical plane.

In the case where fluid can only flow into the control volume through one of the control planes and flow out through another one, (12) can be simplified as follows:

$$Q_m (\omega_2 - \omega_1) = \sum F_z, \quad (13)$$

where Q_m is the fluid discharge in the control volume; ω_2 is fluid flow-out velocity; ω_1 is fluid flow-in velocity.

Suppose that the time interval is quite small; then (13) can be rewritten as follows:

$$\Delta Q_m \Delta \omega = \Delta F_z. \quad (14)$$

TABLE 1: Parameter values of the full-sized tank and the tank model.

	The full-sized tank	The tank model
Length	8 m	2 m
Radius of circular cross section	1.1 m	0.275 m
Height of the elliptical head	0.55 m	0.1375 m
Longitudinal deceleration	-3 m/s ²	-3 m/s ²

The transient flow rate is generally expressed by

$$\Delta \omega = \frac{\Delta Q_m}{Ad}. \quad (15)$$

Therefore, (14) can be reexpressed as follows:

$$(\Delta \omega)^2 dA = \Delta F_z. \quad (16)$$

According to (16), the ratio of the longitudinal sloshing force generated in the full-sized tank to that in the tank model in the same situation can be expressed as follows:

$$\delta_F = \frac{\Delta F_t}{\Delta F_m} = \frac{(\Delta \omega_t)^2 dA_t}{(\Delta \omega_m)^2 dA_m} = \delta_v^2 \delta_A = \delta_L \delta_L^2 = \delta_L^3. \quad (17)$$

In other words, the ratio is 1 : δ_L^3 .

2.2. Baffle Design. The dimensions and other variables of the full-sized tank and the tank model are listed in Table 1. The length-scale factor is set at 4. The two ends of the tank are covered by 2 : 1 elliptical heads. According to the current Chinese standard, GB/T 25198-2010 *Heads for pressure vessels*, the height of the head is just a quarter of the diameter and its radius varies between the major and minor axis.

In real life, the braking deceleration of a heavy truck can be expressed as follows:

$$a_x = g\varphi, \quad (18)$$

where a_x is the vehicle's longitudinal deceleration and φ is the road adhesion coefficient.

For heavy trucks, the coefficient of road adhesion is about 0.68. Therefore, the maximum longitudinal deceleration of tank trucks is approximately 6.7. In this paper, a deceleration of 3 is chosen for the tank model.

Before baffle design, the current national standard in China should be considered. The current Chinese standard, GB/T 18564-2001 *General specification for normal pressure tank body of transportation liquid dangerous goods*, states the following.

- (1) The transverse baffle should be capable of bearing the following longitudinal inertia force:

$$F_x = m \times 2g, \quad (19)$$

where m is the mass of liquid bulk.

- (2) The free area of the transverse baffle should be greater than 40% of the tank's cross-sectional area. For tanks

with a volume less than or equal to 25 m^3 , the volume between the two tank cross-sections at which adjacent baffles locate should be less than or equal to 3 m^3 ; for those with a volume greater than 25 m^3 , the volume between adjacent cross-sections can be up to 7 m^3 .

- (3) A manhole should be cut in the center of the baffle, and the top and bottom segments should be removed to allow air and fluid, respectively, to drain out.

For this paper, three kinds of baffle have been designed according to the national standard, namely, conventional, circular, and staggered baffles. The baffles are designed to fit the geometric size of the tank model.

2.2.1. The Conventional Baffle. The conventional baffle is depicted in Figure 1. Its free area is set to comprise 57.34% of the tank's cross-sectional area. The removed top segment comprises 15% and the removed bottom segment 27.66% of tank's cross-sectional area in order to enable liquid to drain away quickly and to play the role of manhole at the same time.

To investigate the influence of the baffle installation angle on liquid sloshing reduction, different oblique angles are configured, as shown in Figure 2. The oblique angle is set to be 0-, 5-, 10-, 15-, and 20-degree. The corresponding baffles are labeled N0, N5, N10, N15, and N20. It should be noted that all of the conventional baffles have the same projection area in the vertical plane.

2.2.2. The Circular Baffle. A series of circular baffles have been designed to investigate how the sizes of pierced holes influence the liquid sloshing reduction, as shown in Figure 3. The free area of all the circular baffles is set to comprise 57.34% of the tank's cross-sectional area. The removed top segment is equal to 5% of cross-sectional area and the removed bottom segment 10% in order to enable liquid to drain away quickly.

In Figure 3(a), only a manhole is cut in the center of the baffle. In Figures 3(b)–3(e), the radius of the manhole is decreased and a few small holes are arranged around it. The circular baffles plotted in Figures 3(a)–3(e) are labeled C0, C2, C4, C6, and C18, respectively, according to the number of small holes.

2.2.3. The Staggered Baffle. Unlike the conventional and circular baffle, a complete staggered baffle is divided into two parts, and the free area of each part is set to be 28.67% of the tank's cross-sectional area. The two parts look just the same, as shown in Figure 4. The baffle has the same radius as the tank's cross-section so as to the baffle match with the tank well. The two baffle parts are located in two vertical planes, and there is a longitudinal distance between them.

The angle from the X-coordinate to the line connecting the middle point of the baffle chord and the center of the tank's cross-section is defined as the arrangement angle, as shown in Figures 5–6.

The staggered baffle can be arranged in two ways: either the arrangement angles of the two adjacent baffle parts are opposite to each other, as presented in Figure 5, which is defined as the reverse staggered baffle or the arrangement

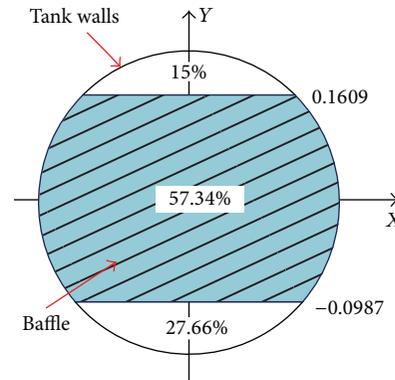


FIGURE 1: Schematic diagram of the conventional baffle.

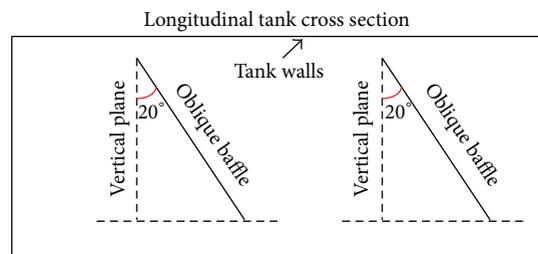


FIGURE 2: Oblique angle of conventional baffles.

angles are equal to one another, as shown in Figure 6, which is defined as the similar staggered baffle. For each option, the absolute value of the arrangement angle is set as 10-, 15-, and 20-degree in this paper, and labeled SR10, SR15, and SR20 (reverse staggered baffles) and SS10, SS15, and SS20 (similar staggered baffles).

2.2.4. The Spatial Distribution of Baffles. The volume between the two adjacent tank cross-sections at which baffles locate should meet the specified requirements. For the conventional and circular baffle, four baffles are needed, and they divide the tank into five parts, as illustrated in Figure 7. Four groups of staggered baffle made up of eight baffle parts are needed, and they divide the tank into nine parts, as Figure 8 shows.

By the principle of inertia, fluid flows forward when tanker brakes, and the fore elliptical head (the wall marked as *wall1* in Figures 7 and 8) bears the maximum sloshing force. To make the illustration of the simulation results easier, the fore and aft tank walls as well as the baffles are all labeled, as shown in Figures 7 and 8.

2.3. FLUENT Simulation Settings. Before liquid sloshing is simulated using FLUENT, tank models equipped with the designed baffles were constructed and meshed. To save on simulation time and ensure simulation accuracy at the same time, the interval size of the mesh volume was set at 0.05 units.

The *volume-of-fluid* (VOF) method was utilized to simulate the two-phase (liquid and air) flow phenomenon. The reference pressure location was arranged at any spatial point

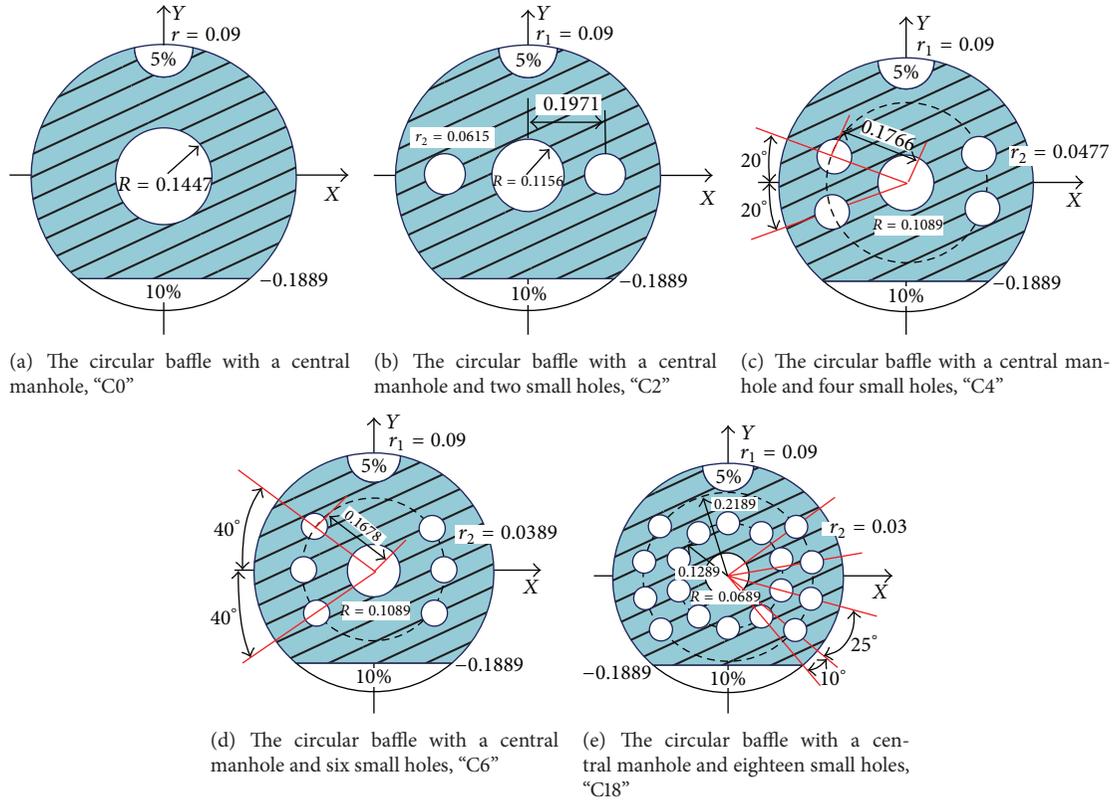


FIGURE 3: A series of circular baffles.

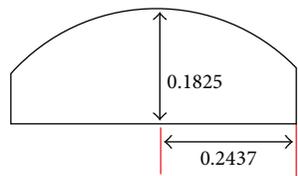


FIGURE 4: A part of a staggered baffle.

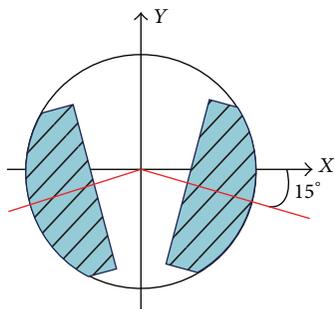


FIGURE 5: The reverse staggered baffle.

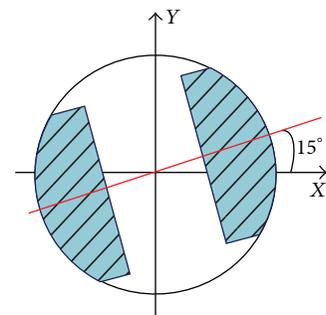


FIGURE 6: The similar staggered baffle.

outside of the tank to ensure that the simulation reflects the actual situation accurately.

Unlike other research, deceleration was applied to the tank continuously instead of being withdrawn at a certain time. Supposing that the initial braking speed is 80 km/h,

while the maximum braking deceleration the road can provide is about 6.7 m/s^2 , the truck needs 3-4 seconds to stop completely. Generally, the simulation time for liquid sloshing is about 2 seconds, which is less than the actual braking time. Therefore, if the simulation time is not being increased, there is no need to withdraw the deceleration during the simulation, and a full-development liquid sloshing process can be monitored and studied.

According to market research, in practice, the liquid fill percentage or fill level is in the range of 0.4–0.8 most of the time. Therefore, the liquid fill level varied from 0.4 to 0.8 with a step size of 0.1 in this paper.

During the simulation, the longitudinal sloshing forces on all of the walls were monitored separately.

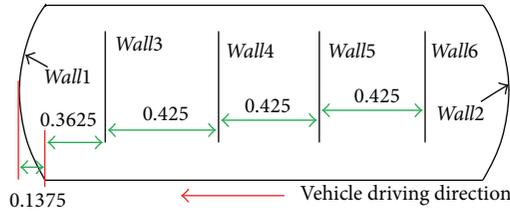


FIGURE 7: The distribution of the conventional and circular baffle.

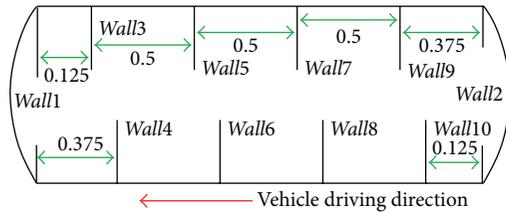


FIGURE 8: The distribution of the staggered baffle.

3. Results and Discussion

Liquid sloshing in tanks without baffles was simulated first and the maximum sloshing force on *wall1* was compared with the result obtained using the quasistatic method to verify the accuracy of the FLUENT simulation settings.

In the quasistatic method, the longitudinal sloshing force on *wall1* is expressed as follows:

$$F_x = ma_x. \quad (20)$$

As the liquid fill level changes from 0.4 to 0.8 with a 0.1 step-size, the fluid volume (fluid in the elliptical heads is included) takes the following values: 0.7975 m^3 , 0.9083 m^3 , 1.0713 m^3 , 1.1987 m^3 , and 1.3616 m^3 .

The results of the comparison were presented in Figure 9. It is known that liquid sloshing forces obtained using the quasistatic method are a little smaller than those obtained by FLUENT simulation. While liquid sloshing obtained using the quasistatic method is very close to the actual, the comparison result reveals the correctness of FLUENT simulation settings.

For each kind of baffle, several forms existed. Therefore, an intergroup comparison of the influence of different baffle forms on the liquid sloshing reduction was carried out first. Then a comparison of the best baffle forms for each kind was made to determine the largest limitation to the liquid sloshing.

3.1. Results of the Conventional Baffle. The longitudinal sloshing forces on *wall1*–*wall6* in tanks equipped with conventional baffles and with the 0.4 liquid fill level of 0.4 are plotted in Figure 10.

In Figures 10(a)–10(b), it is quite obvious that N0, whose oblique angle is 0-degree, has much less of a reducing effect on liquid sloshing than the other baffles. Moreover, the sloshing forces on *wall3*–*wall6* in the tank equipped with N0 are much larger than the forces on the same wall in tanks equipped with

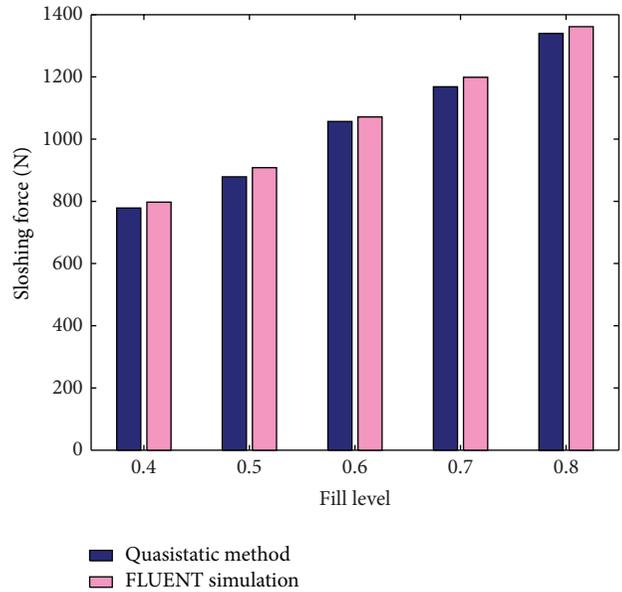


FIGURE 9: Comparison of the longitudinal liquid sloshing forces obtained using FLUENT simulation and the Quasistatic method.

N5–N20, as shown in Figures 10(c)–10(f). Furthermore, N5–N20 have almost the same effect on liquid sloshing reduction. It is known that the larger the sloshing force on the baffle is, the more easily damaged it will be. Therefore, N0 is the worst choice as a transverse baffle when liquid percentage is 0.4.

It can be seen in Figure 10 that the sloshing forces on *wall1* and *wall3* are much bigger than those on the other walls. As the scale ratio between the FLUENT simulation sloshing force and the real-life case is 1 : 64, *wall1* and *wall3* should be paid the most attention. Therefore, for liquid fill levels of 0.5–0.8, only the sloshing forces on *wall1* and *wall3* are presented (see Figure 11). The sloshing force on *wall1* reflects the baffle effect in terms of reducing longitudinal liquid sloshing in partially filled tanks, and the sloshing force on *wall3* is just the maximum sloshing force the baffle bears.

It can be seen in Figures 11(a1)–11(d1) that the sloshing forces on *wall1* differ very little between tanks equipped with conventional baffles of different arrangement angles at the same liquid fill level. The same can be said of the sloshing forces on *wall3*. That is to say that, at the same liquid fill level, conventional baffles of different oblique angles have almost identical effects in terms of reducing liquid sloshing, and the maximum sloshing forces on the baffles are also quite similar.

However, a closer watch reveals that the following conclusions can be drawn. For liquid fill level changing from 0.5 to 0.6, N0 does not reduce liquid sloshing to the same extent as the other baffles, and the force on *wall3* is also larger than those in tanks equipped with the baffles of N5–N20. For the liquid fill level of 0.7, the sloshing forces on *wall1* in tanks equipped with conventional baffles of N0–N20 are quite close to each other, so are the sloshing forces on *wall3*. Finally, for the liquid fill level of 0.8, N10 bears the largest sloshing force and is not as effective as the other baffles in limiting sloshing.

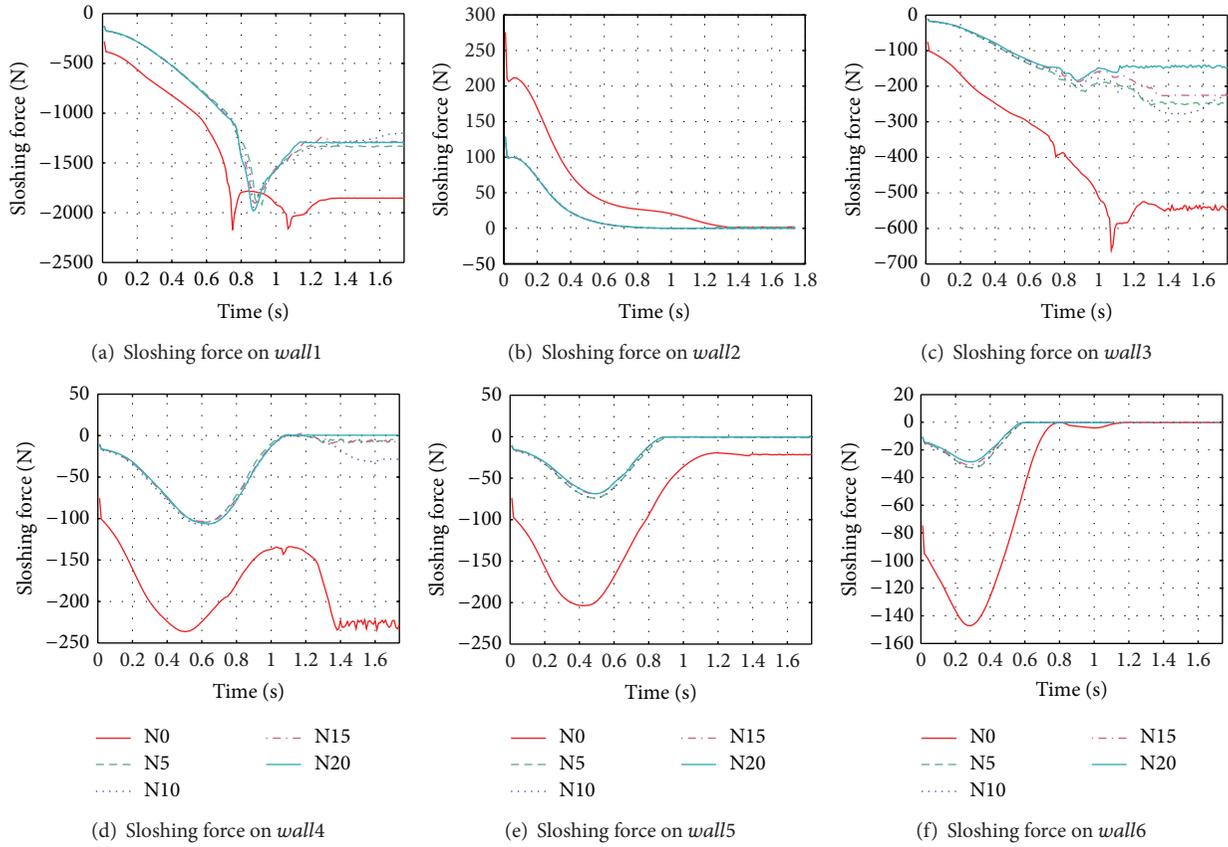


FIGURE 10: Sloshing forces on wall1–wall6 in tanks with conventional baffles and the liquid fill level of 0.4.

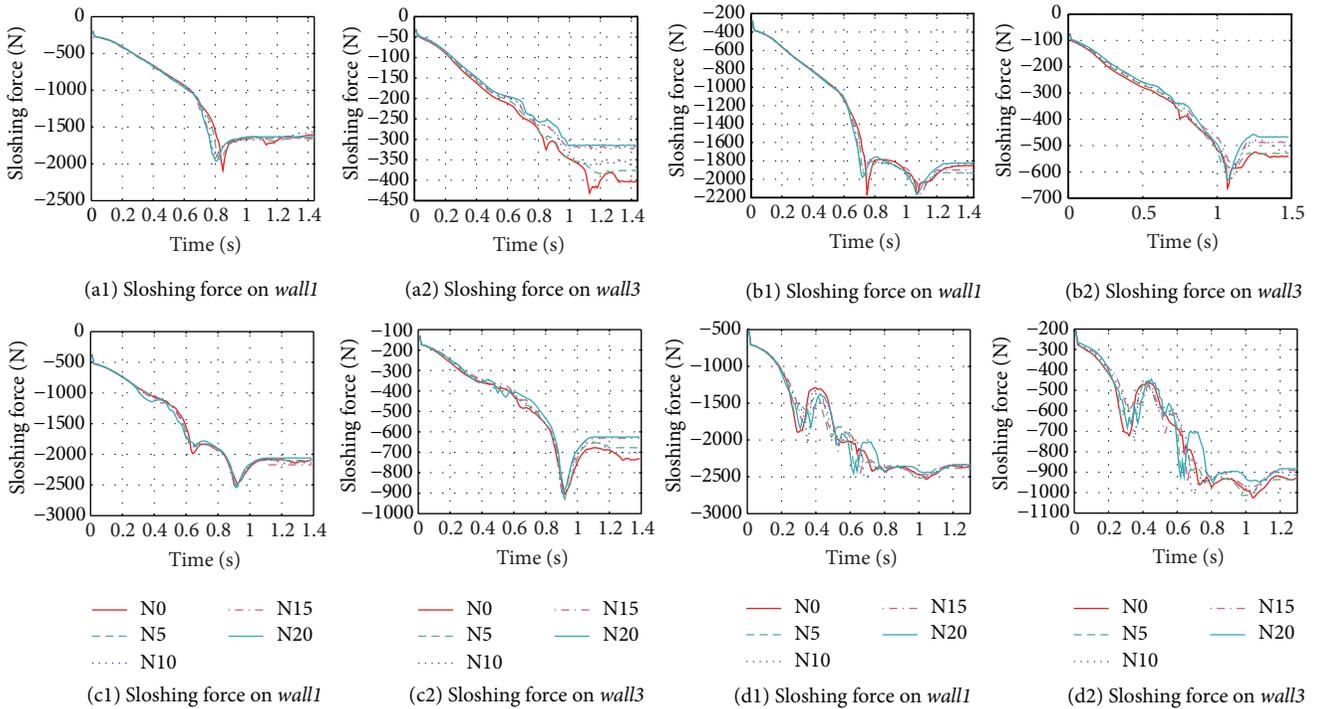


FIGURE 11: Sloshing forces on wall1 and wall3 in tanks with conventional baffles and liquid fill levels of 0.5–0.8: (a) fill level = 0.5; (b) fill level = 0.6; (c) fill level = 0.7; (d) fill level = 0.8.

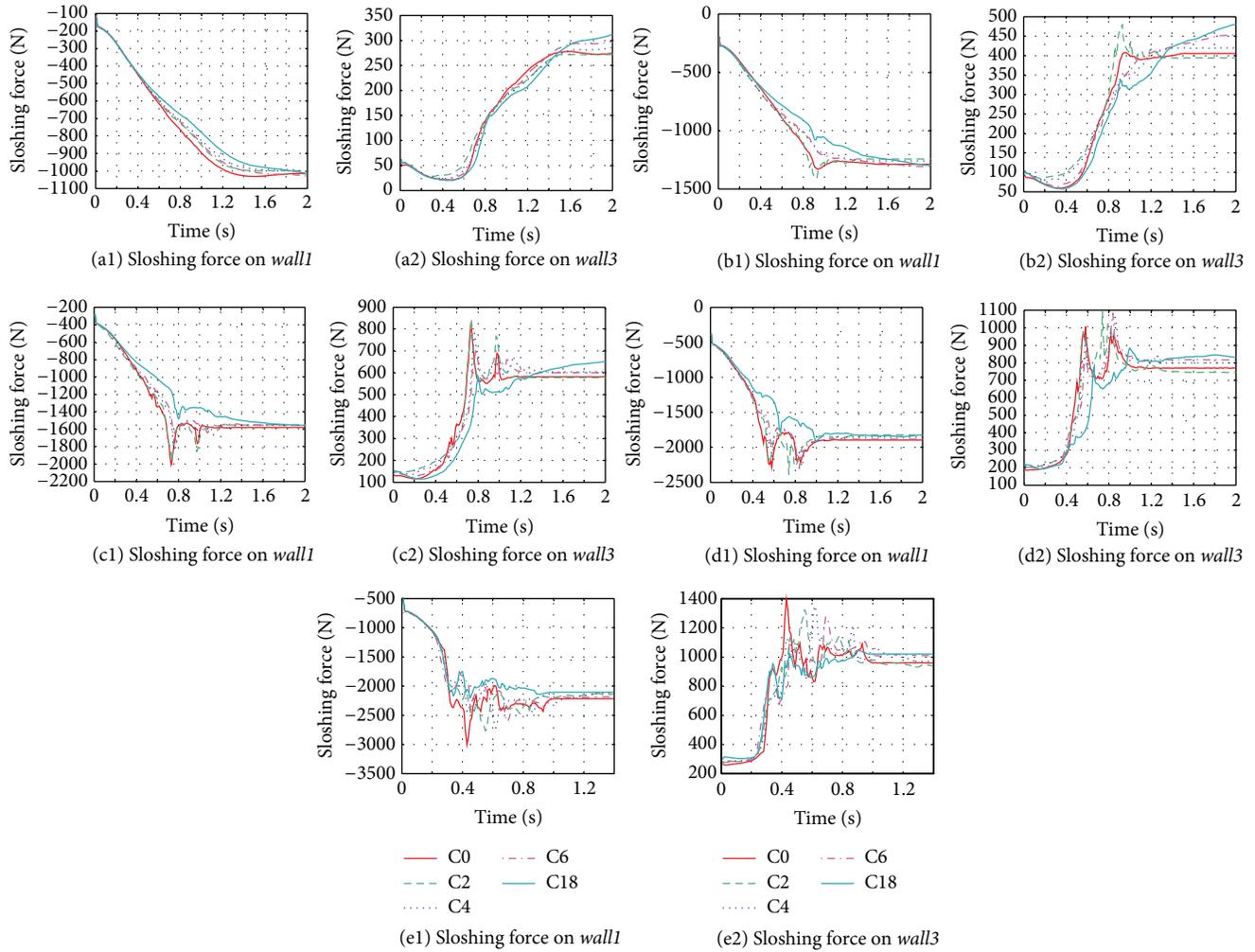


FIGURE 12: Slushing forces on *wall1* and *wall3* in tanks with circular baffles and liquid fill level changing from 0.4 to 0.8: (a) fill level = 0.4; (b) fill level = 0.5; (c) fill level = 0.6; (d) fill level = 0.7; (e) fill level = 0.8.

According to the above analysis, N5, N15, and N20 are the ideal conventional baffles as they limit slushing the most. However, while baffles with different oblique angles have the same projection area in the vertical plane, the larger the oblique angle is, the heavier the baffle will be. That will increase the structural container's mass and reduce the useful load capacity of the tank vehicle. Therefore, N5 is the best conventional baffle.

3.2. Results of the Circular Baffle. The number of circular baffle arranged in the tank model is equal to that of conventional baffle used in the previous section. Similar with the former, the longitudinal slushing forces on *wall2* and *wall4*–*wall6* are much smaller than those on *wall1* and *wall3*. Therefore, only the slushing forces on *wall1* and *wall3* are provided here, again for liquid fill level changing from 0.4 to 0.8. The simulation results were used to evaluate the slushing restraint effects of different circular baffles and the maximum slushing forces the baffles have to bear.

Under liquid slushing in tanks equipped with different circular baffles, the forces on *wall1* and *wall3* change steadily

while the liquid fill level changes from 0.4 to 0.5, as shown in Figure 12. As the liquid fill level increases, the variation in the forces becomes irregular and harsh, and many sharp peaks and valleys appear. These abrupt changes in the forces acting on the tank walls and the baffles could reduce the lifetimes of both. Therefore, such large and abrupt force change is not welcome. The steadier the force change is, the better force analysis the object has, even if the steady value of the force is a little higher.

It can be seen in Figures 12(a1)–12(e1) that, for circular baffles with the same free area, the smaller the radii of the holes (including both the manhole and the small holes around it) are, the more holes the baffle must have, and thus the larger the reduction in slushing will be. For the tank equipped with C18, the slushing force on *wall1* is the smallest, and the force variation on this baffle is much smoother. Therefore, C18, the baffle with a central manhole and eighteen small holes around it has the best effect in terms of reducing liquid slushing.

The slushing forces on the baffles tend to steady values with an increase in simulation time. However, the steady slushing forces on *wall1* and *wall3* in the tank equipped with

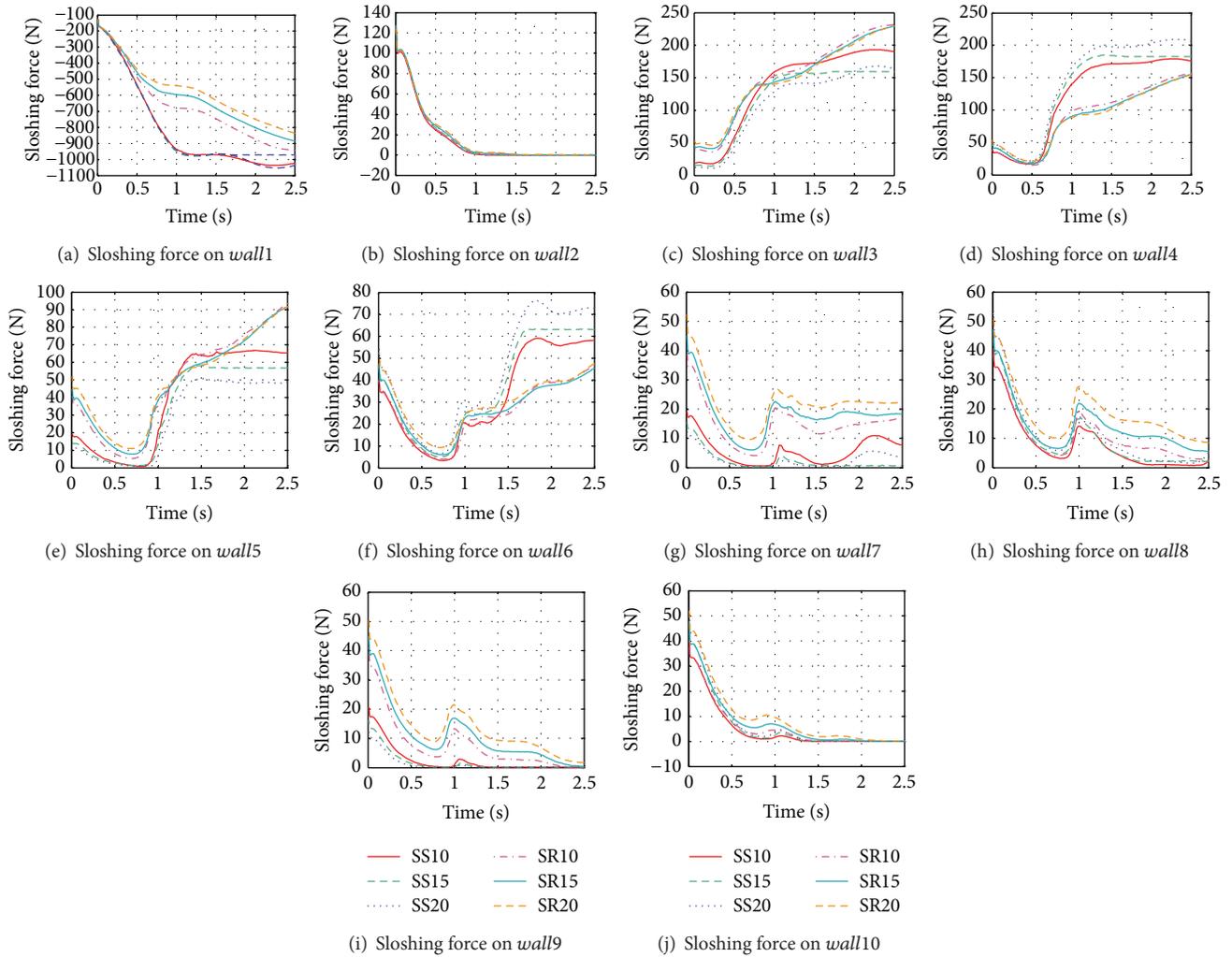


FIGURE 13: Slashing forces on wall1–wall10 in tank equipped with staggered baffles and with the liquid fill level of 0.4.

C18 are not the smallest when the liquid sloshing develops fully. As mentioned earlier, while we pay much more attention to the steady and gentle change of sloshing forces, their absolute values are less of importance. Therefore, C18, the baffle with a central manhole and 18 small holes around the manhole, is the best circular baffle.

We can also discover that circular baffle with more holes has a better antisloshing effect. According to the force distribution of the baffle, sloshing forces near the holes are much bigger. By Newton’s third law, liquid sloshing force is greatly degraded when liquid flows through the holes. Moreover, the small holes make much bigger contribution to the antisloshing effect. Therefore, if we could get a baffle with many small holes pierced on it, the baffle must have an excellent behavior on reducing liquid sloshing. Limited by mesh-grid size and simulation time, simulation for liquid sloshing in a tank equipped with such baffles is not conducted.

3.3. Results of the Staggered Baffle. Unlike the conventional and circular baffles, a complete staggered baffle is composed

of two parts, as shown in Figures 5 and 6. They are arranged in a staggered fashion so as to restrain fluid sloshing in partially filled tank vehicles. Therefore, there are four baffle pieces fitted on the left side of the tank and four on the right, as shown in Figure 8.

In tanks equipped with different staggered baffles and with the liquid fill percentage of 0.4, the longitudinal sloshing forces on wall1–wall10 are plotted in Figure 13. It can again be seen that the sloshing forces on wall1 and wall3 are much bigger than those on the other walls. As in the cases of the conventional and circular baffles, wall1 and wall3 bear the biggest sloshing forces and are the most easily damaged parts in the tank.

In Figure 13(a) we can see that SR20 has the best effect in terms of reducing liquid sloshing. However, as can be seen in Figures 13(c), 13(e), and 13(g)–13(j), the sloshing forces on wall3, wall5, and wall7–wall10 in the tank with SR20 are the largest. In this situation, the sloshing restraint effect is chosen as the criterion for evaluating the best staggered baffle. This is due to the fact that sloshing force on wall1 is much bigger than that on wall3.

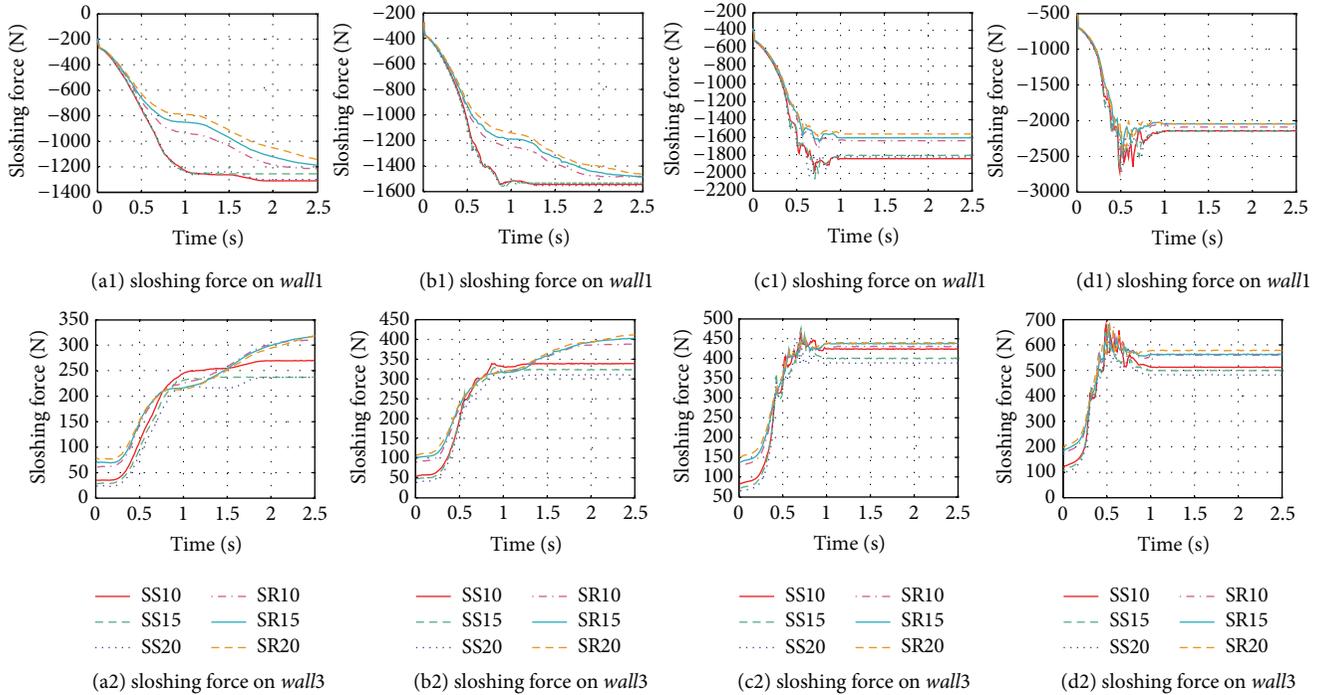


FIGURE 14: Slushing forces on *wall1* and *wall3* in tanks equipped with staggered baffles and with liquid fill levels of 0.5–0.8: (a) fill level = 0.5; (b) fill level = 0.6; (c) fill level = 0.7; (d) fill level = 0.8.

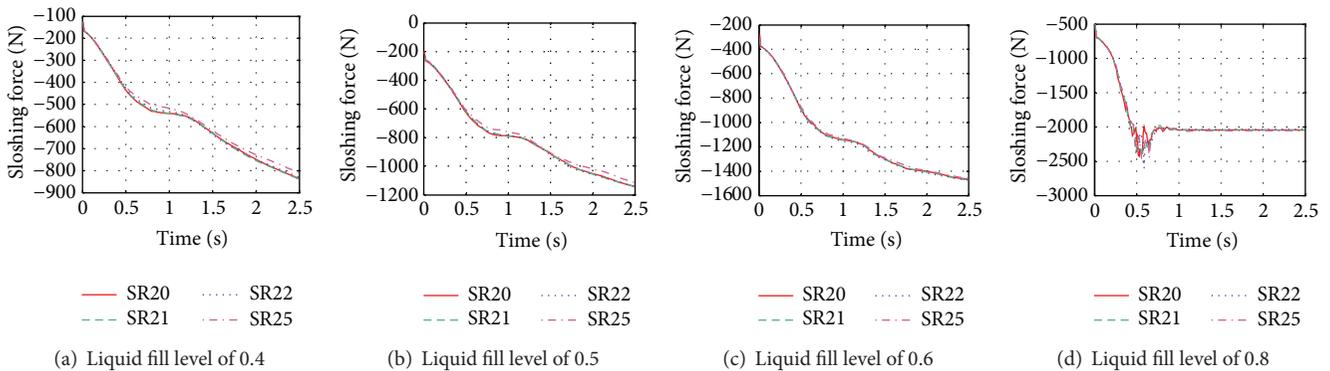


FIGURE 15: Slushing forces on *wall1* in tanks equipped with reverse staggered baffles and with liquid fill levels of 0.4–0.8.

The slushing forces on *wall1* and *wall3* in tanks with staggered baffles while the liquid fill level changes from 0.5 to 0.8 are presented in Figure 14. It is very apparent that SR20 has the best effect on reducing the liquid slushing, and the SR20 baffles bear the largest slushing forces at the same time. As the fill level increases, the slushing forces on the walls become irregular and harsh, and force peaks and valleys appear. The variations in the slushing forces generated in the tank equipped with the SR20 baffle are much smoother than those generated in tanks equipped with the other staggered baffles. Therefore, the SR20 baffle can be considered as the best staggered baffle.

It can be drawn in Figures 14(a1)–14(d1) that, for the reverse staggered baffle, there is a direct relationship between the arrangement angle and the baffle’s antisloshing effect.

With the increase of arrangement angle, the antisloshing effect of the baffle becomes better. Consequently, a few more reverse staggered baffles whose arrangement angles are 21-degree, 22-degree, and 25-degree were designed to investigate whether SR20 is indeed the best staggered baffle, and the complementary reverse staggered baffles are labeled SR21, SR22, and SR25. An arrangement angle that is bigger than 25-degree was not considered to ensure that the fluid could still drain away quickly and easily.

Therefore, for tanks equipped with reverse staggered baffles whose arrangement angle is 20-, 21-, 22- and 25-degree, the slushing forces on *wall1* while liquid fill level changing from 0.4 to 0.8 are plotted in Figure 15. As the curves on *wall1* were quite similar to fill levels of 0.7 and 0.6, only the curves for the liquid fill level of 0.6 are shown.

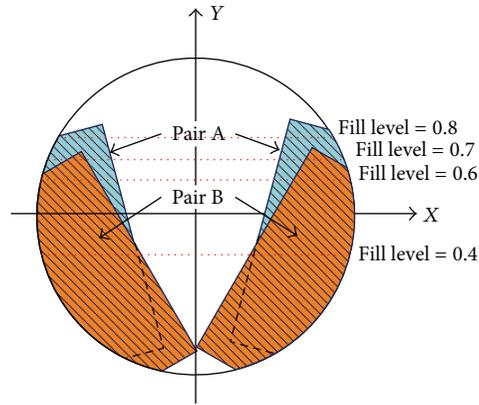


FIGURE 16: Schematic diagram for the antisloshing effect of staggered baffles.

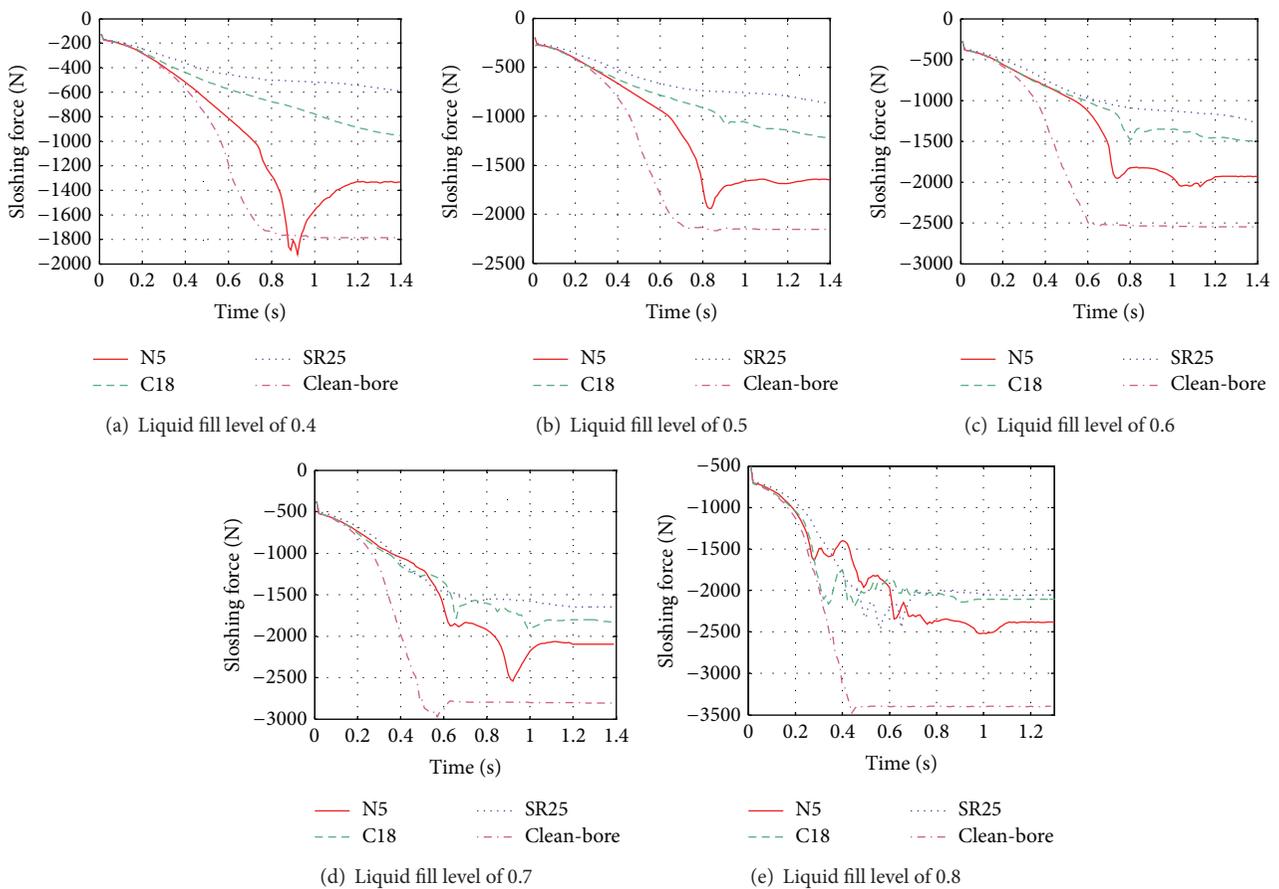


FIGURE 17: Sloshing forces on *wall1* in tanks equipped with different kinds of baffles and with liquid fill levels of 0.4–0.8.

The increase in the arrangement angle strengthens the reduction effect of the reverse staggered baffle on liquid sloshing. However, for sloshing forces generated in tanks with the reverse staggered baffles of 20- and 25-degree arrangement angle, with the increase of the liquid fill level, the difference in these two cases becomes smaller. Furthermore, the change in the baffle’s sloshing restraint effect when the baffle arrangement angle change from 20- to 25-degree is smaller than that when the baffle arrangement angle changes from 5- to

20-degree (under the same conditions). While it is not suggested that the arrangement angle should exceed 25-degree, it can be deduced that the reverse staggered baffle with an arrangement angle of 25-degree is the best staggered baffle.

An attempt was made to discover why the bigger the arrangement is, the better antisloshing effect the staggered baffle has. With the increase in the arrangement angle, for liquid fill level changes from 0.4 to 0.8, the area of staggered baffles used for reducing liquid sloshing grows, as shown

in Figure 16. While the external acceleration or deceleration keeps constant, the bigger the antisloshing area is, the better antisloshing effect the baffle has.

3.4. Effect of Different Kinds of Baffles on the Reduction of Liquid Sloshing. Now, the three baffles with the best liquid sloshing restraint effects, one of the conventional baffle, one of the circular baffle, and one of the staggered baffle, are compared to each other so as to select the most effective one. The comparison curves are plotted in Figure 17.

It is obvious that the reverse staggered baffle with an arrangement angle of 25-degree has much more better effect on reducing liquid sloshing than the other baffles when the liquid fill level is in the range of 0.4–0.7. When the liquid fill level is higher than 0.7, C18, the circular baffle with a central manhole and eighteen small holes gives an identical performance to the latter. The conventional baffle performs quite poorly when the liquid fill level is smaller than 0.7. Therefore, across the whole range of liquid fill levels, the reverse staggered baffle with an arrangement angle of 25-degree would be the first choice for reducing liquid sloshing in a tank with a circular cross-section.

4. Conclusions

To reduce longitudinal liquid sloshing in a partially filled tank, the conventional, circular, and staggered baffles were designed and several forms in each kind were used to investigate the impact of baffle installation angle, the sizes of holes in the baffle, and their arrangement on the antisloshing effect. A full-sized tank was scaled down to a tank model according to the theory of similarity to save on simulation time. The FLU-ENT software was used to simulate liquid sloshing in tanks with different fill percentages, equipped with kinds of baffles and subject to constant braking deceleration. A time-series analysis on forces that act on tank walls and transverse baffles was carried out to obtain the viable optimal baffle designed in this paper.

It was found that the fore elliptical head and the baffle closest to the fore head are the most easily damaged parts in a partially filled tank equipped with baffles. Therefore, sloshing forces on *wall1* and *wall3* must be examined to ensure the structural integrity of both the tank and the baffle. By a time-series analysis of the sloshing forces, it was found that for the conventional baffle, baffle equipped with a 5-degree oblique angle is the best one on reducing liquid sloshing, and the sloshing force on this baffle is the smallest. For the circular baffle, baffle has a central manhole and 18 small holes around it is the most effective one on reducing liquid sloshing, and the changing of forces on this baffle is much smoother. For staggered baffles, the reverse staggered baffle with an arrangement angle of 25-degree has the best antisloshing effect; however, the force on this baffle is the biggest. In the end, by comparison of the three baffles that are chosen from each kind of baffle, the staggered baffle is the optimal one on reducing liquid sloshing.

During the analysis procedure, it was discovered that the baffle area used for antisloshing has a direct relation with

baffle's liquid sloshing reduction effect. The bigger the area used for antisloshing is, the better antisloshing effect the baffle has. Besides, the distance between adjacent baffles can make great contribution to the development of liquid sloshing; hence, a smaller distance is suggested. Combined with the limitation on the structural container's mass, the reverse staggered baffle might be the best transverse baffle on reducing liquid sloshing.

Since we just illustrate liquid sloshing in tanks with different transverse baffles, the function mechanism of transverse baffle on antisloshing as a further study should be conducted.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research is supported by Research on Evaluation and Detection Technology of Handling and Driving Stability for Commercial Vehicles (no. 2009BAG13A04), the National Natural Science Foundation of China (Grant no. 51375200), and the National Natural Science Foundation of China (Grant no. 51208225).

References

- [1] Research and Innovative Technology/Bureau of Transportation Statistics, "Transportation-2007 commodity flow survey," 2010.
- [2] J. Woodrooffe, "Evaluation of dangerous goods vehicle safety performance," Transport Canada Report, TP 13678 E, 2000.
- [3] C. Kwon, "Conditional value-at-risk model for hazardous materials transportation," in *Proceeding of the Winter Simulation Conference (WSC '11)*, pp. 1703–1709, Phoenix, Ariz, USA, December 2011.
- [4] L. E. Ghosh, S. M. Rapik, and P. L. B. Christopher, "Release risk model to evaluate highway cargo tank component-specific safety performance in transporting hazardous materials," in *Proceeding of the 91st Annual Meeting Transportation Research Board (TRP '12)*, Washington, DC, USA, 2012.
- [5] K. Z. Liu, H. X. Liu, Y. C. Yan et al., "Analysis of the road accidents with the tank-vehicle transportation of hazardous materials," *Journal of Safety and Environment*, vol. 21, no. 2, pp. 18–21, 2012.
- [6] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research Part C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [7] M. I. Salem, H. Victor, M. Mucino et al., "Review of parameters affecting stability of partially filled heavy-duty tankers," in *Proceeding of the SAE International, Truck & Bus Meet Exposition*, 1999-01-3709, Detroit, Mich, USA, November 1999.
- [8] X. Kang, S. Rakheja, and I. Stiharu, "Cargo load shift and its influence on tank vehicle dynamics under braking and turning," *The International Journal of Heavy Vehicle Systems*, vol. 9, no. 3, pp. 173–203, 2002.
- [9] F. Solaas and O. M. Faltinsen, "Combined numerical and analytical solution for sloshing in two-dimensional tanks of general

- shape,” *Journal of Ship Research*, vol. 41, no. 2, pp. 118–129, 1997.
- [10] K. Modaressi-Tehrani, S. Rakheja, and I. Stiharu, “Three-dimensional analysis of transient slosh within a partly-filled tank equipped with baffles,” *Vehicle System Dynamics*, vol. 45, no. 6, pp. 525–548, 2007.
- [11] T. Kandansamy, S. Rakheja, and A. K. W. Ahmed, “An analysis of baffles designs for limiting fluid slosh in partly filled tank trucks,” *The Open Transportation Journal*, vol. 4, pp. 23–32, 2010.
- [12] D. Liu and P. Lin, “Three-dimensional liquid sloshing in a tank with baffles,” *Ocean Engineering*, vol. 36, no. 2, pp. 202–212, 2009.
- [13] H. Akyildiz, “A numerical study of the effects of the vertical baffle on liquid sloshing in two-dimensional rectangular tank,” *Journal of Sound and Vibration*, vol. 331, no. 1, pp. 41–52, 2012.
- [14] J. R. Cho and H. W. Lee, “Numerical study on liquid sloshing in baffled tank by nonlinear finite element method,” *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 23–26, pp. 2581–2598, 2004.
- [15] N. Kang and K. Liu, “Influence of baffle position on liquid sloshing during braking and turning of a tank truck,” *Journal of Zhejiang University: Science A*, vol. 11, no. 5, pp. 317–324, 2010.
- [16] A. Dasgupta, *Effect of tank cross-section and longitudinal baffles on transient liquid slosh in partly-filled road tankers [M.S. thesis]*, Department of Mechanical and Industrial Engineering, Concordia University, 2009.
- [17] N. Lloyd, E. Vaiciurgis, and T. A. G. Langrish, “The effect of baffle design on longitudinal liquid movement in road tankers an experimental investigation,” *Transactions of the Institution of Chemical Engineers, Part B*, vol. 80, no. 4, pp. 181–185, 2002.
- [18] M. A. Xue, J. H. Zheng, and P. Z. Lin, “Numerical simulation of sloshing phenomena in cubic tank with multiple baffles,” *Journal of Applied Mathematics*, vol. 2012, Article ID 245702, 21 pages, 2012.
- [19] K. Threepopnartkul and C. Suvanjumrat, “effect of baffles on fluid sloshing inside the moving rectangular tanks,” in *Proceeding of the 2nd TSME International Conference on Mechanical Engineering*, Krabi, Thailand, October 2011.
- [20] K. Modaressi-Tehrani, S. Rakheja, and I. Stiharu, “Role of transverse baffle designs on transient three-dimensional liquid slosh in a partly-filled circular tank,” in *Proceeding of the SAE Commercial Vehicle Engineering Congress and Exhibition*, 2005-01-3594, Chicago, Ill, USA, November 2005.
- [21] M. J. Fabela-Gallegos, A. Ramires-Valencia, C. A. Favala-Gallegos et al., “Experimental assessment of baffles and their effect on the longitudinal sloshing force in a scaled elliptical tank,” in *Proceeding of the SAE Commercial Vehicle Engineering Congress and Exhibition*, 2005-01-3577, Chicago, Ill, USA, November 2005.
- [22] K. Modaressi-Tehrani, S. Rakheja, and I. Stiharu, “Three-dimensional analysis of transient slosh within a partly-filled tank equipped with baffles,” *Vehicle System Dynamics*, vol. 45, no. 6, pp. 525–548, 2007.
- [23] M. J. Fabela-Gallegos, D. Vazquez-Vega, M. Martinez-Madrid et al., “Experimental evaluation of baffles geometry and their configuration in partially filled scaled elliptical tank,” in *Proceeding of the SAE Commercial Vehicle Engineering Congress and Exhibition*, 2005-01-3581, Chicago, Ill, USA, November 2005.
- [24] G. R. Yan, S. Rakheja, and K. Siddiqui, “Baffle design analysis for a road tanker transient fluid slosh approach,” *SAE International Journal of Commercial Vehicles*, 2008-01-2670, 2008.
- [25] I. M. Ibrahim, “Anti-slosh damper design for improving the roll dynamic behavior of cylinder tank trucks,” in *Proceeding of the SAE International Truck and Bus Meeting and Exposition*, 1999-01-3729, Detroit, Mich, USA, November 1999.

Research Article

An Adaptive Model for Calculating the Correlation Degree of Multiple Adjacent Signalized Intersections

Linhong Wang¹ and Yiming Bie²

¹ College of Transportation, Jilin University, Changchun 130022, China

² School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150091, China

Correspondence should be addressed to Yiming Bie; yimingbie@126.com

Received 5 July 2013; Revised 24 August 2013; Accepted 4 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 L. Wang and Y. Bie. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an important component of the urban adaptive traffic control system, subarea partition algorithm divides the road network into some small subareas and then determines the optimal signal control mode for each signalized intersection. Correlation model is the core of subarea partition algorithm because it can quantify the correlation degree of adjacent signalized intersections and decides whether these intersections can be grouped into one subarea. In most cases, there are more than two intersections in one subarea. However, current researches only focus on the correlation model for two adjacent intersections. The objective of this study is to develop a model which can calculate the correlation degree of multiple intersections adaptively. The cycle lengths, link lengths, number of intersections, and path flow between upstream and downstream coordinated phases were selected as the contributing factors of the correlation model. Their jointly impacts on the performance of the coordinated control mode relative to the isolated control mode were further studied using numerical experiments. The paper then proposed a correlation index (CI) as an alternative to relative performance. The relationship between CI and the four contributing factors was established in order to predict the correlation, which determined whether adjacent intersections could be partitioned into one subarea. A value of 0 was set as the threshold of CI. If CI was larger than 0, multiple intersections could be partitioned into one subarea; otherwise, they should be separated. Finally, case studies were conducted in a real-life signalized network to evaluate the performance of the model. The results show that the CI simulates the relative performance well and could be a reliable index for subarea partition.

1. Introduction

When the network-wide traffic signal control strategy is implemented to an urban road network, subarea partition must be conducted to divide the network into some small subareas. The subarea can be classified into three types. The first type only includes one signalized intersection. The second type includes several intersections located on one artery. The third type includes several intersections located on some intersected arteries. For the three type subareas, isolated signal control mode, arterial signal coordination mode, and area signal coordination mode should be implemented to the intersection(s), respectively. Accordingly, subarea partition is the basis for network-wide signal control. A well subarea partition algorithm can improve the adaptive level of the signal control system significantly [1–3].

In fact, the most important part of subarea partition is calculating the correlation degree among adjacent signalized

intersections. The purpose of subarea partition is to improve the whole control performance of the network. If the performance could be improved, then adjacent intersections can be partitioned into one subarea. Otherwise they should be partitioned into different subareas. Thus, a straightforward method for subarea partition is to calculate and compare the control performances under the coordinated control mode and the isolated control mode. However, this method is not easy to implement in practice although it is theoretically straightforward. This is because it would be time consuming to iteratively estimate the performances of hundreds of adjacent intersections in the network. Thus, an effective alternative is to propose an index that reflects the changes in control performance indirectly; this is called the correlation index (CI) in this paper.

Because of the importance of subarea partition, the related researches have been conducted from the 1970s [4–11]. Yagoda et al. defined a coupling index (I) as $I = V/L$.

A higher value of the index would indicate that it would be preferable to have the two intersections coupled. V is the traffic volume and L is the link length between two adjacent intersections [12]. Chang brought forward the interconnection desirability index by considering both flow pattern and platoon dispersion characteristic. If the value of the index was greater than 0.35, the two intersections should be grouped together [13]. Moore and Jovanis [14] thought that adjacent intersections can be partitioned together when their v/c ratios were close. However, the threshold of v/c ratios was not determined. Lin and Huang [15] established a simple model to identify whether two adjacent signalized intersections can be coordinated based on the block length. The block length is affected by original platoon size and platoon completeness ratio. However, the block length cannot depict the impacts of all contributing factors on subarea partition. Lu et al. [16] selected link length, link volume, and cycle length as contributing factors and then developed two models to describe the correlation between two adjacent intersections and between multiple adjacent intersections, respectively, according to expert knowledge. Bie et al. [17] chose not only cycle length but also platoon length as contributing factors. An integrated correlation model was developed to quantify the correlation between two adjacent intersections. However, the determination of the weighted coefficients in the model and the threshold value were dependent on past experience and lacked a theoretical basis.

As described above, some valuable results have been achieved on the subject of correlation model development. However, there are still some deficiencies that needed to be resolved, which are summarized as follows.

- (1) The contributing factors to the correlation index and the model structure are mainly determined according to the past experience of traffic engineers. Most weighting factors and parameters are fixed and cannot vary as the traffic state changes. Thus, the developed models cannot quantify the correlation between adjacent intersections accurately.
- (2) The threshold value for the correlation has not been theoretically justified. The threshold value directly determines whether adjacent intersections can be partitioned into one subarea. However, in existing studies, the threshold value is again determined based on the experience of engineers. Accordingly, the subarea partition results may not be optimal.
- (3) The existing studies have not yet developed a model that can measure the correlation among multiple intersections. In most cases, one subarea includes more than two intersections. Thus, it is inevitable to calculate the correlation among multiple intersections. However, existing studies only calculate the correlation between two adjacent intersections.

Targeting the real-time traffic control systems, an appropriate correlation index should reflect the dynamic nature of traffic flow and interactions among adjacent intersections. In this paper, we propose an approach that can estimate the dynamic correlation among multiple adjacent intersections

for the purpose of subarea partition. The key innovation of this study is that the contributing factors are expressed as variables and the relation model between correlation and the variables is developed. The correlation can be calculated dynamically as the traffic state changes. Moreover, our modeling approach provides a reasonable way to determine the threshold value of correlation at which adjacent intersections should be grouped into one subarea.

2. Contributing Factors for Subarea Partition

The determination of contributing factors for subarea partition is the basis for developing the correlation model. Correlation model is the core of subarea partition, thus the principles of subarea partition must be obeyed when developing the model. If all intersections in the urban network execute isolated signal control mode, it is unnecessary to develop the correlation model because isolated control mode only concerns the intersection itself. However, when we identify whether multiple adjacent intersections can be partitioned together to execute signal coordination mode, the factors that may affect signal coordination should be analyzed. Thus, the factors that can improve signal coordination performance are also those that affect subarea partition. In this paper, the following four factors are selected.

(1) *Difference between Cycle Lengths of Adjacent Intersections.* When intersections in a subarea execute signal coordination mode, they should run the common cycle length (CCL). The CCL can help the signal controller to achieve maximum green wave bandwidth and maintain optimal offsets. Usually CCL equals the maximum cycle length (MCL) of all optimal cycle lengths (OCL) of the intersections. OCL is the cycle length of one intersection when it executes isolated control mode. However, the implement of CCL would increase the vehicle delay, because the CCL is larger than the OCLs. The increased delay is directly proportional to the difference between the CCL and the OCL. If the difference is large enough, the increased vehicle delay may be larger than the reduction in delay achieved by the signal coordination.

(2) *Link Length between Two Adjacent Intersections.* Signal coordination can reduce vehicle delay by providing green time for the approaching traffic platoon. A compact platoon is proven to be effective in improving coordination benefit because more vehicles can pass through the stop line during green time. According to Robertson's platoon dispersion model [18], the original platoon that departs from the upstream stop line would disperse to some extent and the dispersion level is directly proportional to the link length. Accordingly, a long link would pay a negative impact on the arterial progression.

(3) *Path Flow between Upstream and Downstream Coordinated Phases.* When a platoon travels to the downstream intersection, some vehicles in the platoon may turn to uncoordinated downstream phases. Therefore, the path flow between the two coordinated phases would be less than the original flow that departs from the upstream coordinated

phase. A greater path flow indicates that more vehicles are traveling between the two phases and there is greater opportunity to reduce vehicle delays through signal coordination. Thus, the coordination benefit is directly proportional to the path flow.

(4) *Number of Intersections.* According to the field experiences of engineers, the control performance of signal coordination would decrease with the increase in the number of intersections in the subarea. In fact, the reason for performance decrease is due to the cumulative negative impacts incurred by the other three contributing factors. In reality, there may be large differences between the cycle lengths of adjacent intersections, long link length, and low path flow. Therefore, from this perspective, the number of intersections is not a direct contributing factor to subarea partition, but an indirect one caused by the other three factors.

However, if we assume there is no difference between the cycle lengths of adjacent intersections, the link lengths are exactly right, and the path flows are great enough, then these three factors would not have negative impact on signal coordination performance. In such situation, the larger the number of intersections in the subareas, the better the coordination performance would be. Therefore, from this perspective, the number of intersections would be a direct contributing factor to subarea partition.

3. Modeling the Correlation

3.1. *Performance Index.* Let us take two adjacent signalized intersections i and j as an example. Their total vehicle delays per cycle are denoted by D_i and D_j when they are running the isolated control mode. D_i and D_j can be obtained using the delay function from HCM 2000 [19]. If the two intersections are grouped into one subarea, they will run signal coordinated mode. In this case, the total vehicle delays per cycle are denoted by \widehat{D}_i and \widehat{D}_j . Therefore, we propose an index PI, and it indicates the improved performance achieved by grouping i and j into one subarea and can be calculated using

$$\begin{aligned} \text{PI} &= D - \widehat{D} \\ &= \left(D_i \frac{3600}{C_i} + D_j \frac{3600}{C_j} \right) - \left(\widehat{D}_i + \widehat{D}_j \right) \frac{3600}{\max(C_i, C_j)}. \end{aligned} \quad (1)$$

In (1), C_i is the OCL of i when it is run in the isolated control mode. $\max(C_i, C_j)$ is the CCL when the signal coordinated mode is run, D is the total vehicle delay per hour under the isolated control mode, and \widehat{D} is the total vehicle delay per hour under the coordinated mode. All delays are unified into the total vehicle delay per hour (the unit of PI is s/h). Because the OCL of a nonseed intersection may be smaller than the CCL, the number of arriving vehicles in each cycle may also be smaller than that when the signal coordinated mode is operated.

If PI is larger than 0, then the signal coordinated mode achieves better control performance than the isolated control

mode; accordingly i and j can be grouped into one subarea. Otherwise, they should be assigned to different subareas.

The performance of the signal coordination is also affected by the coordination type (i.e., one-directional or two-directional coordination); different coordination types will produce different partition results. In this paper, we only study the relations between the correlation index and the contributing factors under two-directional coordination conditions, which is useful to judge whether adjacent intersections can be partitioned into one subarea to execute two-directional coordination algorithm.

The best way to analyze the impact of a contributing factor on PI is to determine the relationship between them based on theoretical derivation and then calculate the change in PI resulting from a unit change in the contributing factor. However, the relationship is difficult to determine due to the complex nature of vehicle movements between consecutive intersections. There is no universal function that can be used to calculate \widehat{D} . Thus, numerical experiments are used for this paper. Experimental data are obtained relating the PI to the contributing factors, and then relation models between CI and the contributing factors are fitted in Sections 3.2 to 3.5.

An important issue that should be noted is that the traffic control objective is closely related to the traffic state of the intersection. For example, the control objective for an oversaturated intersection is usually set as minimizing the queue length to avoid queue spillovers, while that of an unsaturated intersection is usually set as minimizing vehicle delay. In this paper, we focus on unsaturated intersections and minimizing vehicle delay is selected as the control objective. Thus, in (1) PI indicates the reduced vehicle delay when using the signal coordinated mode instead of the isolated control mode.

Differences in cycle lengths only affect the control performances of the seed intersection and nonseed intersection. The link length and path flow only affect the performances of two adjacent intersections. Therefore, a network composing of two intersections is selected to conduct numerical experiments on these three factors.

3.2. *Impact of the Difference in Cycle Lengths on Correlation.* Again taking intersections i and j as an example, let the link length L equal 400 m and let the average vehicle speed be 12 m/s. j is the seed intersection and i is the nonseed intersection. Thus, when run in coordinated mode, the OCL of j is set as the CCL. The difference between the OCL of i and the CCL is increased using a step of 5 s, while the link length and path flow are fixed. Table 1 shows the results for two experimental scenarios, where the CCL equals 125 s and 110 s, respectively. In Table 1, the improvement is shown by the ratio of PI to D . The larger is the ratio, the better are the signal coordination benefits.

For example, in scenario 1 when the OCL of i equals 90 s, the difference between the cycle lengths is 35 s and PI equals 20,044 s. The improvement ratio is only 5.7%. However, when the OCL of i increases to 125 s, the improvement ratio is 17.24%. The difference between cycle lengths is negatively correlated to the improvement ratio and a similar trend

TABLE I: Results of the experimental scenarios.

Scenario	CCL/s	OCL of i/s	$D/s \cdot h^{-1}$	$PI/s \cdot h^{-1}$	Improvement ratio
1	125	90	351576	20044	5.70%
		95	358915	25438	7.09%
		100	371277	34919	9.41%
		105	378642	40453	10.68%
		110	390619	50311	12.88%
		115	399471	57848	14.48%
		120	408876	65726	16.07%
		125	416925	71878	17.24%
2	110	90	322504	22407	6.95%
		95	331007	28399	8.58%
		100	342658	37425	10.92%
		105	351739	44573	12.67%
		110	360691	51658	14.32%

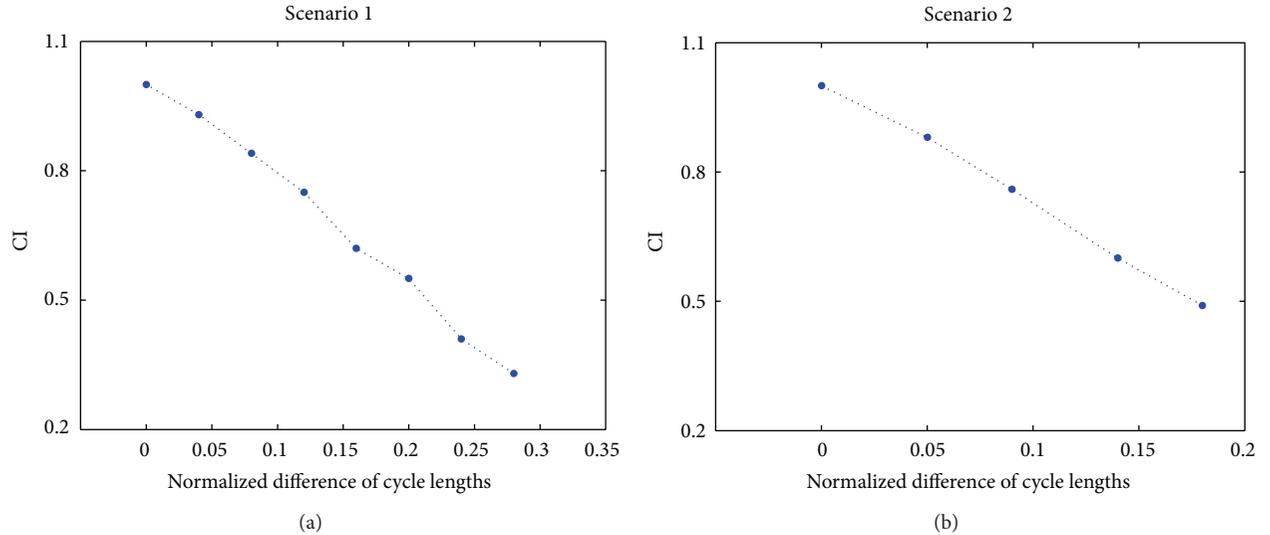


FIGURE 1: Scatter diagrams between CI and the normalized difference in cycle lengths.

is found in other scenarios. The results indicate that the difference between cycle lengths has a substantial impact on PI.

The data shown in Table 1 can also be interpreted as follows. When the difference equals 0, signal coordination achieves the maximum benefits, which means that the correlation between the two intersections is maximal and the correlation index (CI) is 1.0. However, when the difference equals 35 s, the ratio of the improvement to that when the difference equals 0 is only 33%; that is, CI decreases to 0.33.

As can be seen from Table 1, CI is affected not only by the difference in cycle lengths, but also by the CCL. Thus, the difference in cycle lengths is normalized as follows:

$$C_D = \frac{C_j - C_i}{C_j}. \quad (2)$$

Scatter diagrams between CI and C_D are shown for the two scenarios in Figure 1.

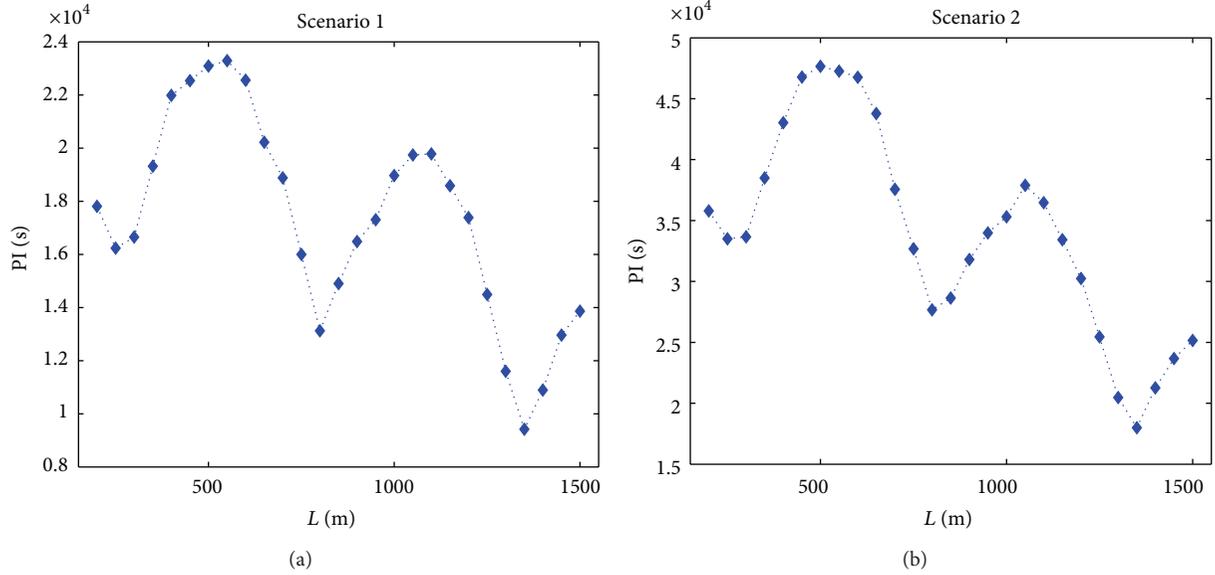
There is a strong linear relationship between CI and C_D , and a linear function can be used to model their relationship. The linear functions for the two scenarios are as follows:

$$CI = \begin{cases} -2.471C_D + 1.024, & R^2 = 0.995, \\ -2.894C_D + 1.009, & R^2 = 0.996. \end{cases} \quad (3)$$

Linear functions can fit the scatter diagrams very well, as indicated by the high R -square values. However, the problem is that the slopes of the two lines are different from each other. This indicates that, in different scenarios, C_D will have a different impact on CI. Therefore, a universal function is needed that depicts the relationship between the two variables, in which the slope varies as the traffic state changes. The linear functions in (3) can be replaced by the following format:

$$CI = \alpha_1 C_D + b_1, \quad (4)$$

where α_1 is the slope of the function, and b_1 is the intercept.


 FIGURE 2: Scatter diagrams between L and PI for the two scenarios.

Theoretically, b_1 should equal 1. However, due to fitting errors in (3), marginal differences exist between the intercepts and 1. In this study, however, the differences are neglected and the value of b_1 is simply set to 1. Therefore, only α_1 needs to be fitted.

α_1 represents the change in CI resulting from a unit change in C_D . It may be affected by many factors, such as the green split and the volume to capacity ratio (v/c) of coordinated phases, the v/c of uncoordinated phases, the total v/c , the CCL and the intersection saturation degree x . To study the impacts of these factors on α_1 , a multivariate regression can be used. First, the factors are introduced into the fitting function and then the variables are eliminated stepwise according to the correlation until the required precision is achieved. Then, the residual variables in the function are those with a significant impact on α_1 . The amount of data in Table 1 is not sufficient to carry out a regression. Accordingly, another 10 experiments are carried out, providing a total of 12 groups of data. The resulting regression function is shown in

$$\alpha_1 = 14.916 - 53.963x + 4.831\lambda_c + 36.281Y, \quad (5)$$

$$R^2 = 0.920.$$

In the function, λ_c is the green split of the coordinated phase and Y is the total v/c . Only these three are retained; the other variables are excluded. An F test is carried out to identify whether the real value and fitted value of α_1 differ significantly (under a significance level of 0.05). The results of the test are acceptable.

3.3. Impact of Link Length on Correlation. Numerical experiments are conducted to establish the relationship between the link length L and the CI. During the experiments, L is increased from 200 m to 1500 m with a step of 50 m while the cycle lengths of the two intersections and the path

flow between coordinated phases are fixed in each scenario. Figure 2 shows the results for two typical scenarios. The cycle lengths of intersection i and j in the two scenarios are set to the same value, 90 s.

In Figure 2, L and PI have a curvilinear relationship and the two curve shapes are similar to each other. As the value of L increases, the value of PI increases up to a maximum, which is attained when L equals about $0.5C_jV$. As L increases further, PI decreases and reaches a minimum when L equals about $0.75C_jV$. The cycle is repeated as L increases with a step size of $0.5C_jV$ meters. The difference is that the values of the maximum and minimum decrease step by step. This changing tendency is significantly different to our usual view, because traditionally we would assume that PI would decrease linearly with an increase in L . However, the traditional view is based on one-directional signal coordination, and in this study we mainly focus on the dual-directional signal coordination of two adjacent intersections. When L equals an integer multiple of $0.5C_jV$, dual-directional interactive coordination is the best approach and thus the traffic control can achieve the maximum benefits. When L equals an odd multiple of $0.25C_jV$, dual-directional coordination produces the worst benefit and thus PI is minimal.

In numerical experiments, L is increased with a step of 50 m; however, the integer multiples of $0.25C_jV$ are not equal to the integer multiples of 50. Thus, to obtain the exact curvilinear relationship between L and PI , the values of PI corresponding to integer multiples of $0.25C_jV$ are included.

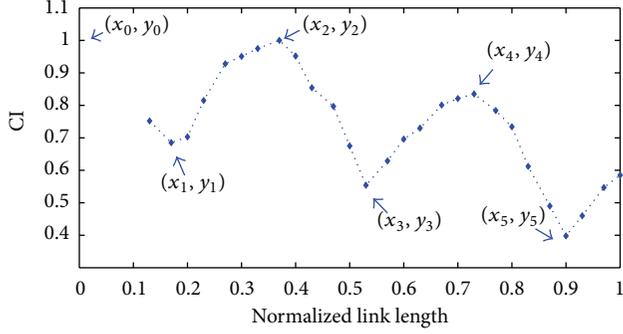
When PI achieves its maximum at L equals $0.5C_jV$, grouping the two intersections into one subarea achieves the maximum control benefits; in this case, CI equals 1. In other cases, CI equals the ratio of PI to the maximum value of PI . L is normalized as follows:

$$L_c = \frac{L}{L_{\max}}, \quad (6)$$

where L_c is the normalized link length.

TABLE 2: Values of CI corresponding to the extreme points.

Scenario	First minimum	First maximum	Second minimum	Second maximum	Third minimum
1	0.663	1	0.533	0.858	0.4
2	0.674	1	0.527	0.816	0.372

FIGURE 3: Scatter diagram between L_c and CI (scenario 1).

L_{\max} is set as 1500 m in this study because signal coordination is not suitable when L is larger than 1500 m according to engineering experience. The relationship between CI and L_c is similar to that shown in Figure 2. Thus, the scatter diagrams are not displayed for these two variables.

Following detailed analysis, it was concluded that a piecewise linear function may be suitable to fit the curve. Moreover, the differences between adjacent minimal values of CI are nearly constant, while the differences between each maximal value of CI and its previous minimal value are also nearly constant. For scenarios 1 and 2, the extreme values of CI corresponding to integer multiples of $0.25C_jV$ are shown in Table 2.

In scenario 1, the difference between the first maximum and the first minimum equals 0.337. The difference between the second maximum and the second minimum equals 0.325. The two numbers are very close. The difference between the first minimum and the second minimum equals 0.130, while the difference between the second and third minimum equals 0.133. These numbers are also very close. In scenario 2, the same phenomenon is observed. The extreme points of the L_c -CI curve are labeled and shown in Figure 3.

In the numerical experiments, L ranges from 200 to 1500. However, theoretically, it is only known that L should be larger than 0. Thus, in Figure 3, the curve is extended to the interval $[0, 200]$. However, the values of CI corresponding to that interval do not have any practical meaning because, when L is smaller than 200 m, the objective of the signal control should be to minimize the queue length or avoid queue spillovers, rather than to reduce vehicle delay. On the basis of the changing tendency of CI, y_0 should be larger than 1. But again this has no practical meaning, so it is set equal to 1 for convenience.

To fit the curve in Figure 3 using a piecewise linear function, the values of y_1 and y_3 are needed. The difference

between y_2 and y_1 represents the reduction in CI when the link length decreases from $0.5C_jV$ to $0.25C_jV$. Similarly, the difference between y_2 and y_3 represents the reduction in CI when the link length increases from $0.5C_jV$ to $0.75C_jV$. The size of the reduction is affected by many factors, such as the CCL, the green split and the v/c of the coordinated phase, the v/c of the uncoordinated phases, the total v/c , and the saturation degree, x . The multivariate regression method described in Section 3.2 is also adopted here to determine which of the above factors have a significant impact on y_1 and y_3 . In total, 10 groups of data are used in the regression.

The regression functions are shown in (7). In the two functions, only x , y_c , and Y are retained and the other variables are excluded. y_c is the v/c ratio of the coordinated phase:

$$\begin{aligned}
 y_1 &= 0.381 + 0.992 \cdot x - 0.870 \cdot Y \\
 &\quad + 0.291 \cdot y_c, \\
 R^2 &= 0.950, \\
 y_3 &= -2.716 + 11.682 \cdot x - 9.814 \cdot Y \\
 &\quad + 1.942 \cdot y_c, \\
 R^2 &= 0.960.
 \end{aligned} \tag{7}$$

Then, y_n in the L_c -CI curve equals

$$y_n = \begin{cases} 1.0, & n = 0, \\ 0.381 + 0.992 \cdot x - 0.87 \\ \quad \cdot Y + 0.291 \cdot y_c, & n = 1, \\ 1.0, & n = 2, \\ -2.716 + 11.682 \cdot x - 9.814 \\ \quad \cdot Y + 1.942 \cdot y_c, & n = 3, \\ y_{n-1} + y_{n-2} - y_{n-3}, & n \geq 4 \text{ and even integer,} \\ y_{n-2} + y_{n-4} - y_{n-2}, & n \geq 4 \text{ and odd integer.} \end{cases} \tag{8}$$

The n th one in the sets of lines can be fitted by

$$\begin{aligned}
 \text{CI}(n) &= \alpha_2 \cdot L_c + b_2, \\
 \alpha_2 &= \frac{y_n - y_{n-1}}{x_n - x_{n-1}}, \quad b_2 = \left(y_{n-1} - \frac{y_n - y_{n-1}}{x_n - x_{n-1}} \cdot x_{n-1} \right), \tag{9}
 \end{aligned}$$

where α_2 is the slope of the line, and b_2 is the intercept.

TABLE 3: Impact of the difference between q_s and $q_{s \max}$ on control benefits.

Scenario	P_q	D_u/s	D_c/s	P_{dk}/s	P_{du}/s	PI_n	P_n
1	0%	265103	77701	0	0	0	0.00%
	5%	271860	72611	5090	6757	1667	3.08%
	10%	276639	69391	8310	11536	3226	5.95%
	15%	283045	67884	9817	17942	8124	14.99%
	20%	287528	64228	13473	22425	8951	16.52%
	25%	297323	59695	18006	32220	14214	26.23%
	30%	299637	57025	20676	34534	13858	25.57%
	35%	305460	53310	24391	40357	15965	29.46%
	40%	309357	49541	28160	44254	16094	29.69%
2	0%	198840	59700	0	0	0	0.00%
	5%	202996	58878	613	4156	3543	5.95%
	10%	206948	57892	3032	8107	5075	8.52%
	15%	210717	56741	5926	11877	5950	9.99%
	20%	212268	56461	7176	13428	6253	10.49%
	25%	215648	54948	9405	16808	7403	12.43%
	30%	218853	53261	11658	20013	8354	14.02%
	35%	221884	51394	12925	23044	10118	16.98%
	40%	224744	49342	14916	25903	10988	18.44%

For scenario 1, the function between L_c and CI is as follows:

$$CI = \begin{cases} -1.872 \cdot L_c + 1.0, & 0 \leq L_c \leq 0.18, \\ 1.872 \cdot L_c + 0.326, & 0.18 < L_c \leq 0.36, \\ -2.594 \cdot L_c + 1.934, & 0.36 < L_c \leq 0.54, \\ 1.871 \cdot L_c - 0.478, & 0.52 < L_c \leq 0.72, \\ -2.595 \cdot L_c + 2.738, & 0.72 < L_c \leq 0.90, \\ 1.871 \cdot L_c - 1.282, & 0.90 < L_c \leq 1.0. \end{cases} \quad (10)$$

The results of the F test show that there is no significant difference between the real values and the fitted values of CI. Thus, (10) can be used to depict the relationship between L_c and CI in scenario 1.

3.4. Impact of Path Flow on Correlation. Let q_s denote the path flow. The maximum value of q_s (denoted by $q_{s \max}$) equals the maximum flow of the upstream coordinated phase. Since signal coordination is not suitable to be implemented to the near saturated intersection, $q_{s \max}$ is taken to be the maximum historical traffic volume of the upstream coordinated phase when the saturation degree is no larger than x_c . In this paper, it is recommended that x_c be set at 0.90.

Again taking intersections i and j as an example, two scenarios are tested. Intersections i and j have the same cycle length and saturation degree. The $q_{s \max}$ values in the two scenarios are 640 pcu/h and 670 pcu/h, respectively. The path flows in each direction are the same. q_s is decreasing from $q_{s \max}$ with a decrement of 5%, while the cycle length and link length remain fixed.

Table 3 shows the results. In the table, P_q is the ratio of $(q_{s \max} - q_s)$ to $q_{s \max}$, which increases from 0 to 40%. D_u is the total vehicle delay per hour of uncoordinated phases. To keep the cycle length fixed in each scenario, the reduced traffic

flow is transferred equally among the uncoordinated phases. Thus, the traffic flow and vehicle delay in the uncoordinated phases will increase. D_c is the total vehicle delay per hour of coordinated phase. Because the traffic volume of the coordinated phase decreases gradually, D_c also decreases. P_{dk} is the decreased vehicle delay of coordinated phase compared with that when P_q equals 0. P_{du} is the increased vehicle delay of uncoordinated phases compared to that when P_q equals 0. PI_n is the negative performance index resulting from the decrease in q_s . P_n is the ratio of PI_n to the PI when P_q equals 0.

In Table 3, the signal control between the two intersections achieves maximum benefits when P_q equals 0. Therefore, CI equals 1. P_n indicates the degradation of PI due to an increase in P_q . Taking scenario 1 as an example, P_n is 14.99% when P_q equals 15%. In this case, the control benefits equal 85.01% of the PI when P_q equals 0; that is, CI equals 0.8501. CI is the difference between 1 and P_n . P_q can be used as the normalized index of path flow. The scatter diagrams between P_q and CI for the two scenarios are shown in Figure 4.

As can be seen from Figure 4, there is a strong linear relationship between P_q and CI, so a linear function can be fitted to the scatter diagrams. The fitted functions are shown in (11). The R squares of the two functions are both larger than 0.94, which indicates that a linear function fits the relationship well:

$$CI = \begin{cases} -0.828 \cdot P_q + 0.997, & R^2 = 0.950, \\ -0.401 \cdot P_q + 0.973, & R^2 = 0.940. \end{cases} \quad (11)$$

The problem is that, as in the earlier case, the slopes of the two functions are different. The slope may again be affected by our previously used set of factors, namely the CCL, the green split and the v/c of the coordinated phase, the v/c of the

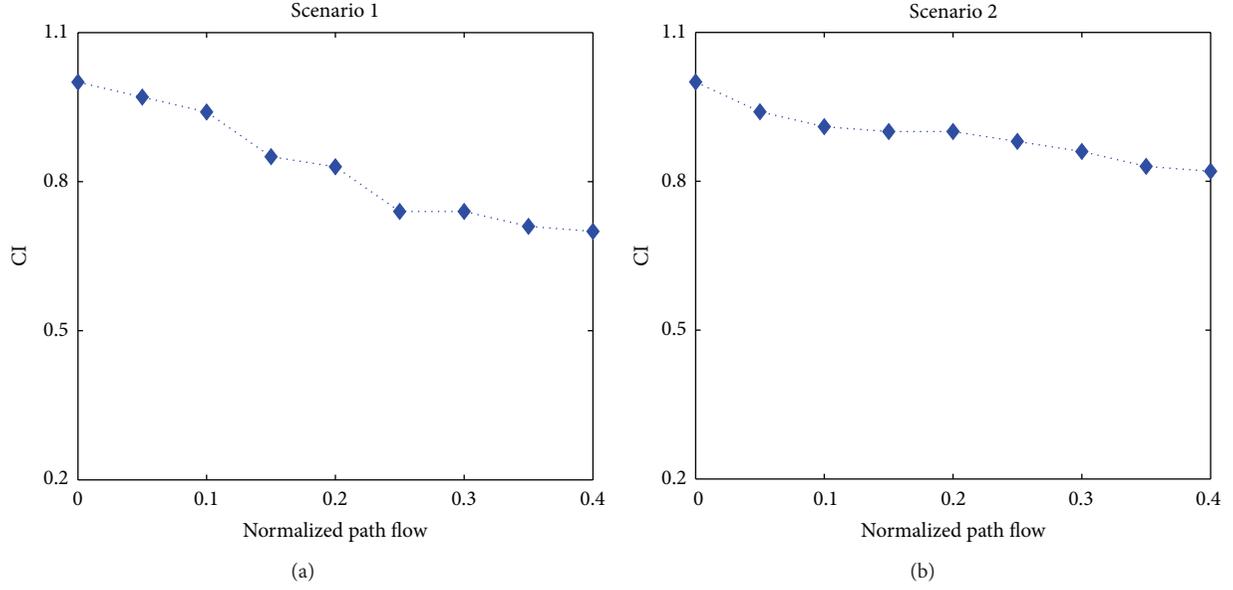


FIGURE 4: Scatter diagrams between normalized path flow and CI.

uncoordinated phases, the total v/c , and the saturation degree x . The multivariate regression method shown in Section 3.2 is again adopted here to distinguish whether any of these factors has a significant impact on the slope. The universal function between P_q and CI may be written as

$$CI = \alpha_3 \cdot P_q + b_3, \quad (12)$$

where α_3 is the slope and b_3 is the intercept.

As in the earlier case, b_3 equals 1 theoretically but there are marginal differences between the intercepts in (11) and 1 because of fitting error. Once again, in this study, the differences are neglected and b_3 is set as 1.

A total of 10 groups of data are used in the regression and the result is shown in

$$\begin{aligned} e^{\alpha_3} &= 5.946 - 2.664 \cdot \log(\text{CCL}) + 0.140 \cdot y_u \\ &\quad - 0.049 \cdot \lambda_c, \quad (13) \\ R^2 &= 0.93. \end{aligned}$$

α_3 is affected by the CCL, y_u , and λ_c . Equation (13) can be rewritten as

$$\begin{aligned} \alpha_3 &= \ln(5.946 - 2.664 \cdot \log(\text{CCL}) \\ &\quad + 0.140 \cdot y_u - 0.049 \cdot \lambda_c). \quad (14) \end{aligned}$$

In the above numerical experiments, the path flows in the two directions are assumed to be the same. However, in reality they may be different. Further numerical experiments are conducted and show that P_q in (12) can be calculated as follows:

$$P_q = \frac{(P_{q1} + P_{q2})}{2}, \quad (15)$$

where P_{q1} and P_{q2} are the ratios between the current path flow and the maximum historical path flow in each direction.

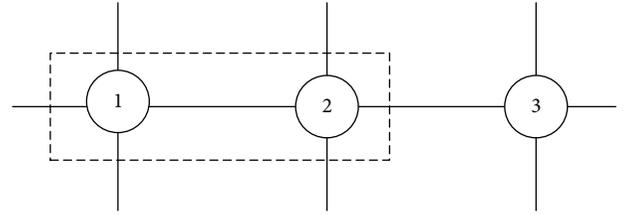


FIGURE 5: Sketch of three adjacent signalized intersections.

3.5. Impact of Number of Intersections on Correlation. The number of intersections in the subarea affects the number of coordinated traffic streams. As was shown in Figure 5, the arrival streams at intersection 1 in the direction of west to east and that at intersection 3 in the direction of east to west cannot be coordinated because they arrive from outside of the subarea. The ratio of coordinated streams to total streams of coordinated phases in the network is $2/3$. With an increase in the number of intersections, the ratio will decrease. Thus, the proportion of vehicles that can be coordinated will increase and the coordinated control performance will increase. To validate the above deduction, numerical experiments are conducted.

In the numerical experiments, the values of the three other contributing factors are fixed. The difference in cycle lengths is set to 0. The link lengths between adjacent intersections are $0.5C_cV$. The path flows of adjacent intersections are set equal to their maximum historical values. In scenarios 1 and 2, all intersections have four phases. In scenarios 3 and 4, all intersections have three phases. To study the impact of the number of intersections on the coordination performance, we simulate subareas with 2, 3, 4, 5, 6, and 7 intersections in each scenario. The results for the four scenarios are shown in Table 4.

TABLE 4: Results for the four experimental scenarios.

Scenario 1		Scenario 2		Scenario 3		Scenario 4	
n	Improvement ratio						
2	8.81%	2	15.35%	2	13.70%	2	17.10%
3	12.35%	3	20.30%	3	17.39%	3	20.39%
4	14.92%	4	23.37%	4	18.21%	4	21.65%
5	15.96%	5	24.56%	5	19.76%	5	22.36%
6	16.34%	6	25.40%	6	20.55%	6	23.16%
7	17.20%	7	26.20%	7	21.38%	7	23.90%

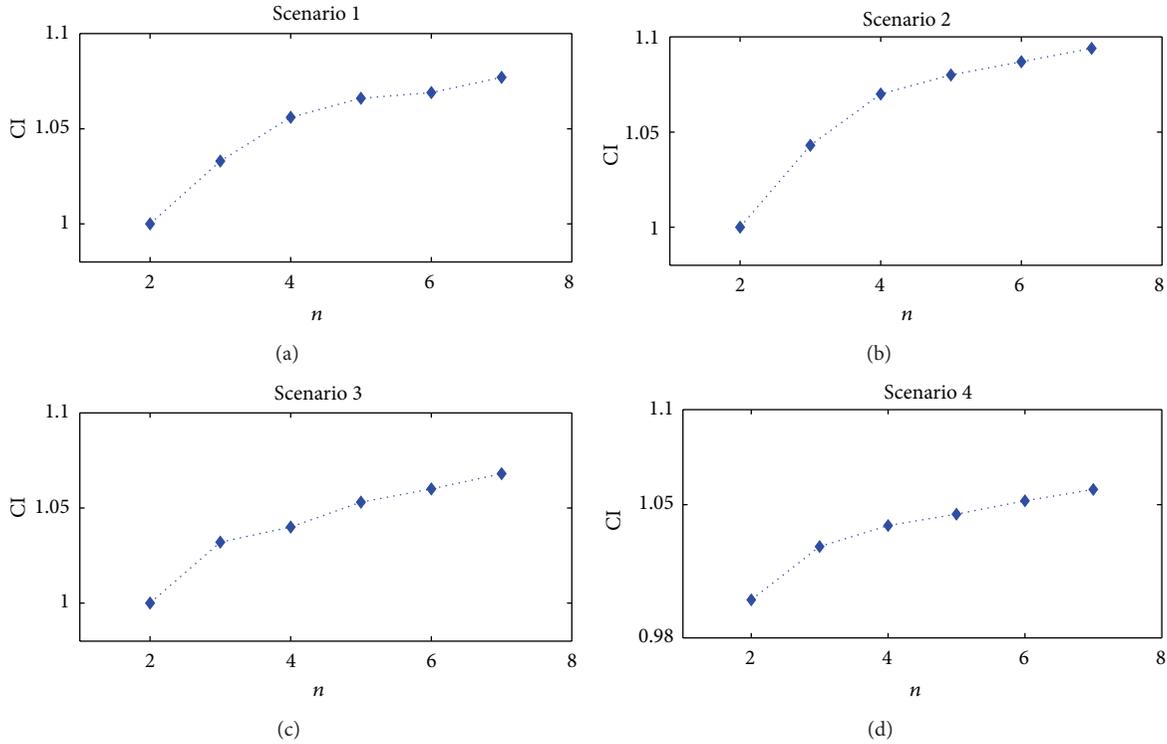


FIGURE 6: Scatter diagram between number of intersections in the subarea and CI.

In Table 4, n is the number of intersections in the subarea. Taking scenario 1 as an example, when n equals 2, the improvement ratio is 8.81%. With an increase in n , the improvement ratio also increases. When n equals 7, the improvement ratio reaches 17.2%. This pattern can also be seen in the other scenarios, demonstrating that the number of intersections indeed impacts upon subarea partition.

Since there is no difference between the cycle lengths in each scenario, the lengths of the links are optimal and the path flows are equal to their maximum historical values. Therefore, when n equals 2, the value of CI for the two adjacent intersections is 1. However, the improvement ratio increases as n increases, so CI also increases. This means that CI will be larger than 1 when n is larger than 2. Based on the above analysis, CI can be calculated as follows:

$$CI(n) = \frac{1 + PIC(n)}{1 + PIC(2)}, \tag{16}$$

where $PIC(n)$ is the improvement ratio when there are n intersections in the subarea.

The correlations between the intersections shown in Table 4 are calculated according to (16), and the scatter diagrams between n and CI for the four scenarios are shown in Figure 6.

From the figure, we can see that, with an increase in n , the increase in CI decreases. The logarithmic function can be used to fit the relation between n and CI. For the four scenarios, the fitting functions are as follows:

$$CI = \begin{cases} 0.061 \ln(n) + 0.963, & R^2 = 0.962, \text{ Scenario 1,} \\ 0.074 \ln(n) + 0.957, & R^2 = 0.960, \text{ Scenario 2,} \\ 0.052 \ln(n) + 0.969, & R^2 = 0.975, \text{ Scenario 3,} \\ 0.044 \ln(n) + 0.974, & R^2 = 0.968, \text{ Scenario 4.} \end{cases} \tag{17}$$

The R squares of the four functions are all larger than 0.96, which indicates that the logarithmic function fits the scatter

diagrams well. However, the slopes of the functions differ from each other, which demonstrates that the impact of the number of intersections on the correlation changes as the scenario changes. The slope is also affected by the same set of factors discussed in the previous sections (the CCL, the green split, etc.). The multivariate regression method shown in Section 3.2 is again adopted here to determine which if any of these have a significant impact on the slope. The universal function between n and CI is given by

$$CI = \alpha_4 \cdot n + b_4, \quad (18)$$

where α_4 is the slope.

CI equals 1 when n is 2. Therefore, $b_4 = 1 - \alpha_4 \ln(2)$. Another 6 groups of data are obtained based on numerical experiments so that a total of 10 groups of data can be used to determine α_4 . The outcome is as follows:

$$\alpha_4 = -0.263 - 0.511y_c + 1.12x - 0.255 \log(C_c). \quad (19)$$

The R square of the regression model is 0.952. α_4 is affected by three factors, the v/c of the coordinated phase, the saturation degree, and the CCL.

3.6. Integrated Correlation Model of Multiple Adjacent Intersections. In Sections 3.2 to 3.5, detailed relationships between various contributing factors and the CI are determined. To show their joint impact on the CI, the following integrated correlation model is proposed:

$$\begin{aligned} CI &= CI(n) \\ &\cdot \left\{ 1 - \left[\sum_{j=1}^n [(1 - CI_{i,j}(C_D)) + (1 - CI_{i,j}(L_c)) \right. \right. \\ &\quad \left. \left. + (1 - CI_{i,j}(P_q))] \right] (n-1)^{-1} \right\}, \quad j \neq i, \\ &= CI(n) \cdot \left\{ 1 - \left[\sum_{j=1}^n [3 - (CI_{i,j}(C_D) + CI_{i,j}(L_c) \right. \right. \\ &\quad \left. \left. + CI_{i,j}(P_q))] \right] (n-1)^{-1} \right\}, \end{aligned} \quad (20)$$

where $CI_{i,j}(C_D)$, $CI_{i,j}(L_s)$, $CI_{i,j}(P_q)$ and $CI(n)$ can be obtained from (4), (9), (12), and (18), respectively.

In (20), i is the seed intersection of the subarea, and j is a nonseed intersection. The term $1 - (CI_{i,j}(C_D))$ indicates the decrease in the correlation because of the difference between the cycle lengths of i and j . The terms in square brackets indicate the total decrease in the correlation due to adding j into the subarea. When j is not adjacent to i , $CI_{i,j}(L_c)$ is the link length correlation between j and the intersection that is adjacent to j and located between j and i . Similarly, $CI_{i,j}(L_c)$ is the path flow correlation between j and the intersection that is adjacent to j and located between j and i .

As analyzed above, the difference in cycle lengths, the path flow, and the link length only affect the correlation of two intersections. Therefore, the three related correlations are subtracted from 1. However, $CI(n)$ is related to all of the intersections in the subarea. Thus, it is set as the base correlation of the subarea and multiplied by the other terms.

3.7. Threshold Value of CI. PI is the improved network performance when signal coordination is implemented. When PI is larger than 0, the coordination of adjacent intersections can obtain better performance than the isolated control mode. In such situations, the two can be grouped into the same subarea; otherwise, they must be placed in different subareas.

In this paper, the CI is developed in order to simulate PI. When PI is larger than 0, CI is also larger than 0. Therefore, 0 is set as the threshold value of CI, to determine whether adjacent intersections should be placed in the same subarea.

4. Case Study

In this section, field experiments are conducted to validate the correlation model developed in this paper. The experimental setup and then the results are described in the following subsections.

4.1. Experimental Setup. A small road network composed of four signalized intersections in Harbin, China, is selected as the experimental network. The four intersections are all located in the central area of the city. A sketch and phasing diagram of the network are shown in Figure 7. Songshan road is an arterial road. All intersections carry out a signal timing scheme with four phases. When executing the signal coordinated mode, the phases in the north-south direction are selected as the coordinated phases. The green lost time of each phase is 3 s and the saturation flow rate of each lane is 1700 pcu/h. For convenience, the four intersections are labeled intersections 1, 2, 3, and 4, respectively, from north to south. The link lengths between adjacent intersections are 730 m, 510 m, and 500 m.

An excellent CI should reflect the performance of the coordinated control mode relative to the isolated control mode precisely. Namely, the correlation should vary as the relative performance varies. Thus, five scenarios are tested. Because the signal coordinated mode is not suitable for those intersections facing oversaturated or close to oversaturated conditions, the five scenarios are all selected during off-peak hours. The duration of each scenario is 15 minutes. Table 5 displays the survey times of the five scenarios.

Traffic data are collected every 5 minutes and the timing schemes are optimized and updated every 15 minutes. Thus, the correlation among the four intersections is calculated once in each scenario. The calculation of the correlation requires some basic traffic parameters, such as cycle lengths, green splits of coordinated phases, and saturation degrees. The values of the parameters are shown in Table 6. C_i is the optimal cycle length of each intersection when operating in isolated control mode.

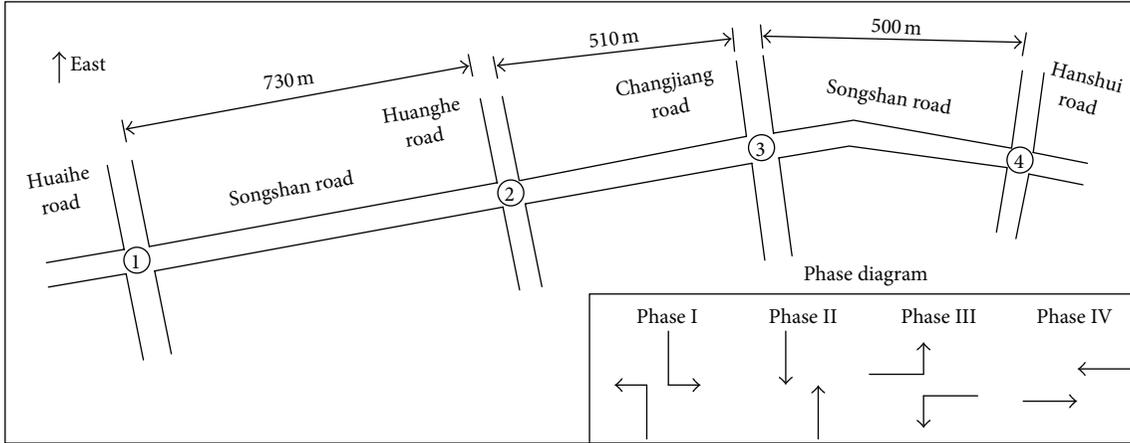


FIGURE 7: Sketch and phasing diagram of the investigated network.

TABLE 5: Survey times of the five experimental scenarios.

Scenario	Survey time
1	9:01 am~09:15 am
2	9:16 am~09:30 am
3	9:31 am~09:45 am
4	9:46 am~10:00 am
5	10:01 am~10:15 am

As can be seen from Table 6, the cycle length of intersection 2 is the longest in all scenarios. Thus, intersection 2 is set as the seed intersection when operating the coordinated mode.

4.2. Experimental Results and Analysis. Table 7 shows the results for the five scenarios.

From scenario 1 to scenario 5, the integrated correlations of the four intersections are 0.42, 0.38, 0.34, 0.51, and 0.55, respectively. Because the threshold value is 0, in the five scenarios, the four intersections can be grouped into one subarea. To test whether CI reflects PI accurately, we also give the values of PI and the improvement ratio of the signal coordination mode to the isolated control mode.

However, CI and the improvement ratio of PI have different dimensions and cannot be compared directly. Thus, the data for CI and the improvement ratio in Table 7 are normalized using a linear function. The normalized data are shown in Table 8.

The scatter diagrams for the scenarios between the normalized CIs and improvement ratios are shown in Figure 8.

From Figure 8 we can see that the two variables have similar changing tendency, but not exactly the same. This is because the correlation model was developed based on numerical experiments and there are some differences between the default values of the numerical experiments and the field values for the traffic parameters.

An *F*-test and a *t*-test are carried out to identify whether the two groups of normalized data differ significantly under a significance level of 0.95. The Data Analysis tool in Microsoft

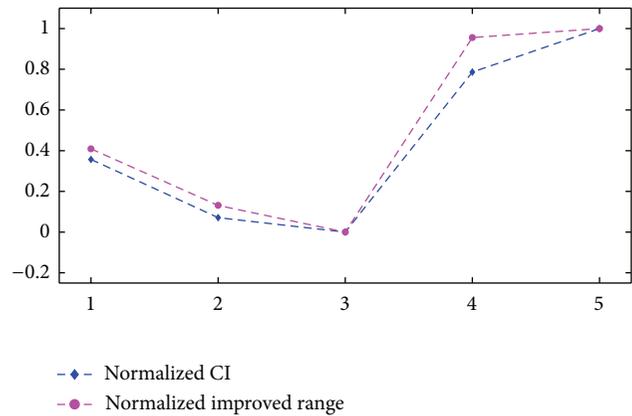


FIGURE 8: Comparison between normalized CI and improvement ratio.

Excel is convenient for executing these tests. The statistical results are shown in Table 9 and all are acceptable.

From the experimental results, we can see that the CI is closely related to the change in the PI, which indicates that the CI can be used as an alternative to the PI. Therefore, the correlation model developed in this paper is suitable for subarea partition.

5. Conclusions

This paper developed a correlation model for multiple adjacent intersections by taking into consideration four contributing factors: the difference between cycle lengths, the link length, the path flow between the upstream and downstream coordinated phases, and the number of intersections. A case study was used to explain the model and the results show that it is reliable for subarea partition.

In this paper, we selected four typical signal intersections located on an arterial road under moderately congested traffic conditions for our experiment. Although the traffic and signal timing conditions of the selected sites are pretty

TABLE 6: Parameters used to calculate the correlations among the intersections.

Scenarios	Intersections	C_i/s	λ_c	x	Y	Y_c	Y_u
Scenario 1	1	108	0.31	0.89	0.79	0.27	0.51
	2	114	0.32	0.89	0.80	0.29	0.51
	3	102	0.29	0.88	0.77	0.25	0.52
	4	106	0.27	0.88	0.78	0.24	0.54
Scenario 2	1	105	0.31	0.88	0.78	0.27	0.51
	2	113	0.33	0.89	0.80	0.29	0.50
	3	105	0.3	0.88	0.78	0.26	0.52
	4	97	0.28	0.87	0.76	0.24	0.52
Scenario 3	1	103	0.3	0.88	0.78	0.26	0.51
	2	110	0.29	0.89	0.79	0.26	0.53
	3	101	0.31	0.88	0.77	0.27	0.50
	4	95	0.29	0.87	0.76	0.25	0.51
Scenario 4	1	98	0.28	0.87	0.77	0.24	0.52
	2	101	0.3	0.88	0.77	0.26	0.51
	3	94	0.27	0.87	0.76	0.23	0.52
	4	92	0.31	0.86	0.75	0.27	0.48
Scenario 5	1	95	0.32	0.87	0.76	0.28	0.48
	2	99	0.33	0.87	0.77	0.29	0.48
	3	93	0.31	0.86	0.75	0.27	0.48
	4	88	0.31	0.86	0.74	0.27	0.47

TABLE 7: Results for the five experimental scenarios.

Parameters	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
$CI_{1,2}(C_D)$	0.86	0.81	0.82	0.91	0.89
$CI_{3,2}(C_D)$	0.71	0.81	0.76	0.80	0.83
$CI_{4,2}(C_D)$	0.81	0.62	0.61	0.74	0.69
$CI_{1,2}(L_c)$	0.71	0.74	0.73	0.72	0.74
$CI_{2,3}(L_c)$	0.86	0.84	0.87	0.74	0.82
$CI_{3,4}(L_c)$	0.83	0.88	0.80	0.82	0.82
$CI_{1,2}(P_q)$	0.84	0.78	0.82	0.87	0.85
$CI_{2,3}(P_q)$	0.79	0.75	0.78	0.86	0.88
$CI_{3,4}(P_q)$	0.77	0.84	0.81	0.86	0.91
$CI(n)$	1.07	1.08	1.100	1.100	1.08
CI	0.42	0.38	0.37	0.48	0.51
PI/s	229740	212245	236980	196572	184634
Improvement ratio	26.3%	22.5%	20.7%	33.8%	34.4%

TABLE 8: Normalized data for CI and the improvement ratio.

Normalized CI	0.357	0.071	0.000	0.786	1.000
Normalized improvement ratio	0.409	0.131	0.000	0.956	1.000

TABLE 9: Statistical results of F -test and t -test.

Sample size	Significance level	F -test		t -test	
		F_c	f	t_c	t
5	0.95	0.16	0.90	2.31	-0.20

F_c is the one-tailed critical value for the F -test; t_c is the two-tailed critical value for the t -test.

typical, the conclusions may still not be transferable to other intersections due to the complexity and diversity of real-life situations. Future research is necessary to develop statistical models to quantitatively assess the generality of these findings.

Acknowledgments

The authors are grateful to Liang-Tay Lin and Shou-Min Tsao. Their paper helped us to find some useful related literature and provided us with some meaningful research ideas. This study is supported by the China Postdoctoral Science

Foundation funded project (no. 2013M530159) and National Natural Science Foundation of China (no. 61304198).

References

- [1] L. T. Lin and S. M. Tsao, "A system approach on signal grouping for areawide control of computerized traffic system," in *Proceedings of the 79th Annual Meeting of the Transportation Research Board*, pp. 1–21, Transportation Research Board, Washington, DC, USA, 2000.
- [2] C. Ding, W. H. Wang, X. Wang et al., "A neural network model for driver's lane-changing trajectory prediction in urban traffic flow," *Mathematical Problems in Engineering*, vol. 2013, Article ID 967358, 8 pages, 2013.
- [3] D. F. Ma, D. H. Wang, Y. M. Bie et al., "A method of signal timing optimization for spillover dissipation in urban street networks," *Mathematical Problems in Engineering*, vol. 2013, Article ID 580546, 9 pages, 2013.
- [4] R. J. Walinchus, "Real-time network decomposition and subnetwork interfacing," *Highway Research Record*, vol. 366, pp. 20–28, 1971.
- [5] Z. Y. Liu, Q. Meng, and S. A. Wang, "Speed-based toll design for Cordon-based congestion pricing scheme," *Transportation Research C*, vol. 31, pp. 83–89, 2013.
- [6] Z. Y. Liu, Y. D. Yan, X. B. Qu et al., "Bus stop-skipping scheme with random travel time," *Transportation Research C*, vol. 35, pp. 46–56, 2013.
- [7] K. Lu, *Research on the foundational theory and key techniques of coordinate signal control in urban traffic network [Ph.D. thesis]*, South China University of Technology, Guangzhou, China, 2010.
- [8] J. Y. Park, *Network-wide signal control with distributed Real-time travel data [Ph.D. thesis]*, University of California at Irvine, Irvine, Calif, USA, 2009.
- [9] L. L. Zang, X. X. Meng, L. Jia, and L. Zhao, "An optimal control framework of traffic signals based on dynamical subarea partition for urban main-roads," in *Proceedings of the 8th World Congress on Intelligent Control and Automation (WCICA '10)*, pp. 4533–4538, IEEE, Jinan, China, July 2010.
- [10] Q. Meng and Z. Liu, "Impact analysis of cordon-based congestion pricing on mode-split for a bimodal transportation network," *Transportation Research C*, vol. 21, no. 1, pp. 134–147, 2012.
- [11] H. Etemadnia, K. Abdelghanya, and A. Hassana, "Network partitioning methodology for distributed traffic management applications," *Transportmetrica A*, 2013.
- [12] H. N. Yagoda, E. H. Principe, C. E. Vick, and B. G. Leonard, "Subdivision of signal systems into control areas," *Traffic Engineering*, vol. 43, no. 12, pp. 42–46, 1973.
- [13] E. C. P. Chang, "Evaluation of interconnected arterial traffic signals," *Transportation Planning Journal Quarterly*, vol. 1, no. 15, pp. 137–156, 1986.
- [14] J. E. Moore and P. P. Jovanis, "Statistical designation of traffic control subareas," *Journal of Transportation Engineering*, vol. 111, no. 3, pp. 208–223, 1985.
- [15] L. T. Lin and H. J. Huang, "A linear model for determining coordination of two adjacent signalized intersections," *Journal of Modeling in Management*, vol. 4, no. 2, pp. 162–173, 2009.
- [16] K. Lu, J. Xu, and Y. Li, "Division method of coordinated control subareas based on correlation degree analysis," *Journal of South China University of Technology*, vol. 37, no. 7, pp. 6–9, 2009 (Chinese).
- [17] Y. Bie, D. Wang, Q. Wei, and D. Ma, "Development of correlation degree model between adjacent signal intersections for subarea partition," in *Proceedings of the 11th International Conference of Chinese Transportation Professionals (ICCTP '11)*, pp. 1170–1180, American Society of Civil Engineers, Nanjing, China, August 2011.
- [18] D. I. Robertson, "TRANSYT A traffic network study tool-Report," Tech. Rep. 253, TRRL Laboratory, London, UK, 1969.
- [19] Transportation Research Board, *Highway Capacity Manual*, National Research Council, Washington, DC, USA, 2000.

Research Article

An Efficient Adaptive Denoising Algorithm for Remote Sensing Images

Xiujie Qu, Fu Zhang, and Huan Jia

School of Information and Electronics, Beijing Institute of Technology, Microelectronics, Beijing 100081, China

Correspondence should be addressed to Xiujie Qu; quxiujie@bit.edu.cn

Received 8 June 2013; Revised 3 September 2013; Accepted 10 October 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Xiujie Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Typically, after the capturing, imaging, and transferring processes have been accomplished, the digital images will contain a variety of noise, caused by both the equipment itself and by the complex working environment. Consequently, it is necessary to perform a de-noising process to facilitate the extraction of useful information. This paper presents a fast and efficient denoising algorithm, which combines the advantages of traditional median filters and weighted filter algorithms. In this algorithm, the noise in the figure is determined, and those results are applied to adaptively change the size of the window, while assigning different weights to the pixels in the filter window. The experimental results show that we can significantly remove almost all salt and pepper noise, while retaining full image textures, edges, and other minutiae.

1. Introduction

Drivers' behavior analysis is increasingly becoming important in the study of intelligent transportation systems (ITS) [1]. Due to the importance of driving behavior to vehicle safety, many researchers have attempted to model driving behavior [2]. Ding et al. took the neural network to learn and incorporate the uncertainties to predict the driver's lane-changing behavior more accurately [2]. To make a good prediction, remote sensing technology, according to Ewan et al. [3], is a critical component since advances in road weather remote sensing technologies have made noninvasive road weather sensors a valuable component in many ITS applications. Specifically, satellite land remote sensing has been widely used in the infrastructure and system engineering of highway transportation systems. Hence, there is a necessity for greater accuracy in remote sensing images. With the fast development of hyper spectral remote sensing technology, the image can describe the characteristics of Earth objects more comprehensively and explicitly [4]. However, the images obtained directly from the sensors, are limited by the performance of the sensor device itself, as well as by the impact of the postprocessing circuit, which contains significant amount of noise interference [5]. Although over the last decades the development of imaging spectrometers is

rapid, remote sensing image is still affected by many complex factors during the processing of acquisition and transmission, which will produce a mass of noises [6].

In order to improve identification and to more effectively reflect objective reality, it is necessary for the digital image to go through a denoising procedure. The factors that impact the quality of the remote sensing images are the undercurrent of the Charge Coupled Device (CCD) camera, the zero response offsets, and the response inconsistency [7]. The noise in imaging systems is usually either additive or multiplicative. After processing by a readout circuit, there are mainly Gaussian and impulse noises remaining. The particle properties of the salt and pepper noise will seriously interfere with the characteristics of very small targets in remote sensing images, severely reducing the credibility of the image data [8, 9]. Therefore, effective denoising processing is very important.

The impulse noise degrades the original image by replacing some image pixels by the noise value. This value could be the maximum and minimum gray level of the image that is known as salt and pepper noise, and a random pixel value in the image gray level that is random-valued impulse noise [10]. It is generated during imaging, transmission, and decoding processes. Additionally, the image cutting process will also produce salt and pepper noise.

The first efficient method in impulse is to carry out a low-pass filter process, based on the median filtering algorithm, which is still in the core of many recent denoising methods and can remove the high-frequency portion of the image [11]. The median filter (MF) replaces each pixel with the median of the neighborhood pixels in the square sliding window that is around it. By using this filter, impulse noise can be discarded, but one of its shortcomings is that in a high-density impulse noise, the number of noisy pixels is greater than that of the noise free ones in the window, and so the median value can also be noisy [12]. The other problem of median filter is that when a pixel is noise-free in nature, it may be altered and replaced by the median value that could lead to the blurring of the image [13]. Another common denoising procedure uses a weighted filter (WF) algorithm [14]. The weighted filter algorithm assigns different weights to the pixels that are in the neighborhood of the center pixel, and then takes a mean value instead of the center pixel's value, to achieve denoising [15]. This algorithm is fast in removing noise. However, it loses much of the image detail and yields an imperfect result at the same time. There are some other denoising methods, such as wavelet transformation [16, 17] and sparse representation [18]. As frequency domain transforms algorithms, the wavelet transform (WT) is a mathematical tool which has been widely used in image processing, and the sparse representation is recent and widely used in image segmentation and is usually based on the over-complete dictionary. Both of the two algorithms are effective in removing noise, but are not easy to implement on hardware.

In this paper, we combine the advantages of these two types of algorithms (MF and WF) and propose a fast and more effective and adaptive filtering denoising algorithm. This algorithm can effectively remove the salt and pepper noise, and obtain a good denoising effect, while preserving details of the image such as textures and contours.

2. Weighted Filter Denoising

Shang and Sui [14], proposed an adaptive weighted filter algorithm that assigns different weights to the pixels that are in the neighborhood of the center spot and determines the importance of the center pixel to achieve denoising. First, assume a point $p(x, y)$ in the source image, the function of which is given as

$$q(i, j) = \{p(x + i, y + j) : -1 \leq i, j \leq 1, i \cup j \neq 0\}. \quad (1)$$

This function represents a 3×3 set of pixels surrounding the center pixel, labeled as neighborhood Q , which is defined as

$$A_{i,j} = |p(x, y) - q(i, j)|, \quad -1 \leq i, j \leq 1, i \cup j \neq 0, \quad (2)$$

where $A_{i,j}$ represents the absolute differences between the point $p(x, y)$ and other points in neighborhood Q . Then, the $A_{i,j}$ differences are organized in ascending order, the first four smallest differences are added into the operator $ROAD(x, y)$. The value of $ROAD(x, y)$ represents the similarities of the nearest four pixels in the neighborhood. Then, an evaluation

must be made based on the center pixel. If the point is valid, we do nothing; otherwise, it is followed with a smooth filter process. The criterion is given as follows:

$$f(x, y) = \begin{cases} \Phi(p(x, y)), & \text{if } (ROAD(x, y)) > \sigma_r, \\ p(x, y), & \text{if } (ROAD(x, y)) \leq \sigma_r. \end{cases} \quad (3)$$

That is to say, if the value of $ROAD(x, y)$ is higher than that of the threshold, then the corresponding point, $p(x, y)$, can be considered to be a noise, and there is a need for denoising; otherwise, we take it as an effective message and there is no need for any transformation. The smaller the threshold selected, the more stringent the noise determination is, but the easier it is to lose the details. The algorithm can automatically assign different weights to the pixels in a designated neighborhood, according to their validity. The closer the pixel is to the center pixel, the higher the weight it is assigned. However, some desirable details were also replaced, especially when the window size was large, yielding the restored image being blurred [15].

3. Adaptive Median Filter

A traditional median filtering algorithm is based on the type of pixels in the neighborhood and takes a gray median value to replace the original pixel [11]. The major drawback of standard median filter is that the filter is only effective to work at low noise densities [15]. That is to say, if the pulse noise density was not too large (the probability distribution of the positive and negative pulses was less than 0.2), then the median filter has a good denoising capability, but the filter size needs to be set in advance. If the area has intensive noise distribution, a small filter window will result in an unobtrusive denoising effect, while a large window will smooth out the details of the edge region.

Nair and Mol [12], has suggested a new adaptive median filtering algorithm as an improvement. The algorithm can automatically modify both the size and shape of the filter window, which can, to a certain extent, avoid the lack of median filtering. If, however, the filter template is large, then a mean filter can be used to effectively improve the filtering performance. The functions of the algorithm are as follows.

- (1) Calculate the median value, Z_{med} , the maximum value, Z_{max} , and the minimum value, Z_{min} , of the pixels in the window. When the size of the window reaches its maximum, calculate the mean value instead.
- (2) Calculate the $A_1 = Z_{\text{med}} - Z_{\text{min}}$ and $A_2 = Z_{\text{med}} - Z_{\text{max}}$, and if $A_1 > 0$ and $A_2 < 0$, then the median value is not a noise, and the algorithm is terminated and taken as an output; otherwise, move on to step (3).
- (3) If the template size is greater than (or equal to) the maximum value, then calculate the mean value. Otherwise, increase the template size and change its shape. If the increased result is less than S_{max} , then go to step (1); otherwise, output the mean value as the final result.

The methodology of Masood et al.'s paper [13] is suitable for a different degree of noise scale. It can adaptively change the size and shape of the filter window to achieve the removal of salt and pepper noise, but the algorithm is limited by the size of the window. When the window becomes too large, the mean filter is essentially a smoothing filter, and is largely responsible for the loss of image details. At the same time, for the actual noise, it still uses a traditional median filtering algorithm to carry out the denoising process, which while also removing some minor objective edges, contours, and other minutiae, it only improves the image denoising visual effects to a limited extent.

4. This Paper's Algorithm

In this paper, the advantages of the weighted filter and the improved adaptive median filter algorithms have been combined, with help from the proposed fast noise suppression algorithm, to produce an improved adaptive weighted correction algorithm. The main concept was to use weighted filters to assign different weights to the pixels in the filter window, and to utilize a median filter (for its good salt and pepper noise removal abilities), to replace the center pixel. At the same time, the concept of noise detection was introduced, and an evaluation of the previous noise point was then used to adaptively change the window size. With a premise of removing noise, improvements have been made to the signal-to-noise ratio of the postprocessing image, while preserving the textures, contours, and other feature details.

For a source image, first a point $p(x, y)$ is assumed, and evaluated to determine whether it is a suspected noise point ($p(x, y) = 255$ or $p(x, y) = 0$), and if not, skip to the next point; otherwise, calculate the number of nonextreme points in the window, which are marked as S_length . If the value of S_length is higher than that of the preset threshold, then reserve the coordinates of the point, while expanding the filter window to continue calculations, or otherwise execute a filtering process. That is to say, center on that point and take its 3×3 neighborhood as a filter window. Then, calculate the absolute differences between the point $p(x, y)$ and other points in the neighborhood, and add the differences into the operator $ROAD(x, y)$. At the same time, set a threshold Q_t (using the best result of 100 experimental results), and if $ROAD(x, y)$ is greater than Q_t , make a denoising process; otherwise, output the original value. The criterion is the same for function (3). During the denoising process, different weights must be assigned to the pixels in the neighborhood, according to their correlation to the center. The higher the correlation is, the greater the weight is, and vice versa. The weight is calculated as

$$\omega_{i,j} = e^{-ROAD(i,j)^2/3 \times \sigma_\omega^3}. \quad (4)$$

In which σ_ω is an experimental constant, and set the $\sigma_\omega = 63$, in order to achieve an ideal removal. Obtain the final weight coefficient for each pixel through normalization. The function is

$$\rho_{i,j} = \frac{\omega_{i,j}}{\sum_{i=-1}^1 \sum_{j=-1}^1 \omega_{i,j}}. \quad (5)$$

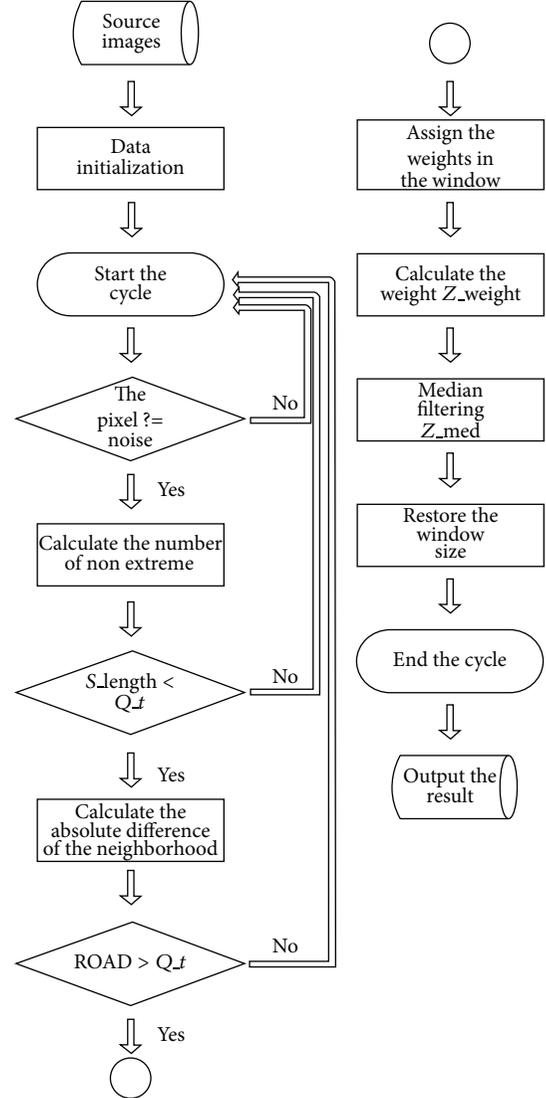


FIGURE 1: The flowchart of the algorithm.

Then, multiply the gray value of each pixel with the corresponding weight coefficient, and add them to get a new center Z_weight , calculated as

$$f(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 \rho_{i,j} \times q(i, j). \quad (6)$$

Take the new result as a replacement for the center pixel and through median filtering, obtain a final value Z_med as the output. We can cycle through all the above steps, until all the pixels have been calculated. The flowchart of the algorithm is shown in Figure 1.

5. Experimental Results and Analysis

In this paper, we took a remote sensing image as a source, and added salt and pepper noise with different densities from 2 to 20%. Through the application of this new algorithm

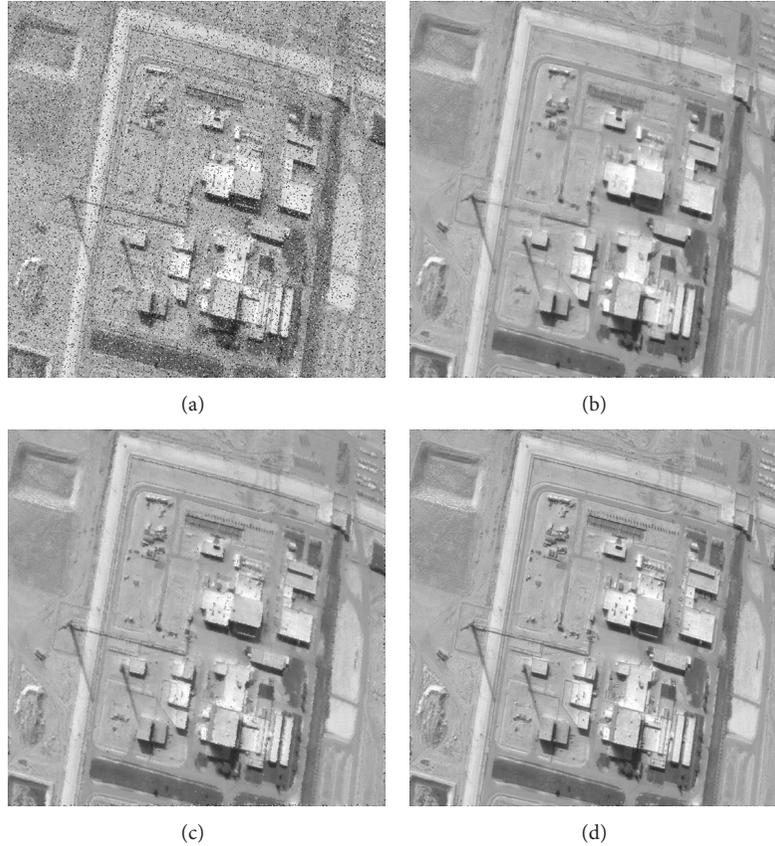


FIGURE 2: (a) Remote sensing image with 10% salt and pepper noise. (b) Weight filter result (3×3 in size of window). (c) Adaptive median filter result. (d) This papers algorithm result.

TABLE 1: Comparison of several de-noising algorithms in PSNR value.

S & P noise density	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
Adaptive median filter	27.9188	27.8494	27.7736	27.5341	27.3256	27.0771	26.9283	26.4809	26.1196	25.6974
Weight filter	33.7074	29.2102	26.0911	23.7517	21.5619	20.5428	19.1369	18.0877	17.2054	16.3512
This paper's algorithm	41.3948	38.7249	37.2327	36.485	34.5214	33.3825	32.684	32.0604	31.7249	31.0123

to the denoising process, and by taking the weighted filter algorithm, and the adaptive median filter algorithm as a comparison, the experimental results shown in Figures 2(a) through 2(d) were obtained, assuming a noise density of 10% for the example.

Judging by the visual effects, the results obtained by the weighted algorithm showed that much of the noise had been removed, but the edges, contours of the runway, houses, and other small targets on the ground, had become blurred, while also losing a large amount of the texture message from the lawns and sandy lands. All of these will be detrimental to the quality and value of a collection of remote sensing information. The denoising results obtained by the adaptive median algorithm were very poor, having only removed some of the noise, and having blurred the contours of the image. Among the three groups, the experimental results obtained with this paper's algorithm were the best. This algorithm removed almost all the noise, and retained some

of the interesting features of small targets such as the tag lines on the runway and the support vehicles, and did not filter out the contours and edges of the remotely sensed images.

Judging from the evaluation index, the greatest value of peak signal-to-noise ratios was obtained by this paper's improved algorithm, when compared with the other two algorithms discussed. With the continuously improving accuracy of the CCD imaging devices, remote sensing image noise density will remain at a low level. The experimental results in Table 1 show that when the noise density is low (less than 10%), the value of the PSNR will be significantly improved by this new algorithm. Therefore, the improved algorithm in this paper has a great practical value.

6. Conclusion

In this paper, the advantages of the weighted filter and the improved adaptive median filter algorithms have been

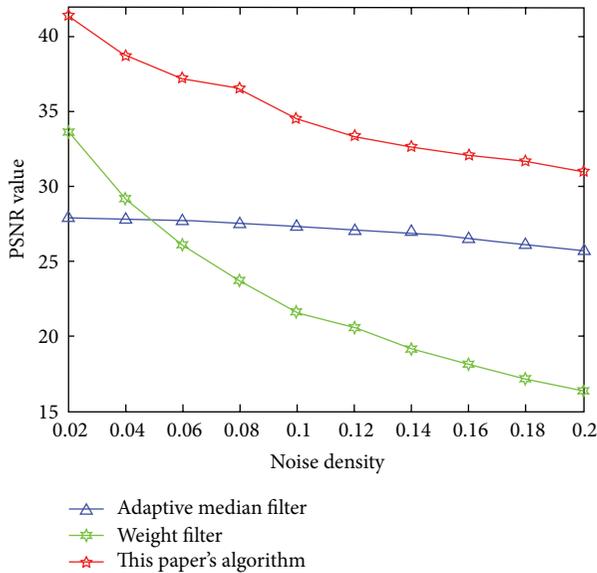


FIGURE 3: PSNR graph.

combined to propose an improved fast weighted-median filter algorithm. The major reason why the proposed method can significantly improve the performance of noise removing is that it adds a noise-detection before the denoising process. By this way, the procedure can quickly find out the noise pot and just directly do the removing process on the noise instead of processing the effective pixels that save a lot of time. This new algorithm can better distinguish the noise and effective information than traditional algorithms, and clean out nearly all salt and pepper noise in the remotely sensed images, while retaining some of the interesting features of small targets such as the contours, edges, and textures. The results obtained by processing with this new algorithm have a high PSNR value, as shown in Figure 3, containing various useful messages which do help to identify the small targets.

References

- [1] W. Wang, H. Guo, H. Bubba, and K. Ikeuchi, "Numerical simulation and analysis procedure for model-based digital driving dependability in intelligent transport system," *KSCE Journal of Civil Engineering*, vol. 15, no. 5, pp. 891–898, 2011.
- [2] C. Ding, W. Wang, X. Wang, and M. Baumann, "A neural network model for driver's lane-changing trajectory prediction in urban traffic flow," *Mathematical Problems in Engineering*, vol. 2013, Article ID 967358, 8 pages, 2013.
- [3] L. Ewan, A. Al-Kaisy, and D. Veneziano, "Remote sensing of weather and road surface conditions is technology mature for reliable intelligent transportation systems applications," *Transportation Research Record*, vol. 2329, pp. 8–16, 2013.
- [4] D. Xu, L. Sun, J. Luo, and Z. Liu, "Analysis and denoising of hyperspectral remote sensing image in the curvelet domain," *Mathematical Problems in Engineering*, vol. 2013, Article ID 751716, 11 pages, 2013.
- [5] X. Zhang, X. Feng, and W. Wang, "Two-direction nonlocal model for image denoising," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 408–412, 2013.
- [6] L. Sun and J.-S. Luo, "Three-dimensional hybrid denoising algorithm in derivative domain for hyperspectral remote sensing imagery," *Spectroscopy and Spectral Analysis*, vol. 29, no. 10, pp. 2717–2720, 2009.
- [7] P. Liu, F. Huang, G. Li, and Z. Liu, "Remote-sensing image denoising using partial differential equations and auxiliary images as priors," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 358–362, 2012.
- [8] G. Y. Chen and B. Kégl, "Image denoising with complex ridgelets," *Pattern Recognition*, vol. 40, no. 2, pp. 578–585, 2007.
- [9] F. Yang, W. Qin, Y. Xie, T. Wen, and J. Gu, "A shape-optimized framework for kidney segmentation in ultrasound images using NLTV de-noising and DRLSE," *BioMedical Engineering Online*, vol. 11, no. 82, 2012.
- [10] M. Nasri, S. Saryazdi, and H. Nezamabadi-pour, "A fast adaptive salt and pepper noise reduction method in images," *Circuits, Systems, and Signal Processing*, vol. 32, no. 4, pp. 1839–1857, 2013.
- [11] S. Esakirajan, T. Veerakumar, A. N. Subramanyam, and C. H. Premchand, "Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter," *IEEE Signal Processing Letters*, vol. 18, no. 5, pp. 287–290, 2011.
- [12] M. S. Nair and P. M. A. Mol, "Direction based adaptive weighted switching median filter for removing high density impulse noise," *Computers and Electrical Engineering*, vol. 39, no. 2, pp. 663–689, 2013.
- [13] S. Masood, A. Hussain, M. A. Jaffar, A. M. Mirza, and T. Choi, "Machine learning and directional switching median-based filter for highly corrupted images," *Knowledge and Information Systems*, vol. 36, no. 3, pp. 557–577, 2013.
- [14] Z. Shang and Y.-K. Sui, "The invariant of the stiffness filter function with the weight filter function of the power function form," *Acta Mechanica Sinica*, vol. 28, no. 6, pp. 1617–1619, 2012.
- [15] C.-T. Lu and T.-C. Chou, "Denoising of salt-and-pepper noise corrupted image using modified directional-weighted-median filter," *Pattern Recognition Letters*, vol. 33, no. 10, pp. 1287–1295, 2012.
- [16] G. Chen, W.-P. Zhu, and W. Xie, "Wavelet-based image denoising using three scales of dependency," *IET Image Processing*, vol. 6, no. 6, pp. 756–760, 2012.
- [17] D. M. Ballesteros L and J. M. Moreno A, "Wavelet de-noising on hardware devices with Perfect Reconstruction, low latency and adaptive thresholding," *Computers & Electrical Engineering*, vol. 39, no. 4, pp. 1300–1311, 2013.
- [18] R. M. Farouk and H. A. Khalil, "Image de-noising based on sparse representation and non-negative matrix factorization," *Life Science Journal-Acta Zhengzhou University Overseas Edition*, vol. 9, no. 1, pp. 337–341, 2012.

Research Article

Railroad Track Deterioration Characteristics Based Track Measurement Data Mining

Peng Xu,¹ Reng-Kui Liu,¹ Feng Wang,² Fu-Tian Wang,¹ and Quan-Xin Sun¹

¹ Beijing Jiaotong University, MOE Key Laboratory for Urban Transportation Complex Systems Theory, 3 Shang Yuan Cun, Beijing 100044, China

² Shanghai Bureau of China Railways, 355 Mo Ling Road, 3rd Floor, Shanghai 200070, China

Correspondence should be addressed to Peng Xu; suepen_aili@126.com

Received 15 July 2013; Accepted 30 September 2013

Academic Editor: Heiner Bubb

Copyright © 2013 Peng Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate information on future railroad track condition is essential to optimally schedule track Maintenance & Renewal activities in order to minimize influences of the activities on rail traffic under constraints of limited budgets and maintaining allowable condition tracks. In this paper, a track measurement data mining method is presented to this aim. It is developed on the basis of track deterioration characteristics. Actual track measurement data is used to analyze errors in track condition predictions by the method. The analysis results show that the proposed method can mine accurate track deterioration rates from historical track measurement data and thus accurately provides future track condition two or three months in advance.

1. Introduction

Traffic accidents have long been social-economical problem which has caused increasing concerns to the public worldwide [1]. According to statistics on train accidents by Office of Safety of US Federal Railroad Administration, 542933 people were injured or killed by railway accidents mainly resulting from railway track from January 1975 to May 2011 [2].

Transportation systems play a critical role in the development of society and economy. Railway system constituted the largest part of national freight ton-miles, for example, 38.2% in 2005 in USA [3] and 49.70% in 2005 in China [4]. Railroad track as a base element of the railway system greatly and directly influences safety and cost efficiency of rail transport. In the process of track management, maintenance-of-way departments have to try to balance the cost associated with potential damages arising from unfavorable tracks and the cost for Maintenance & Renewal activities to minimize the life cycle cost of track. To attain the minimization of the life cycle cost, there are key issues which need to be addressed. One of them is the railroad track condition forecast technology which is able to allow maintenance-of-way departments to acquire accurate track condition information two or three months in advance. Such information is essential to optimally

schedule Maintenance & Renewal activities, constrained by limited budgets and maintaining track in allowable condition, to minimize influences of the activities on rail traffic.

To date, there are several track condition prediction methods developed throughout the world by researchers of universities, technology firms, and railroads. Researchers of the Railway Technical Research Institute of Japan employed Double Exponential Smoothing method to develop track degradation models for predicting standard deviations of track surface and alignment over 100 meters long sections of track [5]. Alfelor and Fateh of the United States Department of Transportation and Carr of the ENSCO Inc. built a track degradation database which stores measurements of Gage Restraint Measurement System, including track gauge restraints and track geometry parameters and contributes factors covering traffic loads, environmental factors, and track structural characteristics [6]. Using the degradation database, a track degradation analysis program was coded, which is able to establish a one-to-one linear relationship between track degradation and a specified contributing factor by employing least square linear regression. Chen et al. of the Academy of Railway Sciences of China proposed an Integrated Factor Method (IFM) to predict values of track geometrical parameters on next month oversampling points

[7]. IFM assumes that track geometrical condition on next month is linearly dependent on the current month's track geometrical condition. The linear evolution rates of a given geometrical parameter are determined according to standard deviations of the geometrical parameter over 200 meters long sections of track. After analyzing the determined evolution rates of a track, Chen et al. categorize evolution rates of a given track geometrical parameter over track sections into 17 groups, each of which has a constant evolution rate. To make the prediction, measurements of the last running of track geometry car are used. Based on a common denominator, "A good track behaves well, while a poorer one deteriorates faster," Veit of the Graz University of Technology and Marschnig, a LCC rail consultant, of Austria considered the relationship between initial and future track condition linear, as well as the relationship between future track condition and elapsed time since the last ballast tamping, cleaning, or renewal exponential, and thus proposed an exponential model to predict condition over 5 meters long sections of track between two adjacent maintenances of given kinds [8]. Quiroga and Schnieder of the Braunschweig University of Technology of Germany proposed an exponential model, like the model of Veit and Marschnig, to predict mean deviations of track surface over 200 meters long sections of track between two adjacent tamping activities [9]. Based on experience of experts in regard to railroad track deterioration, Meier-Hirmer, Riboulet, and Sourget of the SNCF and Rousignol of the Université Paris-Est Marne-la-Vallée of France employed gamma stochastic process to fit the deterioration rate of track surface over 1000 meters long sections of track [10]. According to track deterioration characteristics, Xu et al. proposed a multistage linear method to describe track condition deterioration processes between two adjacent maintenance activities [11]. Based on the research results in [17], Xu et al. employed piecewise linear regression to develop a method for predicting mean values and standard deviations of track condition over unit sections in the future two or three months [12].

According to the track deterioration characteristics, this paper will present a novel method that mines historical track measurements for track condition in future two or three months. The rest of the content will be organized as follows. Section 2 briefly discusses the track deterioration characteristics. Based on the characteristics, the novel track measurement data mining method follows in Section 3. Using actual track measurements, performance of the presented method is analyzed in Section 4. Research conclusions regarding the current research are drawn, and future research areas related to the current research topic are briefly discussed in Section 5.

2. Track Deterioration Characteristics

Generally speaking, railroad track deteriorates as a result of accumulative combinational influences of seven categories of impact factors [13–16]: (1) wheel loads on the rails, (2) track characteristics, (3) materials and manufacture, (4) design and construction, (5) maintenance, (6) environment, and (7) terrain. The influence of the wheel loads is by far the primary

cause for track deterioration. The track characteristics, that is, track configurations and condition of track components, play a critical role in resisting track deterioration and greatly affect dynamic wheel loads. Track deterioration usually begins with small imperfections in the materials and errors in the manufacture of rails and other track components, and performances of the materials and efficacy of the manufactured components are crucial for maintaining track in good condition. Influences of errors during the design and construction of track add to the influences of the materials and manufacture, as the initial source of track deterioration. During a maintenance operation, survey errors, measurement errors, and maintenance machine tolerances may introduce additional track deviations. Moreover, different kinds of maintenance machine usually have different effectiveness. During track deterioration, in addition to the wheel loads, environmental factors directly deteriorate track as well. As the base of railroad track, terrain has obviously direct and great influences on track deterioration. Any variations in terrain will be reflected in rapid deteriorations of track.

During track deterioration, some of these categories of the impact factors interact with each other. For instance, wheel loads deteriorate a track in terms of condition of track components and performances of the materials of the track components. Simultaneously, the deteriorating track increases the wheel loads and reduces the resistance to the deterioration. Such interaction between these three categories of the impact factors continues as trains run over the track.

The above brief introduction shows that there are many kinds of impact factors affecting track deterioration. The combinational influences of all the impact factors vary from one track location to another. In other words, each track point location has its own unique track deterioration process. The uniqueness characteristic of track deterioration has been qualitatively proven during the past several decades of track management practices. To date, only few of the impact factors are measurable, and interactions among the impact factors are unmeasurable. Mainly because of these two facts, when modeled, track deterioration is usually considered random.

3. Track Measurement Data Mining Method

Based on the above introduced track deterioration characteristics, this section will present a track measurement data mining method, which allows maintenance-of-way departments to acquire track condition two or three months in advance. To this aim, brief introductions of track condition and the corresponding measurement data are given first. The mining method is presented last.

3.1. Track Condition and the Corresponding Measurement Data. Track condition is described by track geometrical condition and track structural condition [17]. But track condition usually refers to only track geometrical condition. Our research follows this terming convention as well. Track condition is described by eight geometrical parameters [16]: Gauge, Cross Level, Left/Right Surface, Left/Right Alignment, Twist,

and Curvature. These parameters of tracks under wheel loads are often measured at a specified sampling interval with Track Geometry Car. Within China Railroads, there are four kinds of Track Geometry Cars. GJ-4 is the most extensively used category, and its sampling interval is 0.25 m. In addition to the eight geometrical parameters, measurement data by GJ-4 also includes other two categories of parameters: positioning and comfort. There are two positioning parameters, Milepoint and Auto Location Detection, and four comfort parameters: lateral and vertical box and axle accelerations.

3.2. Characteristics Based Track Measurement Data Mining.

As pointed out in Section 2, track deteriorates uniquely at track point locations. As a result, it is ideal that track condition data should be mined on a track-location basis. In reality, it is impossible to obtain actual values of geometrical parameters over track locations mainly because of errors in milepoint measurements of track measurement data [18, 19]. This means track condition measurements cannot be mined on a point basis. To model track deterioration as accurate as possible, consequently, track deterioration modeling should be done on a short track section (referred to as unit section, hereafter), whose length is determined by the accuracy of milepoint measurements. In previous researches, two fine levels of milepoint error correction model have been developed. After processed with the two correction models, the track condition measurement data can achieve milepoint accuracy most often far below two sampling intervals, that is, 0.5 m. Accordingly, the length of unit section, on which track deterioration will be modeled, is 0.5 meters. Except for unit sections that cover badly damaged rail joints, such track length is reasonable for the track deterioration modeling, because track of 0.5 m in length deteriorates basically similarly.

For a given unit section, the track between two adjacent maintenance operations deteriorates nonlinearly with the accumulative influences of combinational impact factors. But within a short period of time, the track deterioration can be considered approximately linear. In other words, within the short period of time, the track has an approximately identical deterioration rate. It is important that the length of the short time period should match the cumulative combinational influences of all the impact factors. Therefore, if the deterioration rate of a unit section in a short period of time is available, track condition over the unit section can be forecasted. Because track condition is described with 8 geometrical parameters (see Section 3.1), the track deterioration rate of a unit section is characterized by 8 deterioration rates corresponding to the eight geometrical parameters as well.

Because there are at least two sampling points on a unit section, the mean of measurement values of a specified geometrical parameter over the section is used to present the value of the geometrical parameter over the section. Such processing method has two main benefits: (1) reducing the adverse effect of remaining milepoint errors and (2) tolerating the negative effect of noises in geometrical measurements. For a given geometrical parameter on a given unit section,

its deterioration rate within a short period of time can be obtained by employing the least square method to fit values of the parameter in the time period. Unfortunately, for the forecast scenario, the values of the geometrical parameter are unavailable. There is a practical knowledge regarding track deterioration that a section of track with high deterioration rate deteriorates rapidly, and a section of track with slow deterioration rate deteriorates slowly. The reason is that the two categories of impact factors, the track characteristics, and the terrain, over a section of track without maintenance operations involved, are basically unchanged within a short period. The practical knowledge indicates that for a specified geometrical parameter over a given unit section, the deterioration rate in future several months can be estimated by using the historical values of the parameter over the section. The date range, within which the historical values are used to make estimation, must match track deterioration characteristics.

Let $t_{n,x}$ be the number of days between the day when the last maintenance operation was carried out on a track section covering the given section and the day of x days after the n th track geometry car inspection since the last maintenance, and let $a_{n,x}^j$ be the value of the specified geometrical parameter over the j th sampling point in the given section on $t_{n,x}$ days after the last maintenance operation. Assume that the number of sampling points in the given section is l . Let $m_{n,x}$ denote the value of the specified parameter over the given section; that is, $m_{n,x} = (\sum_{j=1}^l a_{n,x}^j)/l$.

Assume that the n th track geometry car inspection is the current inspection and the date range, within which the historical values are used to estimate the deterioration rate, is from $t_{h,0}$ through $t_{n,0}$. Accordingly, the deterioration rate r_n , which is used to forecast track condition, is obtained according to the following:

$$r_n = \frac{(n-h+1) \sum_{i=0}^{n-h} t_{h+i,0} m_{h+i,0} - \left(\sum_{i=0}^{n-h} t_{h+i,0} \right) \left(\sum_{i=0}^{n-h} m_{h+i,0} \right)}{(n-h+1) \sum_{i=0}^{n-h} (t_{h+i,0})^2 - \left(\sum_{i=0}^{n-h} t_{h+i,0} \right)^2}. \quad (1)$$

Track condition in the future can be considered the sum of the current condition and the cumulative combinational effects of all impact factors, which are quantified by the deterioration rate r_n . Therefore, the value of the specified parameter on a day, $t_{n,x}$, over the j th sampling point in the given section can be estimated as $\hat{a}_{n,x}^j = a_{n,0}^j + x * r_n$, where $\hat{a}_{n,x}^j$ denotes the estimation for $a_{n,x}^j$. Let $A_{n,x} = [a_{n,x}^1, \dots, a_{n,x}^l]^T$ and $\hat{A}_{n,x}$ be the estimation for $A_{n,x}$; that is, $\hat{A}_{n,x} = [\hat{a}_{n,x}^1, \dots, \hat{a}_{n,x}^l]^T$. The above given process of mining historical measurement values of the specified track geometrical parameter for future track condition is graphically demonstrated in Figure 1.

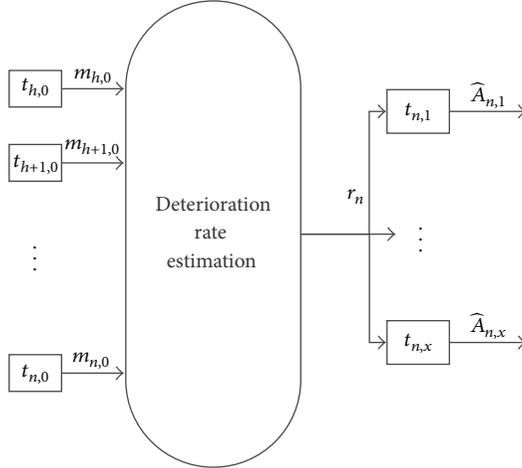


FIGURE 1: The process of mining historical measurement values of a specified geometrical parameter for future condition.

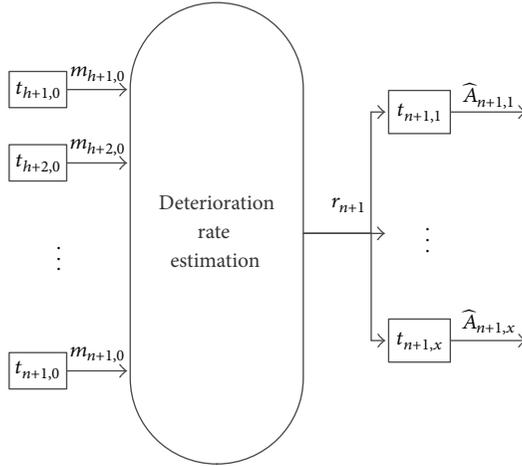


FIGURE 2: The process of mining historical measurements after a new track condition inspection.

As noted previously, on a unit section, the track deteriorates nonlinearly with the cumulative combinational influences of the impact factors. To approximate actual track deterioration rates as close as possible, the estimated deterioration rates should be revised continuously as the impact factors deteriorate the track. Therefore, after a new track condition inspection was carried out, the inspection data is involved in the process of the deterioration rate revision, as demonstrated in Figure 2.

4. Performance Analysis

The Kowloon-Beijing railroad track is very important in the Chinese rail networks. It connects Kowloon to Beijing through 9 provinces. By the million ton kilometers, it ranks in the top sixth of the entire Chinese rail networks; by the million passenger-kilometers, it is ranked fourth. In this section, errors in track condition predictions for a 2 kilometer

TABLE 1: Error statistics for the right alignment predictions.

Date	Mean (mm)	Standard deviation (mm)	Mean absolute error (mm)
Sep. 24	-0.0025	0.3740	0.2412
Oct. 10	0.0005	0.3300	0.2131
Oct. 30	0.0021	0.4165	0.2829
Nov. 13	-0.0033	0.3810	0.2575
Dec. 12	0.0059	0.3859	0.2391
Dec. 25	-0.0031	0.2638	0.1607

long section of the track by the proposed track measurement data mining method will be presented. Specifically, 6 times predictions for values of right alignment and twist over sampling points on the track section are analyzed. These 6 times predictions were made for track condition on September 24, October 10, October 30, November 13, and December 12, and December 25, 2008.

On the presented track section, 60 U71Mn rails were laid and continuously welded, where 60 indicates that a piece of such rail 1 meter in length weighs 60 kilograms, and U71Mn is the material of the rails. The ties on this section are concrete and their model number is II. The ballast is the first class granite rocks, and thickness of the ballast layer is 50 mm.

We made discussions with field engineers on the length of the short time period, within which the track deteriorates approximately linearly. The time period takes the values of 6 months. Considering the traffic characteristics of the track section, track condition of future two months should better be predicted. Therefore, track measurement data in the last four months is used to estimate the deterioration rate.

For measuring the performance of the proposed method, four statistical indices are calculated. They are the mean, the standard deviation, the mean absolute error, and the correlation coefficient between measurement values and predicted values.

4.1. Right Alignment. Errors for every prediction are presented in the histogram, as shown in Figure 3. Four performance measuring statistical indices are calculated for the prediction on each target date as well, as listed in Table 1.

From these 6 histograms, it is concluded that errors of each prediction are normally distributed around 0 mm. The normal distributions of errors indicate that the proposed track measurement data mining method is able to capture the track deterioration trend components and then to use the captured deterioration trends, that is, the deterioration rates, to make predictions.

The above conclusions drawn from the error distributions are quantitatively verified by the calculated statistical indices. The mean error for each prediction is very close to 0 mm; the standard deviation for each prediction is far below 1 mm. From these statistical facts, it is confident that for right alignment, the proposed method can accurately predict its values two months in advance.

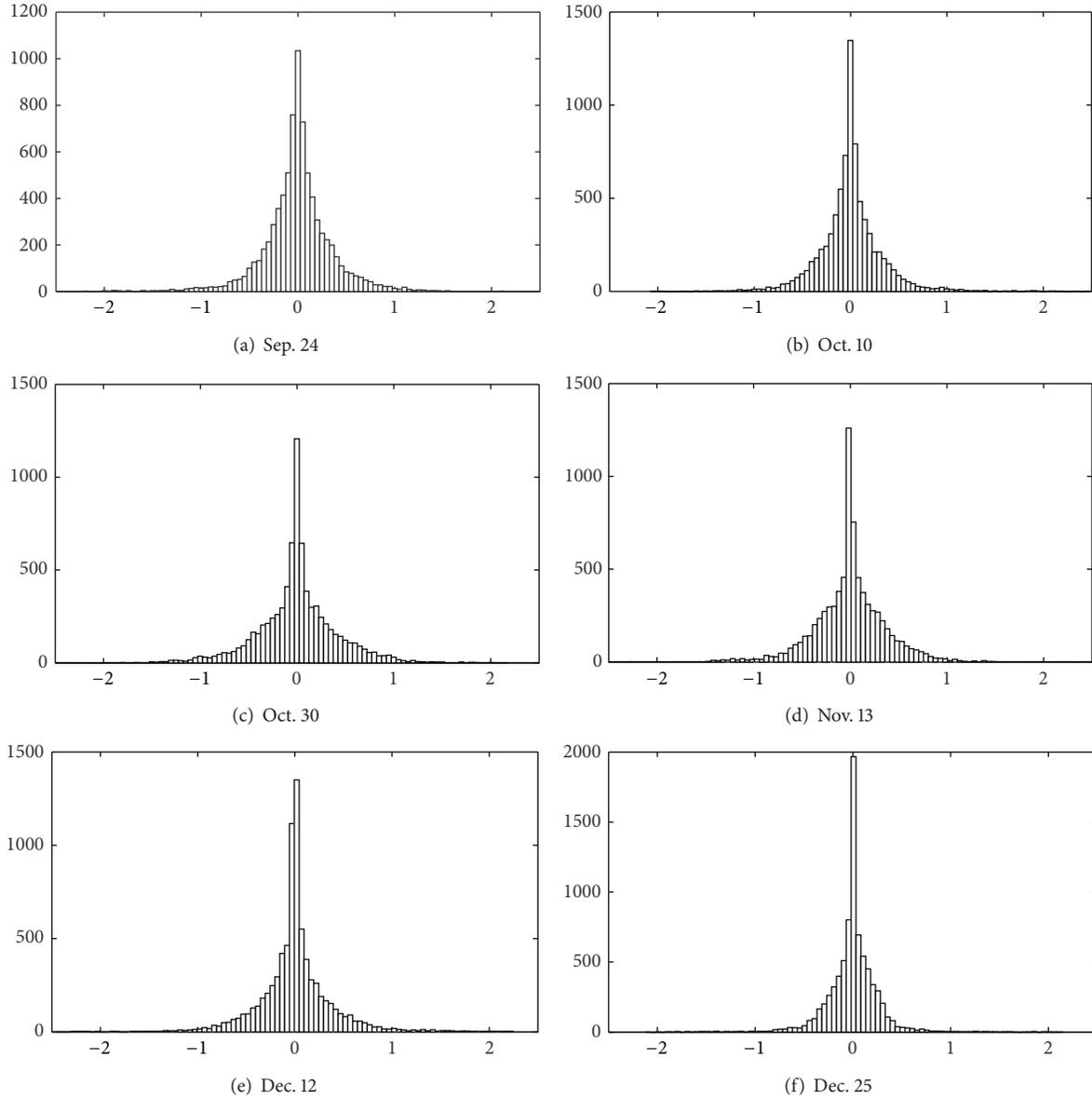


FIGURE 3: Histograms of errors in the right alignment predictions.

TABLE 2: Error statistics for the twist predictions.

Date	Mean (mm)	Standard deviation (mm)	Mean absolute error (mm)
Sep. 24	-0.0070	0.6156	0.3783
Oct. 10	0.0099	0.5360	0.3174
Oct. 30	-0.0033	0.4442	0.2606
Nov. 13	-0.0020	0.5549	0.3592
Dec. 12	0.0016	0.7457	0.4952
Dec. 25	0.0006	0.5979	0.3998

4.2. *Twist.* Following Section 4.1, errors in the twist predictions are graphically shown in the histogram, as illustrated in Figure 4. The statistical indices of errors are worked out for each prediction, as listed in Table 2.

Histograms in Figure 4 show that errors in the twist predictions are normally distributed around 0 mm. These facts indicate that deterioration trends in the twist are captured by the proposed method. Statistical indices in Table 2 confirm the conclusions drawn from Figure 4. The mean error for each prediction is very close to 0 mm; the standard deviation is far below 1 mm. From these statistics, it is straightforward to infer that the proposed method can accurately predict values of twist over sampling points two months in advance.

5. Conclusions and Future Research Areas

In this paper, a track measurement data mining method has been proposed. The method can mine the track condition deterioration trends which are essential to make predictions

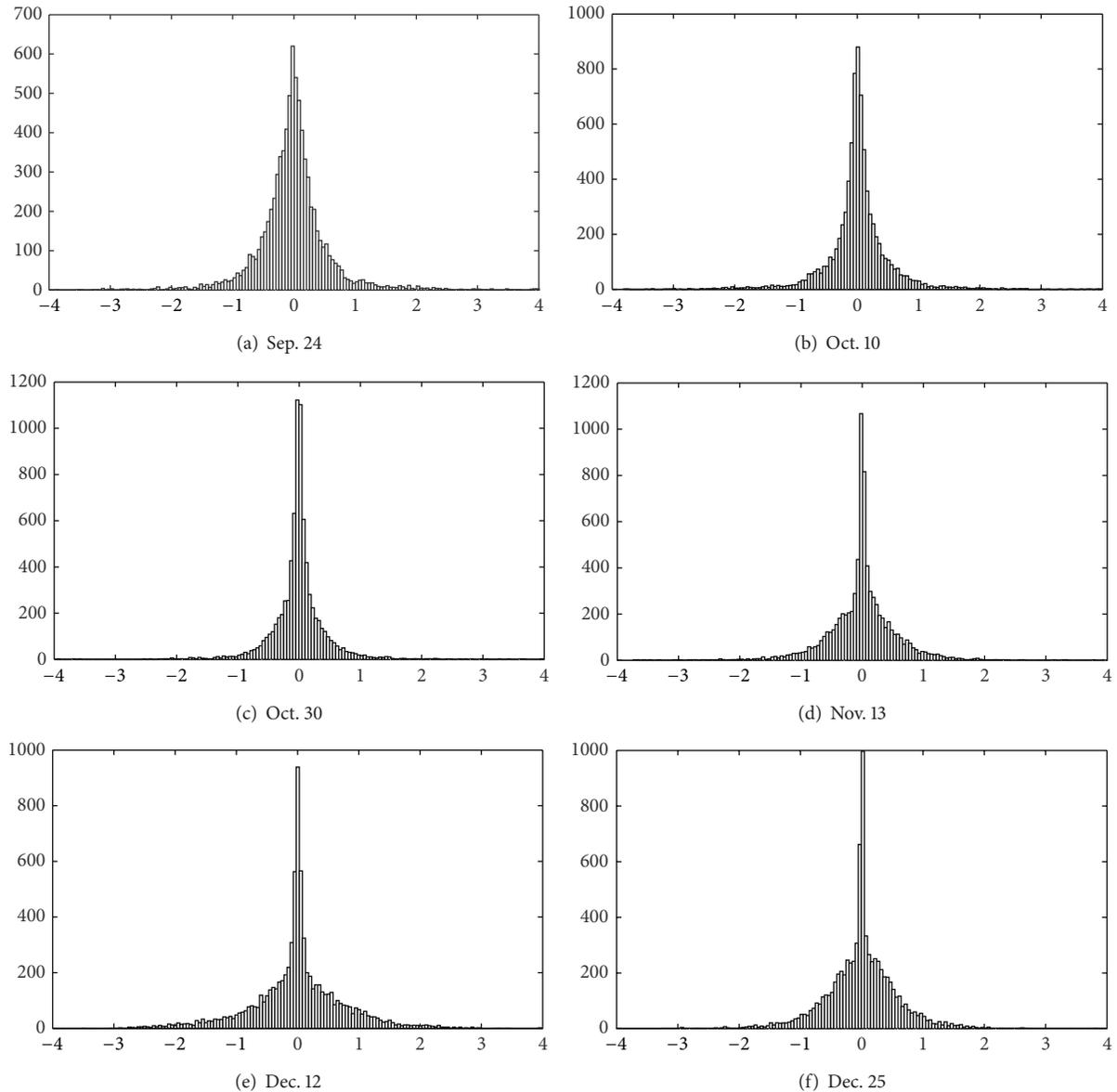


FIGURE 4: Histograms of errors in the twist predictions.

for track condition. Track condition measurement data of the Kowloon-Beijing railroad track was used to analyze errors in track condition predictions by the method. The analysis results show that the proposed method can accurately predict values of geometrical parameters over sampling points.

The analysis on the track deterioration characteristics shows that if the length of the time period (see Section 3.2) can be determined in accordance with actual track deterioration processes, errors in track condition predictions can be reduced further. That means more reliable track condition predictions can be available. Therefore, this direction towards improving the proposed method will be investigated. What is more, the distributions of errors in track condition predictions show that if a normal random variable is incorporated into the proposed method, it will be enhanced again.

Conflict of Interests

The authors state that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was partly sponsored by the National Basic Research Program of China (2012CB725406).

References

- [1] H. Guo, W. Wang, W. Guo, X. Jiang, and H. Bubb, "Reliability analysis of pedestrian safety crossing in urban traffic environment," *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.

- [2] “Federal Railroad Administration Office of Safety Analysis—Train Accidents,” 2011, <http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/Query/inctally1.aspx>.
- [3] RITA, “Table 1: U.S. Domestic Freight Ton-Miles by Mode [WWW];” 2007, http://apps.bts.gov/publications/special_reports_and_issue_briefs/special_report/2007_07_27/html/table.01.html.
- [4] National Bureau of Statistics of China, “China Statistical Yearbook 2012 [WWW];” 2012, <http://www.stats.gov.cn/tjsj/ndsj/2012/indexeh.htm>.
- [5] A. Kawaguchi, M. Miwa, and K. Terada, “Actual data analysis of alignment irregularity growth and its prediction model,” *Quarterly Report of RTRI*, vol. 46, no. 4, pp. 262–268, 2005.
- [6] R. M. Alfelor, G. A. Carr, and M. Fateh, “Track degradation assessment using gage restraint measurements,” *Transportation Research Record*, no. 1742, pp. 68–77, 2001.
- [7] X. M. Chen, L. Wang, F. C. Yang, X. S. Chai, and W. Q. Wu, “Integrating factor method for predicting the developing trend of railway track irregularity,” *China Railway Science*, vol. 27, no. 6, pp. 27–31, 2006.
- [8] P. Veit and S. Marschnig, “Sustainability in track—a precondition for high speed traffic,” in *Proceedings of the ASME Joint Rail Conference (JRC ’10)*, pp. 349–355, April 2010.
- [9] L. M. Quiroga and E. Schnieder, “A heuristic approach to railway track maintenance scheduling,” in *Proceedings of the 12th International Conference on Computer System Design and Operation in the Railways and other Transit Systems (COM-PRAIL ’10)*, pp. 687–699, September 2010.
- [10] C. Meier-Hirmer, G. Riboulet, F. Sourget, and M. Roussignol, “Maintenance optimization for a system with a gamma deterioration process and intervention delay: application to track maintenance,” *Journal of Risk and Reliability*, vol. 223, no. 3, pp. 189–198, 2009.
- [11] P. Xu, R. K. Liu, F. T. Wang, Q. X. Sun, and H. L. Teng, “A novel description method for track irregularity evolution,” *International Journal of Computational Intelligence Systems*, vol. 4, no. 6, pp. 1358–1366, 2012.
- [12] P. Xu, Q. Sun, R. Liu, and F. Wang, “A short-range prediction model for track quality index,” *Proceedings of the Institution of Mechanical Engineers F*, vol. 225, no. 3, pp. 277–285, 2011.
- [13] A. Hamid, K. Rasmussen, M. Baluja, and T.-L. Yang, *Analytical Descriptions of Track Geometry Variations, DOT/FRA/ORD-83/03.1.*, ENSCO, Springfield, Va, USA, 1981.
- [14] T. B. Maertens, *The relationships of climate and terrain to maintenance of way on the Norfolk Southern Railroad between Norfolk, Virginia, and Portsmouth, Ohio [Ph.D. thesis]*, University of Tennessee, Knoxville, Tenn, USA, 1990.
- [15] J. Sadeghi and H. Askarinejad, “Influences of track structure, geometry and traffic parameters on railway deterioration,” *International Journal of Engineering, Transactions B*, vol. 20, no. 3, pp. 291–300, 2007.
- [16] E. T. Selig and J. M. Waters, *Track Geotechnology and Substructure Management*, chapter 15, Thomas Telford, London, UK, 1994.
- [17] J. M. Sadeghi and H. Askarinejad, “Development of track condition assessment model based on visual inspection,” *Structure and Infrastructure Engineering*, vol. 7, no. 12, pp. 895–905, 2011.
- [18] P. Xu, Q. X. Sun, R. K. Liu, and F. T. Wang, “Key equipment identification model for correcting milepost errors of track geometry data from track inspection,” *Transportation Research C*, vol. 35, pp. 85–103, 2013.
- [19] P. Xu, Q. X. Sun, R. K. Liu, R. R. Souleyrette, and F. T. Wang, “Optimizing the alignment of inspection data from track geometry cars,” *Computer-Aided Civil and Infrastructure Engineering*, 2014.

Research Article

Research on Face Recognition Based on Embedded System

Hong Zhao, Xi-Jun Liang, and Peng Yang

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

Correspondence should be addressed to Hong Zhao; 594286500@qq.com

Received 8 July 2013; Revised 5 September 2013; Accepted 25 September 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Hong Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Because a number of image feature data to store, complex calculation to execute during the face recognition, therefore the face recognition process was realized only by PCs with high performance. In this paper, the OpenCV facial Haar-like features were used to identify face region; the Principal Component Analysis (PCA) was employed in quick extraction of face features and the Euclidean Distance was also adopted in face recognition; as thus, data amount and computational complexity would be reduced effectively in face recognition, and the face recognition could be carried out on embedded platform. Finally, based on Tiny6410 embedded platform, a set of embedded face recognition systems was constructed. The test results showed that the system has stable operation and high recognition rate can be used in portable and mobile identification and authentication.

1. Introduction

Face recognition technology [1] emerged in 1980s, developed rapidly, and obtained staged achievements since 1990s. Gradually, it has been applied in feature search system, authentication system [2], and access control system [3]. Due to a number of images feature data, complex calculation, and the larger storage space and high processing capacity, currently, most face recognitions are realized only by PCs with high performance; so portability and mobility in this process are restricted greatly. At present, the embedded system [4] is widely used in the front-end of entrance guard system and attendance system in order to collect face images. Then, the information collected is transferred to the back-end over the network, and the face recognition is carried out by the back-end PCs. However, this working mode heavily relies on recognition by the back-end and is limited by the bandwidth and stability of data transmission network. It still cannot reach the purpose of the move at will.

Since portability and mobility in this process are restricted greatly in the current face recognition system, it is necessary to develop a set of face recognition systems in which both image collection and recognition are realized on the embedded system.

2. Principle of Face Recognition

The process of face recognition is divided into two stages, training and recognition stages, shown in Figure 1.

2.1. Face Training. To ensure convenient face image processing, the original YUV format image is transformed to IplImage format image. Haar-like face detection algorithm (Viola-Jones method) is used to identify face region [5]. In order to enhance the contrast of image, reduce the influence from external factors and improve the following recognition rate; the face image identified is processed with the histogram equalization.

In order to obtain main features of original image, Principal Component Analysis [6] (PCA) subspace of eigenfaces, the PCA is used to extract subspace of eigenface from face image processed. This method can effectively reduce redundant data, and data can be processed in a low-dimensional feature space. Meanwhile, most information of the original image is saved.

2.2. Face Recognition. Similarly, the test of face image is processed through format transformation, Haar-like face detection, and histogram equalization.

The face image processed is projected to the PCA subspace of eigenfaces; thus, projection coefficients on the

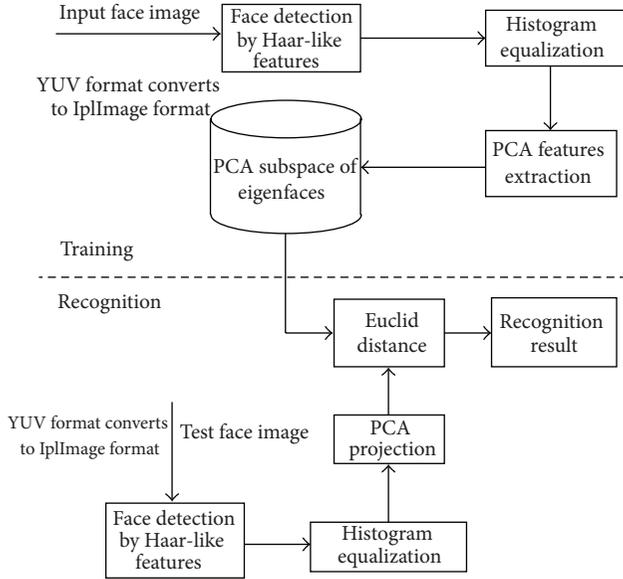


FIGURE 1: Principle of face recognition.

subspace can be obtained. The projection coefficients represent the position of test face image in the PCA subspace of eigenfaces. Then, the analysis contrasts the coefficients with the coefficients of PCA subspace of eigenfaces; finally, the face can be recognized by using Euclidean Distance [7].

3. Major Algorithms

3.1. Haar-Like Face Detection Algorithm

3.1.1. *Haar-Like Features.* Viola-Jones in 2001 published a paper [8] which was a watershed in the real-time face detection technology. The real-time face detection was realized through combining Adaboost algorithm and Cascade algorithm. Papageorgiou and Viola put forward the original Haar-like features when they applied wavelet transformation to extract features from images. The feature library contained features of three types and four kinds. The three types, two-rectangle feature, three-rectangle feature, and four-rectangle feature, are presented in Figure 2. Since this feature library can only describe the structure with specific directions (horizontal, vertical, and diagonal), the features extracted are relatively rough. Subsequently, Lienhart and Maydt put forward a series of extended Haar-like features [9] as listed in Table 1 based on the basis mentioned above; the edge-feature is extended to 4 types, and linear-feature is extended to 8 types adding 2 center-features. These extended Haar-like features make face recognition more convenient and fast.

Haar feature stands for the differential value of gray level sum between corresponding region of the black rectangle and the white rectangle. And Haar feature reflects the degree of image part-graying. Each feature is composed of 2~3 rectangles, and they are applied in the detection of edge-feature, linear-feature, and center-feature. Value of each

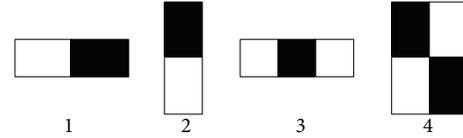


FIGURE 2: Original forms of Haar-like features. (1, 2—edge-feature, 3—linear-feature, 4—diagonal-feature).

feature is made up by the sum of pixel values in corresponding rectangle region, shown as

$$\text{feature}_j = \sum_1^N \omega_i \cdot \text{rectSum}(r_i), \quad (1)$$

where $\text{rectSum}(r_i)$ represents the gray-level integration of image enclosed by rectangular r_i ; N stands for the number of matrices which compose feature_j ; ω_i represents the weight of rectangular region; value +1 represents the weight of white rectangular region; value -1 represents the weight of black rectangular region.

3.1.2. *Integral Image.* Setting the four basic features as an example, permutation, and combination at random in a window with size 24×24 will generate at least hundreds of thousands of features. Calculating eigenvalues of these features will make a large amount of calculation. Therefore, obtaining gray value of the pixel and then making evaluation obviously cannot meet real-time needs. The integral-image method can make fast to process image on calculating the eigenvalues of current subimage.

Integral image is a method of fast calculation of $\text{rectSum}(r_i)$ with the idea of replacing time with space. The sum of pixel values in rectangle region formed from starting point of image to the rest is saved as an array. We can directly use the array to do calculation when we need to calculate $\text{rectSum}(r_i)$ of certain region. This method avoids recalculation of pixels of this region; so calculation speed improves.

3.1.3. *Adaboost Algorithm.* Adaboost algorithm is a kind of classifier algorithms. Its basic idea is constructing an accurate classifier with strong ability of classification by means of combining a large number of simple classifier according to some rules.

(1) *Training of Simple Classifiers.* The form of simple classifier generated by the j th feature is shown as

$$h_j(x) = \begin{cases} 1, & p_j \cdot f_j(x) < p_j \cdot \theta_j \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where h_j is the value of simple classifier; x is a testing sub-window; θ_j is the threshold; p_j is the attribution of sample and it indicates the direction of the sign of inequality; value +1 represents positive samples and -1 negative samples; and $f_j(x)$ is eigenvalue.

Training samples include positive samples and negative samples. The object samples (human face images) to be

TABLE 1: Three kinds of Haar-like features.

Edge-feature		
Linear-feature		
Center-feature		

detected represent positive samples, and any other images represent negative images. All sample images are normalized to the same size with 20×20 .

From formula (2), a weak classifier is decided by both corresponding threshold θ_j and feature f_j . Each feature f_j will be trained to obtain a specified classifier, and the analysis will find the optimal threshold θ_j and minimize the classification error of using this weak classifier to classify all samples under current weight distribution.

For each feature, the corresponding weak classifier is trained. Finally, the weak classifier with lowest classification error ratio for all training samples is selected, called optimal weak classifier.

(2) *Training of Strong Classifiers.* Adaboost classifier includes many optimal weak classifiers which are connected together by some rules and weights. After T times training, T optimal weak classifiers are generated.

A strong classifier has been constructed by T optimal weak classifiers in terms of the following:

$$C(x) = \begin{cases} 1, & \sum_{t=1}^T a_t \cdot h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where T represents the number of optimal weak classifiers included in the strong classifier and $a_t = \log(1 - \epsilon_t)/\epsilon_t$, ϵ_t represents error ratio of the t th optimal weak classifiers.

From formula (3), all weak classifiers have their judgments for testing image. This process is similar to "voting." Then through getting the weighted sum of "voting" in terms of error rates of weak classifiers, the final result can be made by comparing the result of weighted sum of "voting" with the result of average "voting." The result of average "voting" is, namely, probability average value under the condition where the supporting probability is equal to objecting probability.

3.1.4. *Cascade Classifier.* Since the detection process of the strong classifier composed of several weak classifiers costs lots of time, Paul Viola and Michael Jones put forward the cascade

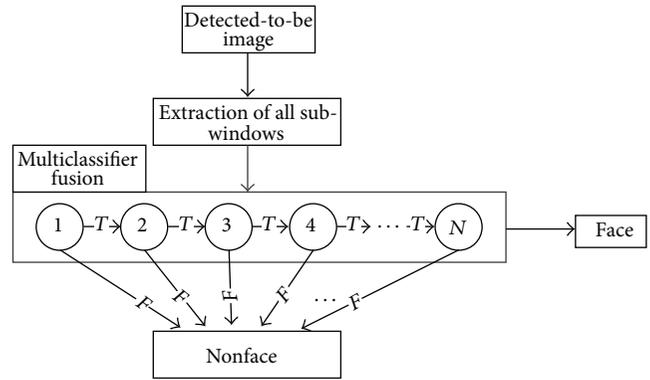


FIGURE 3: Structure of cascade classifier.

face classifier based on Adaboost algorithm. This classifier can detect face quickly and recognize face effectively.

In fact, the multilayer structure proposed by Paul Viola and Michael Jones is a degenerated decision tree. If the current subwindow is a human face when the images go through all simple classifiers in a certain order, the current image will go to the next detection, otherwise, the current sub-window is ceased and the next sub-window will work, as shown in Figure 3.

3.2. *"Eigenface" Recognition Algorithm.* Based on principal component method, "Eigenface" recognition algorithm [10] has been widely applied to face detection and face recognition.

"Eigenface" is the assembly of these eigenvectors corresponding to the large eigenvalues in face covariance matrix. It treats face image as a vector and gets eigenvectors by Karhunen-Loeve transform [11]. The eigenvectors which are similar to the face are called eigenface. The linear combinations of these eigenvectors are used to describe, represent, and recognize the face image.

The preprocessed face image is projected to the subspace composed of "Eigenface". Then projection coefficients on the subspace can be obtained. The projection coefficients that

represent the position of test face image in the PCA subspace of eigenfaces compare with coefficients of the subspace of eigenfaces and finally recognize by using Euclidean Distance.

The method of computing credibility is the key to recognition. Now the best method of computing confidence is base on Euclid Distance, as shown in the following:

$$C = 1.0 - \frac{\sqrt{D / (\text{trainFacesNum} \cdot \text{eigenVectorsNum})}}{255.0}, \quad (4)$$

where C represents credibility; D is the Euclidean Distance between the projection of a test image and the projection of the trained images; trainFaceNum is the number of faces in training; and eigenVectorsNum is the number of face eigenvectors.

4. System Construction

4.1. General Architecture. Based on Tiny6410 embedded platform, a set of embedded face recognition system is constructed. The man-machine interface of the system is programmed with Qt graphic library; the part of video gathering is implemented by video interface of v4l2 [12] in Linux; and OpenCV library is applied in the video processing part.

Two steps in the system work as follows:

- (1) detection stage: system searching for the face region (displayed by rectangle) in the whole image;
- (2) recognition stage: contrasting the face image obtained above to the face image trained in the database, and then judging the person who it is.

If the system recognition is successful, the recognition result will be displayed in white text and the system will pop a dialog box which shows login. If failed, system will pop a warning dialog.

4.2. Working Process

4.2.1. System Training. As shown in Figure 4, test ID is inputted firstly, a frame of image from USB camera is gotten by system, and then the image is grayed and processed with histogram equalization in order to enhance the degree of contrast of image. Next, the preprocessed image is determined whether it will be added to the training set. PCA algorithm is applied to deal with all images in the training set when the number of images in training set reach the presented number. Finally, the XML database is generated.

4.2.2. System Recognition. In the recognition stage, system reads the database file of trained images and applies PCA algorithm to compare test image with the database data. If credibility goes beyond the threshold value, the corresponding user name is displayed on the screen and a message will be popped; otherwise, a warning dialog will be popped.

System can capture several images from camera and compute average credibility of these images in order to improve the accuracy and the reliability.

Figures 5 and 6 are recognition result of pretraining and recognition result of posttraining, respectively.

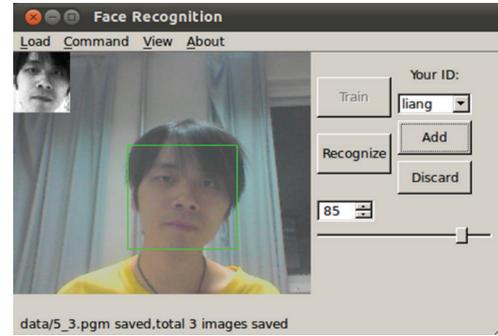


FIGURE 4: Face training.

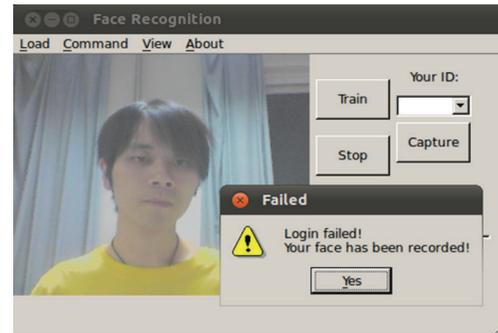


FIGURE 5: Recognition result of pretraining.

5. System Test

The system carries out 9 times tests.

In the No. 1~No. 4 tests, only the face data of person A is added to the face database. At this time, person B and person C are “strangers.” Only Person A can login the system.

In the No. 5~No. 9 tests, the face data of person B is added to the face database. At this time, only person C is a “stranger,” but both person A and person B can login the system.

The results of test are shown as Table 2.

In the 1st test, the time consuming for login person A is 14 s. At this time, system temporarily saves the login information of person A. In the 2nd test, person A can login immediately. In the 3rd and 4th test, since person B and person C did not login before, and person B and person C are not trained, thus B and C cannot login. At this time, system still saves the login information of person A. So person A can login immediately. In the 5th and 6th tests, person B is trained. At this time, person B can login immediately, and system temporarily saves the login information of person B. In the 7th test, because person C did not be trained, person C cannot login. In the 8th test, system still saves the login information of person B when person A login again. So system needs to rescan all sub-windows (time consuming is 12 s), and temporarily saves the login information of person A. In the 9th test, system still saves the login information of person A when person B login again. So system needs to rescan all sub-windows (time consuming is 11 s), and temporarily saves the login information of person B.

TABLE 2: Test results of embedded face recognition system.

No. of test	No. of person	Time for login (s)	No. of frames for recognition	Average credibility	No. of error recognition
1	A	14	2	0.856	0
2	A	—	10	0.866	0
3	B	∞	10	0.675	0
4	C	∞	10	0.778	0
5	A	—	10	0.878	0
6	B	—	10	0.906	0
7	C	∞	10	0.824	0
8	A	12	2	0.883	0
9	B	11	2	0.913	0

(1) Time consuming for login: referring to the time consumed from test user sitting in front of camera to test user login successfully.

(2) Number of frames recognized: referring to the number of frames used to calculate the average credibility.

(3) —: referring to the immediate login.

(4) ∞ : referring to the failure of login.

(5) The threshold value of credibility is set as 0.85.

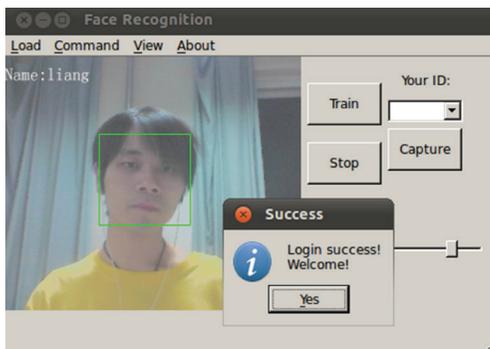


FIGURE 6: Recognition result of posttraining.

6. Conclusion

This paper introduces the specific face recognition technology which is based on embedded platform and puts forward a solution, which stresses on face detection algorithm, face recognition algorithm, and application development. This technology makes full use of the advantage of PCA algorithm on feature extraction and the advantages (such as fast detection speed and high detection rate) of AdaBoost algorithm based on Haar. A set of embedded face recognition system based on Tiny6410 embedded platform is realized. After face recognition testing, the results showed that this system runs stably and has high recognition rate. Thus, it can be widely used in the Things of Internet that needs to verify user identification through portable and mobile methods [13] and in Intelligent Transportation System [9, 10] that needs face recognition technology. In the future research, the Cortex A8 embedded platform that has better ability of floating-point operation will be applied in the system in order to further improve the overall performance of the system.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant no. 61262016, University

Foundation of Gansu Province under Grant no. 14-0220, Natural Science Foundation of Gansu under Grant no. 1208RJZA239, and the Technology Project of Lanzhou (2012-2-64).

References

- [1] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques," *Journal of Information Processing Systems*, vol. 5, no. 2, pp. 41–68, 2009.
- [2] X. Liu, T. Chen, and B. V. K. Vijaya Kumar, "Face authentication for multiple subjects using eigenflow," *Pattern Recognition*, vol. 36, no. 2, pp. 313–328, 2003.
- [3] D. Bryliuk and V. Starovoitov, "Access control by face recognition using neural networks and negative examples," in *Proceedings of the 2nd International Conference on Artificial Intelligence*, pp. 428–436, Crimea, Ukraine, September 2002.
- [4] W. Ping, "Research on the embedded system teaching," in *Proceedings of the International Workshop on Education Technology and Training and the International Workshop on Geoscience and Remote Sensing*, vol. 1, pp. 19–21, Shanghai, China, December 2008.
- [5] M.-T. Pham and T.-J. Cham, "Fast training and selection of Haar features using statistics in boosting-based face detection," in *Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV '07)*, pp. 1–7, Rio de Janeiro, Brazil, October 2007.
- [6] S. Nedeveschi, I. R. Peter, and A. Mandrut, "PCA type algorithm applied in face recognition," in *Proceedings of IEEE International Conference on Intelligent Computer Communication and Processing (ICCP '12)*, pp. 167–171, Cluj-Napoca, Romania, 2012.
- [7] J. Javed, H. Yasin, and S. F. Ali, "Human movement recognition using euclidean distance: a tricky approach," in *Proceedings of the 3rd International Congress on Image and Signal Processing (CISP '10)*, vol. 1, pp. 317–321, Yantai, China, October 2010.
- [8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518, Kauai, Hawaii, USA, December 2001.
- [9] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *Proceedings of the International*

- Conference on Image Processing (ICIP '02)*, vol. 1, pp. 900–903, September 2002.
- [10] E. Gumus, N. Kilic, A. Sertbas et al., “Eigenfaces and support vector machine approaches for hybrid face recognition,” *The Online Journal on Electronics and Electrical Engineering*, vol. 2, no. 4, pp. 308–310, 2010.
- [11] M. Effros, H. Feng, and K. Zeger, “Suboptimality of the Karhunen-Loève transform for transform coding,” *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1605–1619, 2004.
- [12] L. Yinli, Y. Hongli, and Z. Pengpeng, “The implementation of embedded image acquisition based on V4L2,” in *Proceedings of the International Conference on Electronics, Communications and Control (ICECC '11)*, pp. 549–552, Ningbo, China, September 2011.
- [13] Z. Hong, S. Chao, and C. Jie, “Integration schemes for IoT application systems with diverse domain,” *Journal of Beijing Institute of Technology*, vol. 32, no. 12, pp. 201–204, 2012 (Chinese).

Research Article

Empirical Analysis of Traffic Bottleneck at Beijing Expressways

Sheng Jin, Dianhai Wang, and Dongfang Ma

College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China

Correspondence should be addressed to Dongfang Ma; mdf2004@zju.edu.cn

Received 23 July 2013; Revised 16 September 2013; Accepted 16 September 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Sheng Jin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The expressways in Beijing are confronted with more serious traffic congestions. Based on the survey data obtained from the typical sections at the expressways, the time dependent characteristics of traffic flow parameters were analyzed in detail and the data gap was found in this paper. The Fast Fourier Transform (FFT) method is proposed to transfer the data of traffic flow parameters for describing the fluctuation characteristics of traffic flow. Two methods of identification, the graph method and the control line method, were proposed as to the change time of traffic bottleneck forming and dissipating. The findings in this paper have already been applied in traffic management and ramp control at the expressways in Beijing.

1. Introduction

Nowadays, traffic congestions at the expressways have become more serious in Beijing, and the situation is deteriorating. One of the key reasons for traffic congestions is the merging conflicts of the entering and exiting traffic streams caused by the close distances from the upstream entrance ramp (or exit ramp) to the downstream one and the numerous entrance and exit ramps [1–3]. Traffic bottlenecks form easily in the conflict zones along the expressways, and the queue spilled over may result in congestion or even jam when traffic demand increases. Therefore, it is necessary to research on the traffic bottleneck for traffic management or ramp metering.

A prosperous literature can be observed on this topic. Bertini and Cassidy presented some traffic features at freeway bottlenecks. Observations from two-freeway bottlenecks in and near Toronto, Canada, indicated that average vehicles discharge rate from a queue could be 10% lower than the flow measured prior to the queue's formation. The present findings came by virtually comparing transformed curves of cumulative vehicle arrival number versus time and cumulative occupancy versus time measured at neighboring loop detectors [4]. Muoz and Daganzo researched the bottleneck mechanism of a freeway diverging and found that an off-ramp queue may hamper freeway flow much more than an on-ramp bottleneck does [5]. By analyzing bottleneck formation on freeways, Das and Levinson identified “active

bottleneck” locations on freeways and sections where bottlenecks occurred because of disturbances caused by downstream bottlenecks propagating backwards in the form of shockwaves [6]. Ogut and Banks worked on the stability of traffic flow at freeway bottlenecks [7]. Hu and Schonfeld developed a traffic simulation and optimization model to analyze traffic flow in large networks with severe queuing and to transfer traffic volume at bottlenecks [8].

Most of the work about traffic bottleneck mainly focused on the freeways, but a few on urban expressways, especially on those metropolises as in Beijing. Beijing Municipal Institute of City Planning and Design simply compared the traffic flow characteristics of freeways with those in some developed countries. It should be noted that many transport policies are under discussion in Beijing for congestion mitigation, for example, implementing the traffic congestion pricing [9, 10] and prompting the public transport systems [11]. In this paper, in order to analyze the traffic bottleneck formation, the time dependent characteristics of traffic flow parameters are analyzed in detail and the data gap is found on the basis of the survey data of typical sections chosen at Beijing expressways. The graph method and the control line method are put forward in order to assess the critical time of traffic state transition during bottleneck formation and dissipation. The former determines the critical time by adjusting curves of cumulative arrival vehicle number versus time and cumulative occupancy versus time. The latter designs an index,



FIGURE 1: Map of the survey sites.

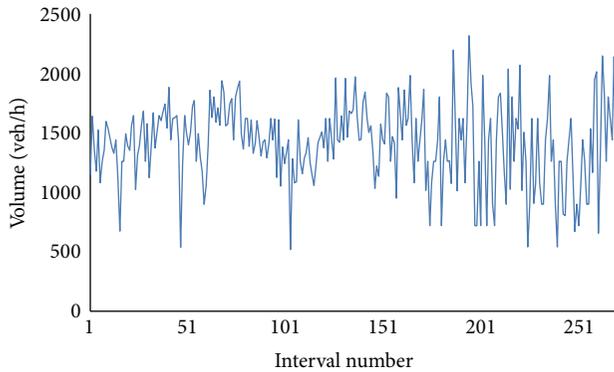


FIGURE 2: Volume on median lane at Site 1.

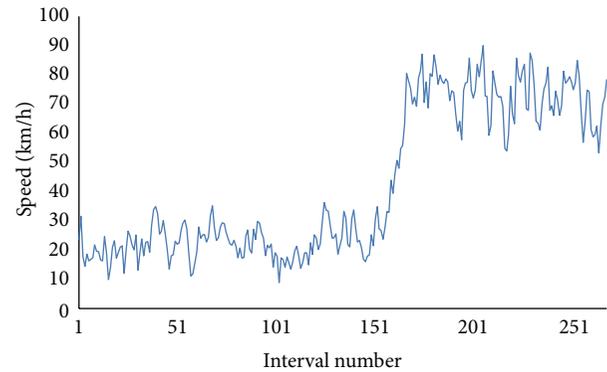


FIGURE 3: Speed on median lane at Site 1.

relative time occupancy which can better describe the traffic state, to calibrate the top and bottom boundary of the data gap according to the quality management principle.

2. Traffic Flow Characteristics Analysis

2.1. Time Dependent Characteristics. The field data (traffic volume, speed, and time occupancy) of a merge area and its upstream and downstream were obtained through the video survey between Sitong Bridge and Lianxiang Bridge in North Ring III at Beijing expressways from 7:00 a.m. to 10:00 a.m. (including morning peak hour) and from 16:00 p.m. to 19:00 p.m. (including evening peak hour) between June and July, 2010. To simplify, the site from Lianxiang Bridge to Sitong Bridge is noted as Site 1, and the opposite direction as Site 2 (See Figure 1).

In data processing, 20 seconds was defined as a statistic time interval, and data were obtained including speed, traffic volume, and time occupancy. The time intervals 20 s, 30 s, 1 min, and 5 min are often used in traffic flow analysis. In order to describe the microcosmic fluctuation characteristics of traffic flow, we use the least time interval which can be provided by advanced transportation information system (ATIS). Here, speed is the harmonic average value of spot speed. Traffic volume is the vehicle number per hour per lane (veh/h/lane). Time occupancy is the ratio of time occupied by vehicles and statistical interval. The data are to analyze traffic flow characteristics and describe the process of traffic bottleneck forming and dissipating.

Figures 2–7 show the change of traffic flow parameters on the median lane at the survey section of Site 1 and Site 2 during the bottleneck dissipating. Figures 2 and 5 indicate traffic

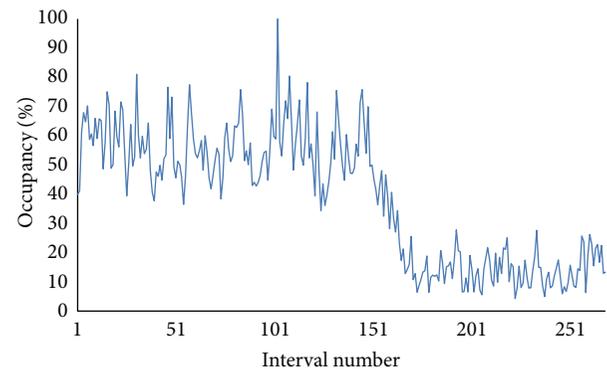


FIGURE 4: Occupancy on median lane at Site 1.

volume without obvious change from bottleneck forming to bottleneck dissipating. Figures 3 and 4 show that the speed (near 20 km/h) is lower and the time occupancy (near 60%) is higher when the traffic flow keeps in congestion state. The speed (near 75 km/h) goes up quickly, and on the contrary, the time occupancy (below 60%) goes down quickly when the bottleneck dissipates. In Figures 6 and 7, the speed (near 75 km/h) is higher and the time occupancy (below 20%) is lower when the traffic flow keeps in the normal state. The speed (near 20 km/h) goes down quickly and the time occupancy (upper 60%) goes up quickly when traffic bottleneck forms, that is, the traffic flow changes from normal to congestion. When it is changed from congestion to normal, the speed goes up and the time occupancy goes down quickly, respectively. Data from the two sites indicates that the curves of speed and time occupancy can reflect the state transition from congestion to noncongestion (or the reverse), but the

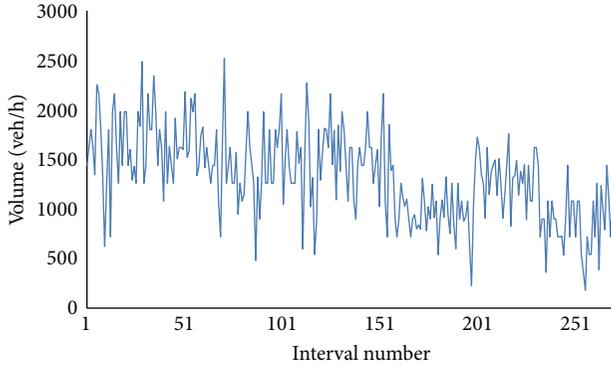


FIGURE 5: Volume on median lane at Site 2.

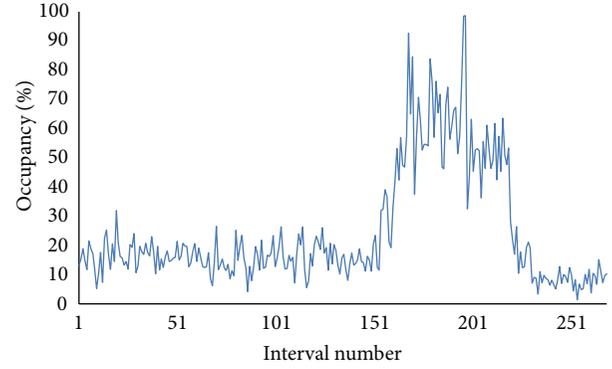


FIGURE 7: Occupancy on median lane at Site 2.

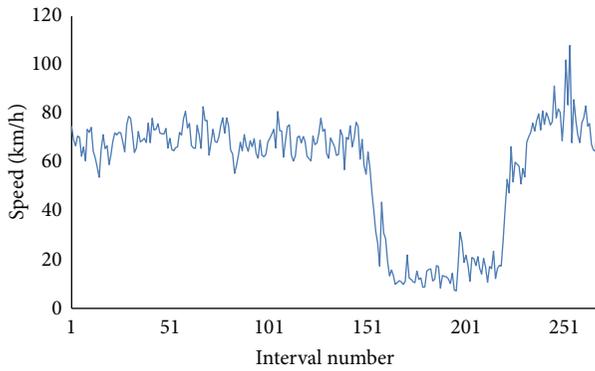


FIGURE 6: Speed on median lane at Site 2.

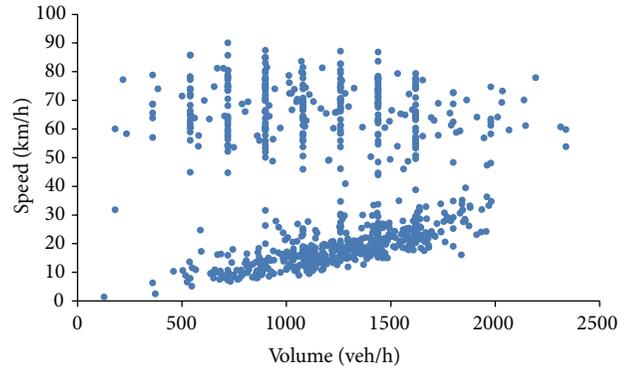


FIGURE 8: Volume-speed plot at Site 1.

curve of traffic volume does not. The change rule of traffic flow parameters is similar for the middle lane and shoulder lane.

Figures 8, 9, and 10 are, respectively, the volume versus speed scatter plot, the occupancy versus volume scatter plot, and the occupancy versus speed scatter plot on all the lanes of the merge area downstream at Site 1.

In Figures 8–10, it is found that data pots cluster into two regions, one represents congestion data, while the other represents noncongestion data, and there is a space with sparse data, without any data in between. The space is defined as data gap [12–18]. The data gap also reflects the transition from one traffic state to another.

2.2. Fluctuation Characteristics. From the time dependent characteristics figures (see Figures 2–7) of traffic flow parameters, it is found that these parameters fluctuate with time when the traffic state is in congestion (see Figures 11, 12, and 13). Some fluctuation characteristics of speeds at freeways had been studied [19–21]. The data obtained at Beijing expressways by us reveal the same characteristics. The fluctuation is regarded as the waves consisting of a series of overlapped simple harmonic oscillation. We can use some periodic functions like *sine*, *cosine*, or both of them to fit the field data. In order to simplify the calculation, the *sine* functions are adopted to fit the fluctuation characteristics of the traffic flow parameters (see (1)). The Fast Fourier

Transform (FFT) method is used to transfer the data of traffic flow parameters (data source from Figures 2–4) from time domain to frequency domain. Then, the simple harmonic oscillation with the largest amplitudes is folded to fit the field data at the object of the minimized relative error. Consider

$$f(t) = a + \sum_{i=1}^n A_i \sin(\omega_i t + \varphi_i), \quad (1)$$

where a is the fluctuation mean value of the traffic flow parameters, A_i is the amplitude, ω_i is the angular velocity, φ_i is the prime phase angle, t is the interval number, and n is the number of sine functions.

Table 1 shows the fitting function values (the amplitudes, angular velocities, and phase angles) of traffic flow parameters obtained from the middle lane of the merge area upstream at Site 2 in congestion state.

From Table 1, with the same angular velocity, satisfactory functions are obtained to describe the fluctuations of traffic volume, speed, and time occupancy. The research shows that there are similar fluctuation rules on the median lane and shoulder lane. Except for the merge area, the same fluctuation facts are found at the other locations. Therefore, the conclusion drawn is that the congested traffic flow parameters (traffic volume, speed, and time occupancy) fluctuate periodically with time, and the periodical fluctuations of these parameters are intrinsic property of urban expressway system.

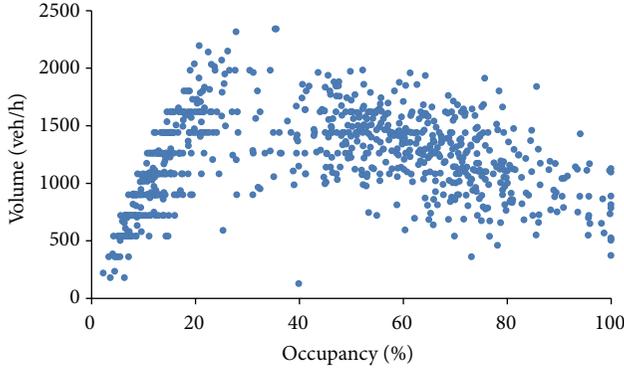


FIGURE 9: Occupancy-volume plot at Site 1.

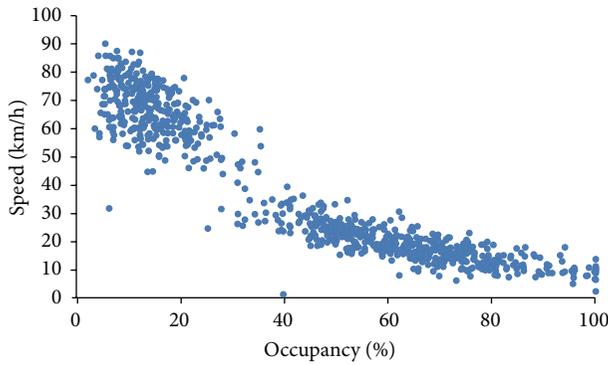


FIGURE 10: Occupancy-speed plot at Site 1.

3. Critical Time Identification

The critical time dividing traffic states is of importance to describe the bottleneck characteristic and is of great value for the traffic control system of urban expressways. Two methods are proposed to define it: the graph method and the control line method.

3.1. Graph Method. According to the traffic flow theory, traffic volume changes with time in accordance with time occupancy when traffic flow is in noncongestion [19, 22]. The accumulative vehicle number curve and the accumulative time occupancy curve can be properly adjusted to make them overlap in the same coordinate plane. But, the time occupancy ascends quickly and the traffic volume descends when traffic flow changes from noncongestion to congestion. Therefore, the slope of accumulative time occupancy curve increases, and that of accumulative vehicle number curve decreases. So, there must be a point where the two adjusted curves start to furcate and the trajectories go in different directions. Then, the time corresponding to the furcated point is the critical time in traffic state transition. During traffic flow resuming normal state, the two adjusted curves gradually come close before they merge. The time corresponding to the joining point is the critical time when traffic state changes from congestion to normal.

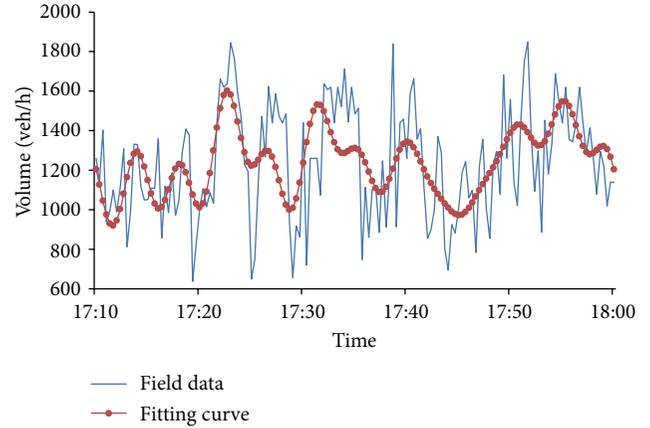


FIGURE 11: Volume fluctuation characteristic on middle lane at Site 2.

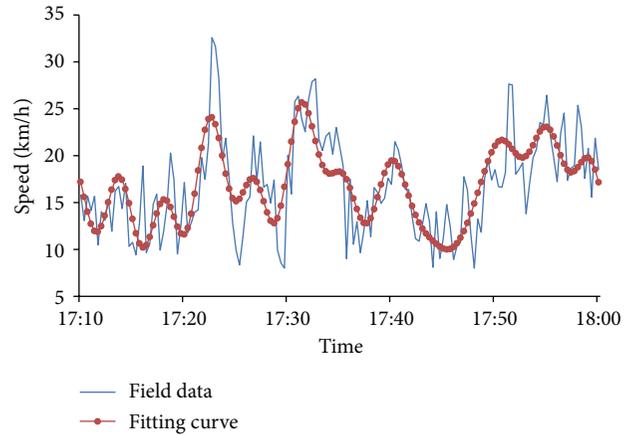


FIGURE 12: Speed fluctuation characteristic on middle lane at Site 2.

TABLE 1: The coefficient value of fitting functions on traffic flow parameters in congestion.

	1	2	3	4	5	6
Volume						
A_i	0.66	0.57	0.35	0.60	0.47	0.42
ω_i	0.08	0.13	0.21	0.25	0.46	0.50
φ_i	2.96	3.67	-0.23	3.56	2.56	1.21
a	6.88		MRE		10.38%	
Speed						
A_i	0.95	0.53	0.59	0.61	0.57	0.39
ω_i	0.08	0.13	0.21	0.25	0.46	0.50
φ_i	2.77	3.75	0.21	3.61	2.74	1.21
a	4.70		MRE		12.09%	
Occupancy						
A_i	0.08	0.03	0.03	0.03	0.04	0.03
ω_i	0.08	0.13	0.21	0.25	0.46	0.50
φ_i	-0.34	0.53	3.46	0.36	-0.21	4.54
a	0.68		MRE		7.52%	

The method to affirm the critical time of traffic state is to adjust the curves of cumulative vehicle arrival numbers

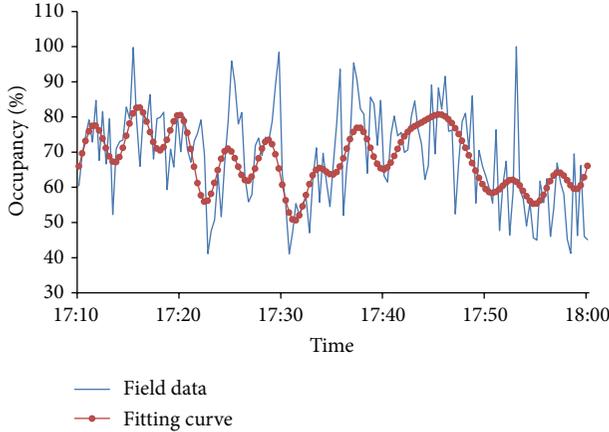


FIGURE 13: Occupancy fluctuation characteristic on middle lane at Site 2.

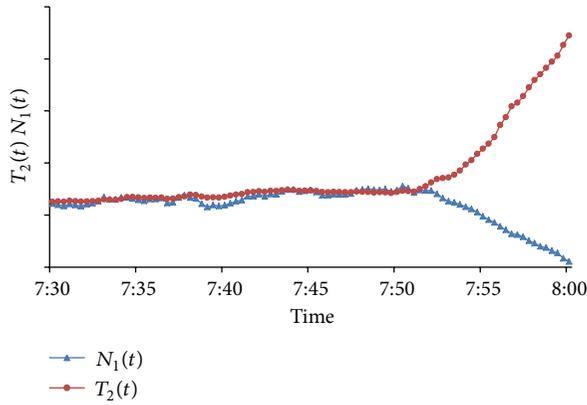


FIGURE 14: Critical time of traffic state change of median lane at Site 1.

and cumulative time occupancy. The adjusted accumulative vehicle number is

$$N_1(t) = N_0(t) - qt', \quad (2)$$

where $N_1(t)$ is the adjusted accumulative vehicle number at time t , $N_0(t)$ is the accumulative vehicle number at time t , q is a flow to be calibrated, and t' is the time starting from a point on the time axis when traffic is in noncongestion. Equation (2) is the change of $N_1(t)$ decided by the flow.

The adjusted accumulative time occupancy is

$$T_1(t) = T_0(t) - bt', \quad (3)$$

where $T_1(t)$ is the adjusted accumulative time occupancy at time t , $T_0(t)$ is the accumulative time occupancy at time t , and b is a coefficient. In (3), the change of $T_1(t)$ is decided by the coefficient b .

In order to make the two adjusted curves overlap in the same coordinate plane when traffic is in noncongestion, the adjusted curve of accumulative time occupancy is amplified by k times. Equation (3) is modified as

$$T_2(t) = k [T_0(t) - bt']. \quad (4)$$

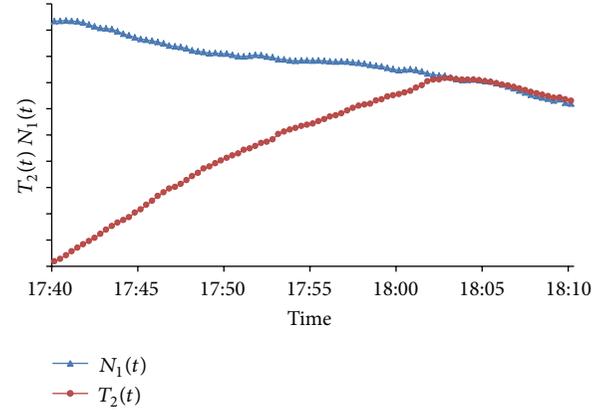


FIGURE 15: Critical time of traffic state on middle lane at Site 2.

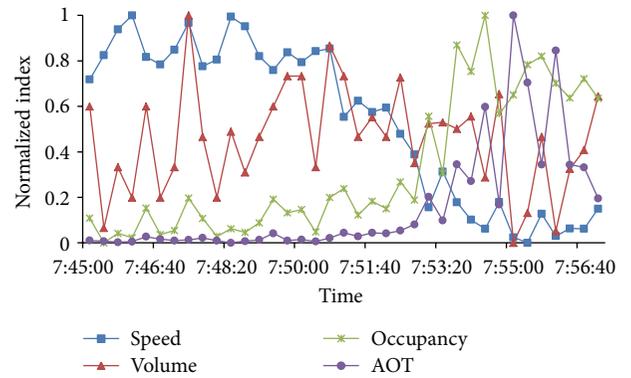


FIGURE 16: Normalization index on median lane at Site 1.

Based on the analysis above, (2) and (4) are applied to plot the two adjusted curves and decide the critical time corresponding to the furcate point or the joining point.

The procedures of affirming the critical time of traffic state are as follows:

- (1) confirm the approximate time of traffic state transition according to the curve of occupancy versus time;
- (2) select the original data about the approximate time;
- (3) confirm the coefficients q , b , and k on the basis of step 2;
- (4) plot the adjusted curves of accumulative vehicle number and accumulative time occupancy according to (2) and (4);
- (5) affirm the critical time of traffic state in terms of the furcate and joining points.

Here, the basic principle to select q , b , and k needs to be specified. When traffic flow changes from normal to congestion, the mean values of traffic volume and time occupancy in the normal situation are defined as q , b , and k are between 10 and 20 in general. When traffic flow changes from congestion to normal, q and k are identical to the above results, but b equals 0.3 [23–25].

The graph method is applied to affirm the critical time by the survey data (7:30~8:00) on the median lane of

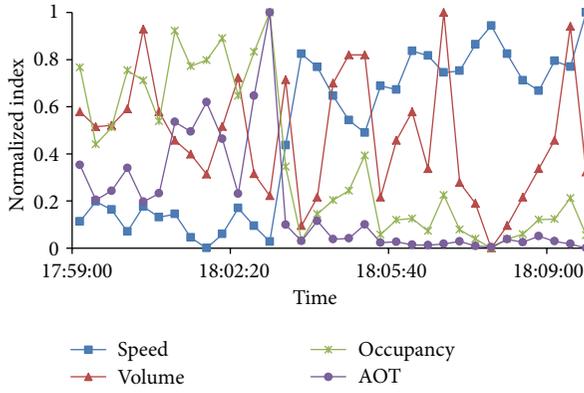


FIGURE 17: Normalization index on shoulder lane at Site 2.

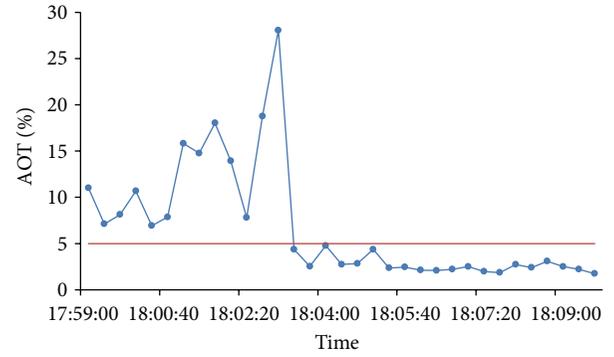


FIGURE 19: Management plots on shoulder lane at Site 2.

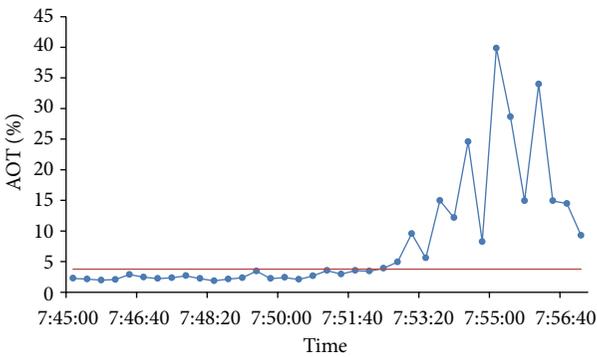


FIGURE 18: Management plots on median lane at Site 1.

the merge area downstream at Site 1. Here, q equals 1455 (veh/h), b equals 0.15, and k equals 15. The adjusted curves of accumulative time occupancy and accumulative vehicle number are shown in Figure 14. The furcate time of two curves is about 7:51:20, which becomes the turning point, where traffic flow changes from normal to congestion, that is, traffic bottleneck is forming. The same work is done to affirm the critical time by the survey data (17:40~18:20) on the middle lane of the merge area upstream at Site 2. Here, q equals 1530 (veh/h), b equals 0.3, and k equals 15. The adjusted curves of accumulative time occupancy and accumulative vehicle number are shown in Figure 15. The joining time of the curves is at 18:03:00. From this point, traffic flow changes from congestion to normal, that is, traffic bottleneck has dissipated. The analysis result accords with video observation.

3.2. Control Line Method. According to the present research, it is inaccurate to use a single traffic flow parameter to determine the state transition at expressways. A compound index, average occupancy time (AOT) is designed to affirm the critical time of traffic state. The expression of index AOT is

$$\text{AOT} = \frac{\sum t_i}{q}. \quad (5)$$

Here, t_i is the time of number i vehicle passing through the detectors, and q is traffic volume.

It is well known that the detector offers time occupancy parameter rather than the parameter $\sum t_i$ [26]. But time occupancy is equal to $\sum t_i$ divided by T ; that is, $o = \sum t_i/T$. Therefore, the index AOT becomes AOT which equals to time occupancy divided by traffic volume; that is, $\text{AOT} = o/q$.

Figures 16 and 17 show the curves of speed versus time, volume versus time, occupancy versus time, and AOT versus time after normalization processing at different site respectively. The index AOT is more stable than the single index (such as time occupancy, and speed) in normal situation, and it has an obvious change when traffic state changes from one to another. So the index AOT adapts to describe the transition of traffic state.

According to the quality management principle, traffic is regarded as a production process following normal distribution. When traffic state is normal, it can be regarded as a normal state, otherwise an abnormal state. By calculating AOT, a quality management diagram can be drawn (AOT- t) to determine the transition of traffic state. If the AOT data falls in the range of $(\mu \pm 3\sigma)$, it indicates that traffic state is of noncongestion. If three consecutive data are more than the upper boundary $(\mu + 3\sigma)$, traffic state will change to congestion at the time. On the contrary, when traffic state is firstly in congestion, if three consecutive data points were lowered to the upper boundary $(\mu + 3\sigma)$, then this indicates that traffic state returns to noncongestion.

The detailed procedure to affirm the transition of traffic state is listed as follows.

- (1) Select data in normal situation to confirm mean value μ and square error σ of AOT, and the data sample number exceeds 45 if adopting 20 s as time interval because the observed period is not less than 15 minutes in general.
- (2) Calculate the boundary value $(\mu \pm 3\sigma)$ of AOT according to step 1. Here, the lower boundary $(\mu - 3\sigma)$ of AOT is not to be considered because all traffic states below the boundary $(\mu + 3\sigma)$ are regarded as normal state.
- (3) Determine traffic state by analyzing the index value of AOT continuously. If three consecutive data points exceed the upper boundary $(\mu + 3\sigma)$, traffic state will change from normal to congestion at the time.

If three consecutive data points are lower to the upper boundary ($\mu + 3\sigma$), traffic flow returns to noncongestion.

Taking, for example, the field data on the median lane of the merge area downstream at Site 1 and on the shoulder lane of the merge area upstream at Site 2, then the control line method is validated (see Figures 18 and 19).

In Figure 18, 50 points in normal situation are chosen to get μ and σ . If μ is 0.0276 and σ is 0.0036, then $\mu + 3\sigma$ is 0.0384. The selected time period is 7:47~7:58. The transition time from normal to congestion is 7:51 by applying the control line method. In Figure 19, 50 points in normal situation are selected to get μ and σ . If μ is 0.0256 and σ is 0.0080, then $\mu + 3\sigma$ is equal to 0.0496. The selected time period is 17:59~18:12. The time of traffic flow changing from congestion to normal is 18:03 by applying the control line method. The outcome accords with the video observation and the graph method.

4. Conclusions

Based on the survey data of the section chosen from Beijing expressways, the time dependent characteristics of the traffic flow parameters are analyzed and the data gap is found. Two methods are proposed: the graph method and the control line method in order to affirm the critical time of traffic state when traffic flow changes from congestion to normal situation (or vice versa). The research outcome shows that the two methods are feasible in accordance with the result of video observation. The findings of the paper have been applied in the control plan of Beijing expressways. Future research will focus on using the field data from different cities to calibration and validation of the proposed models.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (nos. 51278454, 51208462, and 61304191).

References

- [1] X. Qu, Y. Kuang, E. Oh, and S. Jin, "Safety evaluation for expressways: a comparative study for macroscopic and microscopic indicators," *Traffic Injury Prevention*, 2013.
- [2] Q. Meng, X. Qu, K. T. Yong, and Y. H. Wong, "QRA model-based risk impact analysis of traffic flow in urban road tunnels," *Risk Analysis*, vol. 31, no. 12, pp. 1872–1882, 2011.
- [3] W. Wuhong, B. Klaus, and W. Geert, "Discrete dynamics in transportation system," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 234970, 2 pages, 2013.
- [4] R. L. Bertini and M. J. Cassidy, "Some observed queue discharge features at a freeway bottleneck downstream of a merge," *Transportation Research A*, vol. 36, no. 8, pp. 683–697, 2002.
- [5] J. C. Muñoz and C. F. Daganzo, "The bottleneck mechanism of a freeway diverge," *Transportation Research A*, vol. 36, no. 6, pp. 483–505, 2002.
- [6] S. Das and D. Levinson, "Queuing and statistical analysis of freeway bottleneck formation," *Journal of Transportation Engineering*, vol. 130, no. 6, pp. 787–795, 2004.
- [7] K. S. Ogut and J. H. Banks, "Stability of freeway bottleneck flow phenomena," *Transportation Research Record*, no. 1934, pp. 108–115, 2005.
- [8] Y.-C. Hu and P. Schonfeld, "Optimizing traffic diversion around bottlenecks," *Transportation Research Record*, pp. 22–27, 1984.
- [9] Z. Liu, Q. Meng, and S. Wang, "Speed-based toll design for cordon-based congestion pricing scheme," *Transportation Research C*, vol. 31, pp. 83–98, 2013.
- [10] Z. Liu and Q. Meng, "Modelling transit-based park-and-ride services on a multimodal network with congestion pricing schemes," *International Journal of Systems Science*, 2012.
- [11] Z. Liu, Y. Yan, X. Qu, and Y. Zhang, "Bus stop-skipping scheme with random travel time," *Transportation Research C*, vol. 35, pp. 46–56, 2013.
- [12] L. C. Edie, "Car-following and steady-state theory for noncongested traffic," *Operations Research*, vol. 9, pp. 66–76, 1961.
- [13] J. S. Drake et al., "Statistical analysis of density hypotheses," *Highway Transportation Research*, vol. 154, pp. 53–87, 1967.
- [14] A. Ceder, "A deterministic traffic flow model for the two-regime approach," *Transportation Research Record*, no. 567, pp. 16–30, 1976.
- [15] J. C. Muñoz and C. F. Daganzo, "Structure of the transition zone behind freeway queues," PATH Working Paper UCB-ITS-PWP-2000-24, Institute of Transportation Studies, University of California, Berkeley, Calif, USA, 2000.
- [16] M. Koshi, M. Iwasaki, and I. Ohkura, "Some findings and an overview on vehicular flow characteristics," *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, pp. 403–426, 1983.
- [17] Q. Meng and X. Qu, "Estimation of rear-end vehicle crash frequencies in urban road tunnels," *Accident Analysis and Prevention*, vol. 48, pp. 254–263, 2012.
- [18] S. Jin, X. Qu, and D. Wang, "Assessment of expressway traffic safety using gaussian mixture model based on time to collision," *International Journal of Computational Intelligence Systems*, vol. 4, pp. 1122–1130, 2011.
- [19] H. Nathan, *Monograph on Traffic Flow Theory*, The Federal Highway Administration (FHWA), 1996.
- [20] B. S. Kerner, *The Physics of Traffic*, Springer, 2004.
- [21] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [22] Q. Meng and X. Qu, "Uncertainty propagation in quantitative risk assessment modeling for fire in road tunnels," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 42, pp. 1454–1464, 2012.
- [23] Transportation Research Board, *Highway Capacity Manual*, National Research Council, 2002.
- [24] M. J. Cassidy and R. L. Bertini, "Some traffic features at freeway bottlenecks," *Transportation Research B*, vol. 33B, no. 1, pp. 25–42, 1999.
- [25] R. M. Shanteau, "Using cumulative curves to measure saturation flow and lost time," *Institute of Transportation Engineers Journal*, vol. 58, no. 10, pp. 27–31, 1988.
- [26] A. D. May, *Traffic Flow Fundamentals*, Library of Congress Cataloging in Publication Data, 1990.

Research Article

The Departure Characteristics of Traffic Flow at the Signalized Intersection

Zhaowei Qu, Yan Xing, Hongyu Hu, Yuzhou Duan, Xianmin Song, TingTing Chai, and Hu Zhang

College of Transportation, Jilin University, Changchun 130025, China

Correspondence should be addressed to Hongyu Hu; huhongyu@jlu.edu.cn

Received 5 July 2013; Revised 22 August 2013; Accepted 25 August 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Zhaowei Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The motion characteristics of the leading vehicle and the following vehicles of the traffic flow at the typical urban intersections are qualitatively analyzed through the kinematical equation and the traffic wave theory. Then, the motion characteristic of the whole traffic flow during the dispersion process is also studied. Based on the spatiotemporal model of kinematics in the departure process and traffic wave model in the dispersion process proposed, the change of the leading vehicle of the departure process and the time of the following vehicles reaching to the stable speed as well as the relationship between the green time and the departure vehicle number at the intersection are acquired. Furthermore, according to the qualitative analysis and the quantitative calculation of the departure traffic flow at the signalized intersection, the dispersion characteristic of traffic flow at the signalized intersection was studied and analyzed, which provides reliable theoretical basis for traffic signal setting at the intersection.

1. Introduction

Study on the departure characteristic of traffic flow at the intersection is not only one of the important contents in research on the traffic theories at urban road but also directly affects the setting of signal timing program at the intersection and the optimization of traffic flow organization program. Moreover, it is closely related with the study on the traffic capacity of the intersection. Therefore, profound analysis on the traffic flow characteristics of starting traffic flow, dispersion traffic flow, and stably departure traffic flow in the departure process of traffic flow at the intersection are helpful to improve the level of service at intersection and relieve the traffic jam at the intersection. Based on this, many experts and scholars have conducted extensive researches on the departure characteristics of traffic flow at the intersection and achieved fruitful research results. For instance, for the micro characteristics of departure vehicles at the intersection, Gazis et al. [1, 2] first proposed stimulus-response model that sensitivity is inversely proportional to time headway on the basis of car-following theory analysis then conducted further researches and deduced the car-following GHR model based on nonlinear algorithm. Through analyzing the micro

behavior of traffic flow moving, Yang et al. [3–7] put forward the concept of random degree of free flow to measure the interactive strength of vehicles in the traffic flow and obtained the traffic dispersion model suitable for the urban roads. From the aspects of drivers' psychological field features and visual attention features, Jin and Tao [8, 9] analyzed the influence of leading vehicles on the following vehicles in the process of vehicle moving then studied and improved the car-following model for the motion characteristics of following vehicles, respectively. Stokes et al. [10, 11] studied the departure characteristics of left-turn lanes at the intersection, obtained departure saturation flow of left-turn lanes by calculating the saturation time headway, and discovered that there are certain differences in various left-turn lanes. For macro characteristics of departure vehicles at the intersection, May et al. [12, 13] first transformed the traditional research methods, described the formation and dispersion of shock wave at the signalized intersection and bottleneck of pedestrian crosswalk by utilizing graphical method, and then analyzed the signalized traffic flow characteristics and its dispersion process. Bando et al. [14, 15] studied the overall features of traffic flow from the macro angle, then described the

macro features of traffic flow like the unstable phenomenon of traffic flow and blocking process, and obtained the mean-field equation for the relationship between the average speed of the stable traffic flow and the vehicle density. Through analyzing the kinematical characteristics of traffic starting wave at the intersection, Wang et al. [16–18] defined the significance of kinematics parameter of traffic flow in the traffic wave and built the kinematical model of traffic starting wave at the intersection. In addition, Guo et al. [19, 20] analyzed the dispersion process of traffic flow with graphical method, calibrated the moving speed, motion trajectory, and dispersion time in the departure process of traffic flow on the time-space rectangular coordinate system, and determined the queue length and overall dispersion time of vehicles at the intersection.

To a certain extent, the above research results improve the development of the study on the dispersion characteristics of traffic flow at the intersection. However, it needs further discussion on two aspects. First, the time from the beginning of departure of the traffic flow to stable moving should be measured, as well as the speed of stable moving and the traffic volumes from the departure to passing the stop lines stably. The relationship between the queue length and the platoon dispersion time was analyzed in the above researches, as well as the total time passed through the intersection and passing speed, although it has few discussions to the different stages of departure traffic flow in the intersection, which is difficult to offer accurate based data for signal timing at the intersection and the capacity optimization. Second, the relationships among the spreading speed of the whole traffic wave which produced by the arrived vehicles and the queued vehicles, green time, and the number of the passing vehicles during the queued traffic flow departing are analyzed. The results presented by the former researchers were only concentrated on the departure process of queued vehicles at the intersection. For the previous studies focus on the release of queuing traffic flow at the intersection, there are few studies on some situations, such as whether the follow-up vehicles could pass through the intersection in the queuing period, the traffic wave propagation situation after the follow-up flow overtakes the queuing flow, and the relationship between the follow-up vehicles and the green time at the intersection and so on. The departure and motion situations of the traffic flow arrived subsequently are closely related to the passing efficiency and signal timing at the intersection. As a result, it needs further research on these aspects.

By summarizing the above research experience, aiming at the traffic flow at the urban typical intersections, this paper applies kinematical equation and traffic wave theory to conduct qualitative analysis for the motion characteristic of the leading vehicle, the motion characteristic of the following vehicles, and the motion characteristic of the whole traffic flow at the intersection during the dispersion process at the signalized intersection. It further builds spatiotemporal model of kinematics in the departure process at the intersection and traffic wave model in the dispersion process at the intersection to determine the changing situations of the leading vehicle at the departure process, the time for the following vehicles to reach the stable speed, and the relationship

between the green time at the intersection and the departure vehicle number at the intersection. In addition, in this paper we analyzed the traffic flow departure situations combined with the vehicle operation data at the signalized intersection. Through qualitative analysis and quantitative calculation, we analyzed the dispersion characteristic of traffic flow at the signalized intersection and provided reliable theoretical basis for the setting of traffic signals at the intersection.

2. Model Building

When the traffic lights in the intersection turn green, the queued vehicles start to move. When the space headway between the leading vehicle and the second one is significantly greater than the average space headway of the queued vehicles, it indicates that the traffic flow is beginning to disperse. As the leading vehicle of the queue can accelerate continuously to the expected speed without the leading vehicle ahead, it can easily extend its space headway to the second vehicle. Moreover, the second vehicle will be stimulated to speed up to widen the space headway to the third vehicle. Accordingly, if the conditions are met, the change of the motion state will transfer to the following vehicles continuously so that the whole queued vehicles will participate in the dispersion process. If the green light time is long enough, the transfer process will last until all queued vehicles reach their expected speed by the stimulation of the leading vehicles ahead. Till then, the whole dispersion process of the queued vehicles at the intersection has been finished.

Based on the practical dispersion conditions in the intersection, the paper divides the vehicle motion characteristics during the whole dispersion process into three parts: the motion characteristic of the leading vehicle at the intersection during the departure process, the motion characteristic of the following vehicles at the intersection during the departure process, and the motion characteristic of the whole traffic flow at the intersection during the departure process. Moreover, we build a qualitative description model and make quantitative calculation analysis for these three parts, respectively, so as to analyze the departure characteristics of the traffic flow at urban road intersections accurately and thoroughly.

2.1. Assumed Conditions. To simplify the research environment, the following assumptions are proposed.

- (1) There is no bus lane at the intersection or bus stop at the influence area.
- (2) There is no new queue formed at the intersection.
- (3) Take one single lane at the intersection as the research object.
- (4) Ignore the individual differences of vehicles and disturbance of pedestrians and nonmotor vehicles.
- (5) Take the traffic characteristics of straight-moving traffic flow at the intersection with four-phase traffic signal as the analysis object.

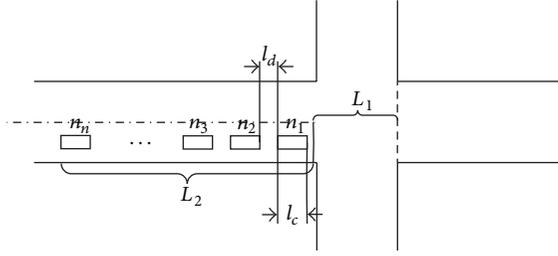


FIGURE 1: The traffic flow diagram at the intersection.

2.2. Solution Parameters. Based on kinematical equation and traffic wave theory, this paper analyzes the motion characteristics of various traffic flow stages at the departure process in the intersection and solves the following traffic parameters:

- (1) the time from the beginning of departure of the queued vehicles to stable moving,
- (2) the departure vehicle number from the beginning of departure of the queued vehicles to stable moving,
- (3) the time headway when the queued vehicles accelerate and stably move,
- (4) the total time needed in dispersion the queued vehicles,
- (5) the spreading speed of the traffic wave after the confluence of the vehicles arrived subsequently and the queued vehicles,
- (6) the relationship between green light time at the intersection and the number of the departure vehicles at different situations.

2.3. Descriptive Analysis on the Motion Characteristic of the Leading Vehicle. Different from other drivers who receive stimulation from the vehicles ahead, the drivers of the leading vehicles in the departure process receive stimulation mainly from the signal lights at the intersection. Therefore, this paper conducts independent study for them, builds a relationship model of moving speed-time-distance, determines the features of accelerating time, accelerating distance and the corresponding spatial positions at different time of the leading vehicle, and provides theoretical basis for the conflict analysis and the setting of red light time at the intersection.

(1) Qualitative Description. This paper takes the typical intersection as the example for analysis, defines the intersection length is L_1 , the queued vehicle length at the intersection is L_2 , and the vehicle distance of the queued vehicles is l_d , ignores the individual performance and size difference of vehicles, and takes the vehicle length as l_c . For details, please refer to Figure 1.

As the starting of the leading vehicle at the intersection is mainly influenced by the intersection signals, this paper considers the acceleration of the leading vehicle at the intersection as a , the stable speed of other vehicles as v_f , the starting time of other vehicles as t_{i0} , the initial position

of various vehicles as x_i , the time for various vehicles to reach stop line as t_{i1} , the time for various vehicles to pass intersection as t_{i2} , and the time for various vehicles to achieve stable speed as t_{if} . In addition, it assumes the real position of other vehicles as x'_i , the corresponding time of a vehicle at the real positions as t'_1 , the real speed of the vehicle moving as v'_1 , and the real distance of the Vehicle i as l_i . According to the kinematical equation, the relations of moving distance-time and speed-time of the leading vehicle at the intersection can be expressed in

$$l_i = x'_1 - x_1 = \int_{t_{i0}}^{t'_1} \int_{t_{i0}}^{t'_1} a dt dt, \quad (1)$$

$$v'_1 = \int_{t_{i0}}^{t'_1} a dt.$$

As the leading vehicle has no obstructive factor ahead after being released, it can speed up till the stable speed for free travel. Hence, the accelerating time t_{1f} of the leading vehicle after being released and the distance covered during the accelerating process can be obtained as

$$v_{1f} = \int_{t_{i0}}^{t_{1f}} a dt, \quad (2)$$

$$s_1 = \int_{t_{i0}}^{t_{1f}} \int_{t_{i0}}^{t_{1f}} a dt dt + v_{1f} (t'_1 - t_{1f}).$$

If $s_1 < L_1$, it means that the leading vehicle reaches stable speed before passing the intersection; if $s_1 > L_1$, it means that the leading vehicle experiences the acceleration process during the whole motion process in the intersection; and if $s_1 = L_1$, it means that the speed of leading vehicle passing the intersection is exactly the stable speed of v_f , and the time for the leading vehicle to pass the intersection is the time needed for it to accelerate to the stable speed.

(2) Quantitative Analysis. This paper then selects a typical intersection in Changchun and collects 16 groups of the moving data of the leading vehicles at the intersection by video, the collected data mainly includes the green light time, the starting time of the leading vehicle, and the time when the leading vehicle passes the mark points in the intersection. Then using these data, the parameters like responsive starting time, the speed variety, and the moving time of the leading vehicle in the intersection can be analyzed. Hereby the intersection length is 40 m. For details, please refer to Table 1.

Through analyzing the data in Table 1, we can get Figures 2, 3, and 4, from which we can find that the response time of various leading vehicles in the traffic flow at the intersection is quite different, lacking certain regularity. This is mainly caused by the different performances of vehicles and drivers. However, the moving speed and the moving time of the leading vehicle in the intersection are relatively stable. This is mainly because the leading vehicles are always at the accelerating stage after being released, and they have no interference from other vehicles. From the data statistics, it

TABLE 1: Moving data of the leading vehicle at the intersection.

Group	Green time	Starting time of the leading vehicle	Mark point 1	Mark point 2	Mark point 3	Mark point 4
Leading vehicle 1	21.40	22.44	26.24	27.72	28.92	30.04
Leading vehicle 2	21.40	22.92	26.72	28.72	30.20	31.44
Leading vehicle 3	21.40	24.00	28.20	29.92	31.32	32.60
Leading vehicle 4	21.40	23.40	28.04	29.80	31.32	32.44
Leading vehicle 5	19.72	21.84	26.60	28.00	29.28	30.32
Leading vehicle 6	19.72	23.20	27.20	29.04	30.32	31.44
Leading vehicle 7	19.72	23.88	28.64	30.24	31.64	32.72
Leading vehicle 8	19.72	22.28	27.32	29.16	30.76	32.08
Leading vehicle 9	24.44	25.08	29.60	31.20	32.48	33.52
Leading vehicle 10	24.44	25.60	30.00	31.84	33.20	34.40
Leading vehicle 11	24.44	27.04	33.20	35.80	37.48	39.08
Leading vehicle 12	24.44	26.20	31.00	32.60	34.12	35.32
Leading vehicle 13	26.00	26.72	30.16	31.40	32.56	33.40
Leading vehicle 14	26.00	27.64	32.56	34.28	35.64	36.76
Leading vehicle 15	26.00	27.56	33.24	35.08	36.52	37.96
Leading vehicle 16	26.00	29.56	33.40	34.84	36.24	37.28

Note: the mark points 1, 2, and 3 are random marks and the mark point 4 is the position of passing the intersection.

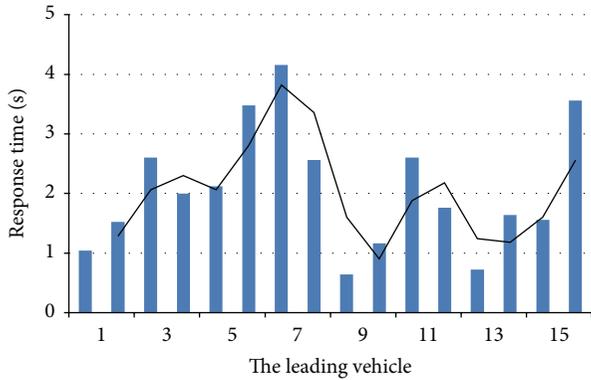


FIGURE 2: Responsive starting time variety of the leading vehicle at the intersection.

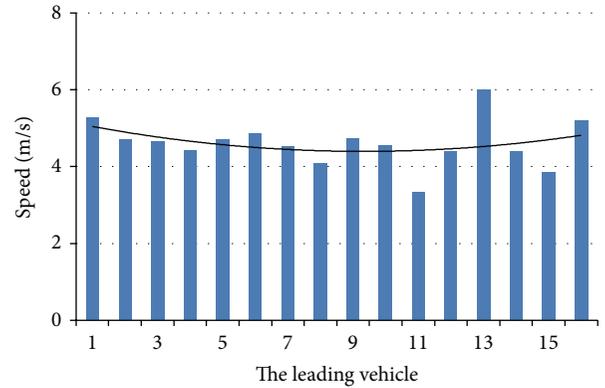


FIGURE 3: Speed variety of the leading vehicle in the intersection.

can be obtained that the average responsive starting time of the leading vehicle is 2.07 s, the average moving speed in the intersection is 4.6 m/s, and the average moving time of the leading vehicle in the intersection is 8.84 s. In addition, the leading vehicle in the intersection is always in the accelerating state, and the average accelerating speed of the investigated vehicles in the intersection is 0.52 m/s^2 . Combined with the position-time-speed of the leading vehicle in the video data as well as (1)~(2), it can be concluded that the Spatiotemporal trajectory of the moving speed of the leading vehicle after being released at the intersection is as shown in Figure 5.

2.4. Descriptive Analysis on the Motion Characteristics of following Vehicles. Generally speaking, the motion characteristics of the following vehicles largely depend on the leading vehicles. The secure time distance between the following vehicles and the leading vehicles determines the following-vehicle driver's sensibility degree for danger. The larger

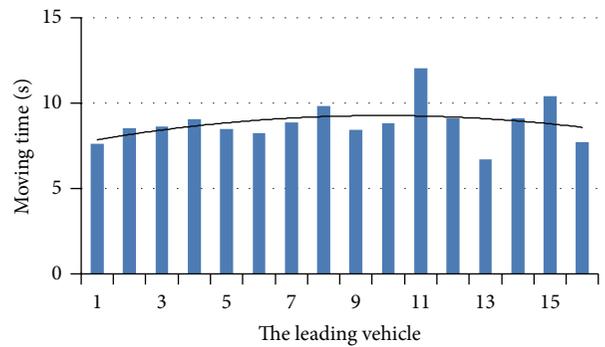


FIGURE 4: Moving time variety of the leading vehicle in the intersection.

the time distance between the vehicles is, the lower the drivers' perceivable risk degree is, and the following vehicles

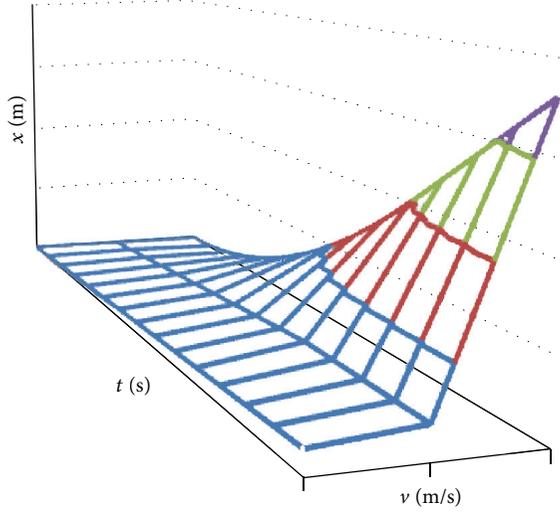


FIGURE 5: Spatiotemporal curve for the moving speed of the leading vehicle at the intersection.

can maintain the speed or accelerate to certain extent. Contrarily, the smaller the time distance between vehicles is, the higher the drivers' perceivable risk degree is, and the following vehicles should maintain the speed or decelerate to certain extent. Consequently, the speed variety (or acceleration) of the following vehicles is closely related to the influence degree of the leading vehicles on the following vehicles. Based on this relation, we can offer acceleration decision model of the following vehicles and establish car-following model and descriptive model for the motion characteristic of the following vehicles.

(1) *Qualitative Description.* In order to establish acceleration decision model of the following vehicles, we firstly need to confirm the functional relations between the acceleration of the following vehicles and their drivers' influence degree from the leading vehicles. Meanwhile, the influence degree of the leading vehicles on the following vehicles' drivers is closely related to the time distance between the following vehicles and leading vehicles. Based on such analysis, the paper primarily establishes the relations between influence degree of the leading vehicles on the following vehicles and the time distance between vehicles. Moreover, through confirming the influence degree of the leading vehicles on following vehicles, we provide the acceleration of the following vehicles whose concrete transformation models are showed as

$$f(h_{ti}) = h_0 - \frac{h_{ti}}{h_0} - h_{\min}, \quad (3)$$

$$\text{IF} = 1 - \sum_{i=1}^n \frac{f(h_{ti})}{n}, \quad (4)$$

$$a_i = \omega * f(\text{IF}). \quad (5)$$

In the equation, a_i is the acceleration of Vehicle i ; $f(\cdot)$ is the transfer function of the influence degree of the leading vehicles on the following ones and the acceleration; IF is the

influence degree of the leading vehicles on the following ones; R is the degree of randomness for the traffic flow; n is the number of the vehicles; $f(h_t)$ is the influence function of the vehicles; h_t is the time distance of the vehicles; h_{\min} is the least necessary time distance for the vehicles.

When $i = 2$, IF is the influence degree between the following and leading vehicles; when $h \geq h_0$, $f(h) = 0$; when $h \leq h_{\min}$, $f(h) = 1$; or when $\text{IF} = 1$, the influence degree reaches the minimum; when $\text{IF} = 0$, it means that the time distance reaches its minimum while the influence degree reaches the maximum.

Then the established acceleration decision model for the following vehicles can be used to describe the motion characteristics of the following vehicles. First, it is supposed that the response time of the queued drivers is t_f , the green light time is t_0 , the real-time moving speed is v_i before the queued vehicles reach the stable speed, and the time required for the following vehicles to reach the stable speed v_f with acceleration a_i is t_f . If the following vehicles have not reached the stable speed before reaching the stop line the relation of the time t_{i1} and the speed v_{i1} when the following vehicles arrive at the intersection stop line can be expressed by (6), while the relation of the distance l_{i1} between the following Vehicle i and the stop line, and the time t_{i1} when the vehicle reaches the stop line can be expressed by (7), thus we can get the length of time of the following car from start time to reaching the stop line time by (8):

$$l_{i1} = (i - 1)(l_c + l_d), \quad (6)$$

$$v_{i1} = \int_{t_0+it_r}^{t_{i1}} \omega * f\left(1 - h_0 - h_{ti} + \frac{h_{t(i-1)}}{2}\right) dt, \quad (7)$$

$$t_{i1} - t_{i0} = \sqrt{\frac{2(i-1)(l_d + l_c)}{\omega} * f\left(1 - h_0 - h_{ti} + \frac{h_{t(i-1)}}{2}\right)}. \quad (8)$$

Meanwhile, the containable vehicles n and the density k of the traffic flow in the intersection can be worked out so as to calculate the traffic wave spreading situations when the traffic flow is departure. First, a relation between the moving time and distance of the following vehicles needs to be built. If the following Vehicle i has not reached the stable speed when moving out of the intersection, the expression can be shown as follows

$$x_{i2} - x_{i0} = \int_{t_0+it_r}^{t_{i2}} \int_{t_0+it_r}^{t_{i2}} \omega * f\left(1 - h_0 - h_{ti} + \frac{h_{t(i-1)}}{2}\right) dt dt. \quad (9)$$

In the equation, x_{i2} is the position when a vehicle passes the intersection.

If the vehicle has reached the stable speed when passing the intersection, the equation can be described as

$$x_{i2} - x_{i0} = \int_{t_0+it_r}^{t_f} \int_{t_0+it_r}^{t_f} \omega * f\left(1 - h_0 - h_{ti} + \frac{h_{t(i-1)}}{2}\right) dt dt + v_f(t_{i2} - t_f). \quad (10)$$

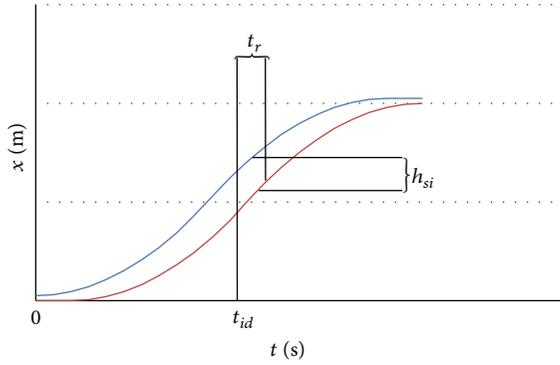


FIGURE 6: The relation diagram of motion trajectory between the following and the leading vehicles.

The secure space between any two vehicles during the departure process can be worked out, on the basis of the exact position and speed of all the vehicles in different time. If the vehicle is accelerating, the calculating model of secure space between the following and the leading vehicles can be shown as (11), while the motion trajectory of the decelerating process (to zero) of the two vehicles is shown in Figure 6:

$$h_{si} = \int_{t_{(i-1)0}}^{t_{id}+t_r} \int_{t_{(i-1)0}}^{t_{id}+t_r} \omega * f \left(1 - h_0 - h_{ti} + \frac{h_{t(i-1)}}{2} \right) dt dt - \int_{t_{i0}}^{t_{id}} \int_{t_{i0}}^{t_{id}} a dt dt. \quad (11)$$

In Figure 6, t_{id} and $t_{id} + t_r$ are the time and position when the vehicles i and $i - 1$ start to decelerate.

If the neighboring vehicles are both moving at a stable speed, the secure space of the two vehicles is the distance covered by the following vehicle within the response time. The expression is as follows

$$h_{si} = v_f t_r. \quad (12)$$

Then, the time headway between the neighboring vehicles can be worked out; this paper analyzes the mentioned time headway h_{ti} , with l_c being the standard length of small-sized cars, summarizing (13) as follows:

$$h_{ti} = \frac{(l_c + h_{si})}{v_{i1}}. \quad (13)$$

Based on the above analysis, the paper can work out the number of the containable vehicles in the intersection n and the traffic flow density k , as shown in

$$n \leq \frac{(L_1 - \sum_{i=1}^n h_{si})}{l_c}, \quad (14)$$

$$k = L_1 - \sum_{i=1}^n \frac{h_{si}}{l_c L_1}. \quad (15)$$

(2) *Quantitative Analysis.* Based on the video-collected six-teen groups of the straight-moving traffic flow data at a typical

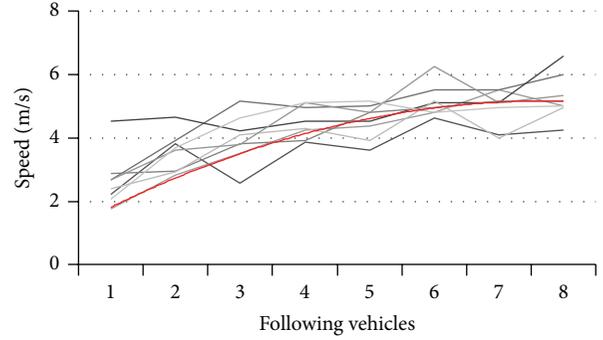


FIGURE 7: Speed variety of the following vehicles.

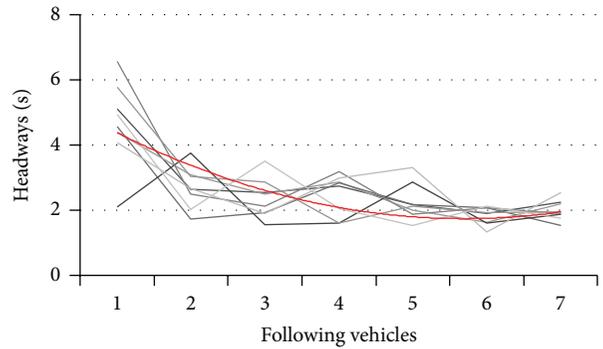


FIGURE 8: Headway variety of the following vehicles.

intersection as in shown in Section 2.3 quantitative analysis, the paper has chosen eight groups and marked the time when the vehicles reach the stop line and the relevant crossing speed. Then the paper analyzes the motion characteristics of the following vehicles at the intersection in the departure process, and the detailed data is shown in Table 2.

According to Figures 7 and 8, the following vehicles speed up gradually in the initial departure stage at the intersection, and it is until the sixth vehicle that the speed is leveling off. The time headway decreases gradually in the initial departure stage, and it is until the fifth vehicle that the time headway is leveling off. At this time, the stable average speed and time headway of the following vehicles are 5.28 m/s and 1.99 s, respectively. Based on the data above, the paper gives the qualitative description of the following vehicles. According to the earlier quantitative analysis, the average acceleration of the vehicles is $a = 0.52 \text{ m/s}^2$, the average response time of the drivers is $t_r = 1.5 \text{ s}$, the average length of the vehicle is 4.8 meters, and the average space of the queued vehicles is 2 meters. Hence, the paper concludes that the moving speed of the sixth vehicle arriving at the intersection stop line is 6.35 m/s, and the time headway when the traffic flow is moving steadily is 2.08 s. Through the comparison, the paper finds that the errors of calculated value and observed average value in speed and time headway are 20% and 4.5%, respectively. So the error values are acceptable, proving that (3) and (8) have a good imitative effect to the moving speed and the time headway of the following vehicles.

TABLE 2: Time and speed of the vehicles in each group when passing the stop line.

Group	Data of the vehicles in each group (v : m/s)							
	1	2	3	4	5	6	7	8
1								
t	22.07	27.17	29.8	32.33	35.07	37.23	39.13	41.37
v	2.22	3.81	2.57	3.87	3.61	4.62	4.10	4.25
2								
t	21.33	25.73	28.8	31.3	34.17	36.17	37.8	40
v	1.75	2.82	3.50	4.25	4.36	4.80	5.52	6.00
3								
t	25.5	30.07	31.8	33.73	36.57	38.73	40.8	42.33
v	2.67	3.90	5.16	4.95	5.00	5.52	5.52	6.00
4								
t	29.67	31.77	35.53	37.07	38.67	41.53	43.13	45
v	4.53	4.66	4.21	4.53	4.53	5.11	5.11	6.58
5								
t	31	37.57	40.07	42.2	45.37	47.23	49.3	51.2
v	2.87	2.94	3.81	3.90	4.80	4.95	5.16	5.16
6								
t	38.57	42.63	45.3	47.2	50.17	53.47	54.8	57.33
v	2.40	2.93	4.10	2.29	3.90	5.16	4.00	4.95
7								
t	40.40	46.17	49.20	52.06	53.67	55.80	57.73	59.67
v	2.67	3.61	3.78	5.11	4.80	6.23	5.11	5.33
8								
t	41.17	45.10	47.13	50.63	52.67	54.20	56.33	58.07
v	2.06	3.69	4.62	5.11	5.16	4.80	4.95	5.00

2.5. Analysis of the Departure Characteristics of Traffic Flow at the Intersection. When the traffic lights began to turn green, queued vehicles formed during red light and early green light period will be released at a relatively stable rate (saturation flow rate). The vehicles will accelerate to a stable speed, while the vehicles arrive after the vehicles ahead are dispersed will cross the intersection at a stable speed. Subsequently, vehicles will queue up again when the traffic lights are not green. The specific Spatiotemporal diagram of the departure vehicles at the intersection is shown in Figure 9.

According to the departure trajectory diagram of traffic flow at the intersection, the departure traffic flow at the intersection mainly includes starting and moving of the traffic flow, which are in correspondence with the start wave and kinematical wave in the traffic wave model. Hence, the traffic wave can be used to study the departure characteristics of the traffic flow at the intersection. Traditional wave model is mainly built on the basis of traffic flow conservation on a certain section within a fixed time. For example, given two neighboring areas with different densities A and B (the densities being k_A and k_B , resp.), and the density boundary is S , the spreading speed of the boundary with the density variety is u_S , and the average moving speeds of the traffic flow in two different areas are u_A and u_B . Figure 10 shows the specific results.

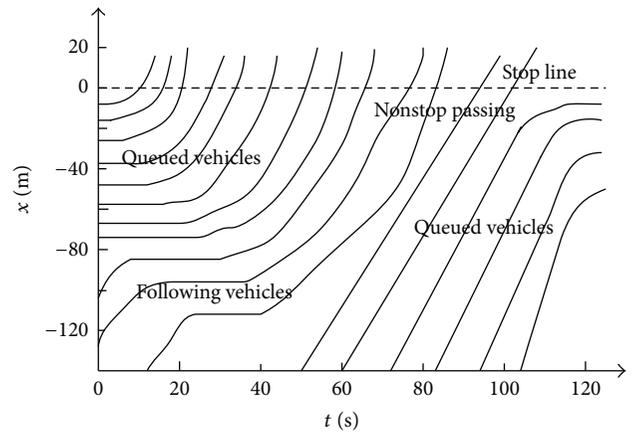


FIGURE 9: Departure trajectory of the traffic flow at the intersection.

According to Greenshield's model, the basic model of traffic flow, the speed of the traffic flow increases as the density decreases. Based on the conservative traffic flow N , (16) can be formed referring to Figure 10:

$$N = (u_A - u_S) k_A t = (u_B - u_S) k_B t. \quad (16)$$

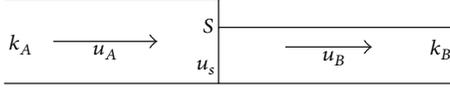


FIGURE 10: Fundamental diagram of the traffic wave.

For there is the relation between traffic flow, traffic density, and moving speed: $= ku$, which put into (16) we have

$$u_s = \frac{(q_A - q_B)}{(k_A - k_B)}. \quad (17)$$

Combined with flow-density-speed relation of Green-shield's model:

$$u = u_f \left(1 - \frac{k}{k_j} \right). \quad (18)$$

In the equation, u_f is free flow speed and k_j is the jam density. Putting (18) into (17), the wave speed expression can be obtained as follows.

Through analysis we can obtain

$$\left[u_f \left(1 - \frac{k_A}{k_j} \right) - u_s \right] k_A = \left[u_f \left(1 - \frac{k_B}{k_j} \right) - u_s \right] k_B. \quad (19)$$

Rearranging the equation one can get

$$u_s = \frac{u_f [k_A (1 - k_A/k_j) - k_B (1 - k_B/k_j)]}{(k_A - k_B)}. \quad (20)$$

When just being released, the traffic flow at the intersection is at the saturation flow rate. If calculating the traffic density of the area from L_3 in head of the intersection stop line, the following results with (15) can be obtained:

$$k_A = L_3 - \sum_{i=1}^n \frac{h_{si}}{l_c L_3}. \quad (21)$$

Meanwhile, when the traffic flow in area A is about to start, the traffic density in area B is approaching $k_B \approx 0$. Then the spreading speed of the traffic wave is as shown in

$$u_s = u_f \left(1 - \frac{L_3}{k_j} - \sum_{i=1}^n \frac{h_{si}}{l_c L_3 k_j} \right). \quad (22)$$

According to the traffic wave speed, the exact time when the traffic wave is approaching every vehicle in the traffic flow at the intersection, namely the starting time of the vehicles, can be calculated by using the following equation:

$$t_{(i+1)0} = t_{10} + \sum_{i=1}^{n-1} \frac{h_{si}}{u_s}. \quad (23)$$

Referring to (23) and kinematical equation, this paper can work out not only the time t_{if} when the vehicles reach a stable

speed in the departure process of traffic flow but also the exact position x_{if} when they reach the stable speed v_f , as shown in

$$t_{if} = t_{10} + \frac{v_f}{a} + \sum_{i=1}^{n-1} \frac{h_{si}}{u_s}, \quad (24)$$

$$x_{if} = x_i + \int_0^{v_f/a} \int_0^{v_f/a} a dt dt. \quad (25)$$

Combining (24) and (25), if the r th vehicle is the first vehicle to arrive at the intersection with a stable speed, the time t_{im} when the i th vehicle comes to the stop line will be as follows:

$$r = \frac{\int_0^{v_f/a} \int_0^{v_f/a} a dt dt}{h_{si}} + 1, \quad (26)$$

$$t_{im} = t_{if} + \frac{v_f}{a}. \quad (27)$$

Referring to (26) and (27), the relation between t_g the green light time at the intersection and N_g the saturated departure number of the queued vehicles can be founded as follows:

$$N_g \leq \sum_{i=1}^r i + \frac{v_f (t_g - t_k)}{h_{si}} \quad (28)$$

However, when the saturated departure of the queued vehicles at the intersection is over, the vehicles arrived subsequently from the upstream intersection are always queuing up, the saturated departure vehicles are supposed to encounter the follow-up traffic flow, resulting in that the original traffic wave disappears and a new one is formed. Given that the saturated departure rate at the intersection is q_m , the saturated departure density is k_m , the flow rate of the vehicles arrived subsequently is q_a , and the relative density is k_a , the following relation based on (16) can be worked out:

$$k_a \left(\frac{q_a}{k_a} - u'_s \right) = k_m (u_s - u'_s). \quad (29)$$

The newly formed traffic wave speed can be expressed by

$$u'_s = \frac{(q_m - q_a)}{(k_m - k_a)}. \quad (30)$$

Suppose that the meeting time and place is t_a and x_a , and the moving directions of the new traffic wave and the previous traffic flow are consistent, the time t'' when the new traffic wave comes to the stop line can be worked out through the employment of kinematical equation:

$$t'' = t_a + \frac{(x_a - x_{10})}{u'_s}. \quad (31)$$

Given that the green light time at the intersection is t_g , and the red light time is t_r , the spatial distance of each queued vehicle is l_s , the maximum queue length is l_{\max} , and

TABLE 3: Relevant departure data of traffic flow at the intersection through investigation.

	The starting time of leading vehicle	The starting time of the last vehicle	The number of queued vehicles	Length of the queue	Speed of traffic wave (m/s)	Stabilized time headway (s)	Stabilized space headway	Green light time (s)	The number of departure vehicles in green light time (s)
1	00.59.09	01.22.11	17	110.5	4.78	2.02	15.23	42	19
2	03.38.13	04.04.37	18	115.2	4.39	1.76	17.6	42	22
3	06.14.17	06.40.67	18	111.6	4.21	1.96	14.92	42	20
4	08.53.50	09.17.22	15	98.3	3.96	1.8	16.17	42	22
5	11.32.77	11.52.60	15	94.5	4.76	1.93	15.44	42	18

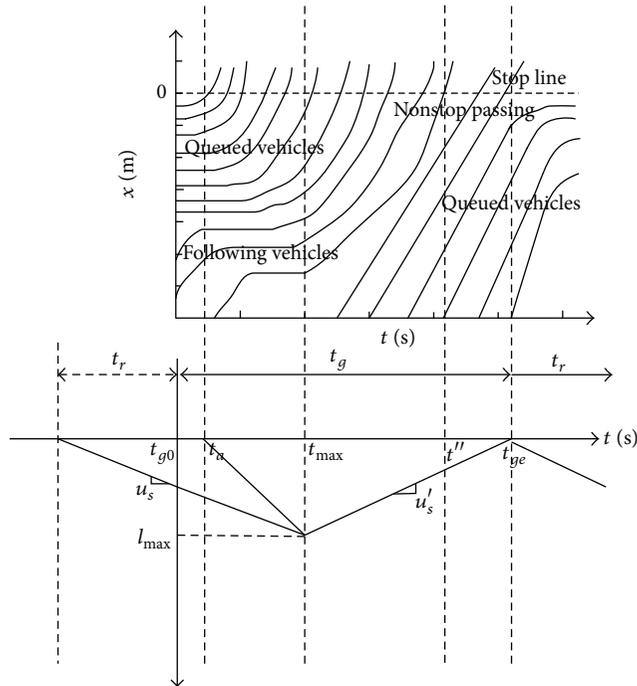


FIGURE 11: Diagram of traffic wave spreading at the intersection.

the corresponding time is t_{max} , the traffic wave spreading situation is shown as in Figure 11.

Through the analysis of Figure 11, (27), (30), and (31), the following equation can be obtained:

$$\begin{aligned}
 t_{ge} &= t'', \\
 t_g &= t_{ge} - t_{g0}, \\
 t_{ge} - t_{max} &= \frac{l_{max}}{u'_s}, \\
 N_{max} &= \sum_{i=1}^r i + q_m (t_{ge} - t_{g0} - t_{km}), \\
 \frac{l_{max}}{l_s} &= N_g.
 \end{aligned} \tag{32}$$



FIGURE 12: Departure data by video capture at the intersection.

Then the relation between the maximum departure length of the queued vehicles at the intersection and the critical green light time at the intersection is expressed in

$$l_{max} = l_s \left[\sum_{i=1}^r i + q_m (t_{ge} - t_{g0} - t_{km}) \right]. \tag{33}$$

To sum up, when $t_g > t'' - t_{g0}$, the traffic flow can be completely released within a cycle; when $t_g = t'' - t_{g0}$, the traffic flow is on the critical point where no new queue is formed; and when $t_g < t'' - t_{g0}$, a new queue will be formed relating to the traffic flow at the intersection.

On the analysis of the straight-moving traffic flow departure characteristics of the southing entrance at a typical intersection in Changchun, the relevant traffic flow parameters have been collected, such as the starting time of the head and last vehicles in the queued vehicles at the intersection, the number, the length, and the moving speed of the queued vehicles, as shown in Figure 12 and Table 3, respectively, so as to have a qualitative analysis of the departure characteristics of the traffic flow at the intersections and test the effectiveness of the model in quantitative description. Besides, qualitative analysis mentioned above is employed to analyze and calculate the traffic wave spreading speed, the relation between green light time, and the number of the departure vehicles, as well as the space and time headway when the traffic flow is dispersing steadily. During the calculation, suppose that the stop line of the intersection is the starting point, the average length of vehicles is 4.8 meters, the average acceleration of

TABLE 4: Relevant departure data of traffic flow at the intersection through calculation.

	The starting time of leading vehicle	The starting time of the last vehicle	The number of queued vehicles	Length of the queue	Speed of traffic wave	Stabilized time of traffic flow	Stabilized space headway	Green light time	The number of departure vehicles in green light time
1	00.59.09	01.22.11	17	110.5	5.62	01.17.80	18.07	42	24
2	03.38.13	04.04.37	18	115.2	4.76	03.50.67	19.36	42	27
3	06.14.17	06.40.67	18	111.6	4.7	06.28.23	15.24	42	21
4	08.53.50	09.17.22	15	98.3	4.57	09.06.07	17.36	42	24
5	11.32.77	11.52.60	15	94.5	5.04	11.44.17	14.34	42	21

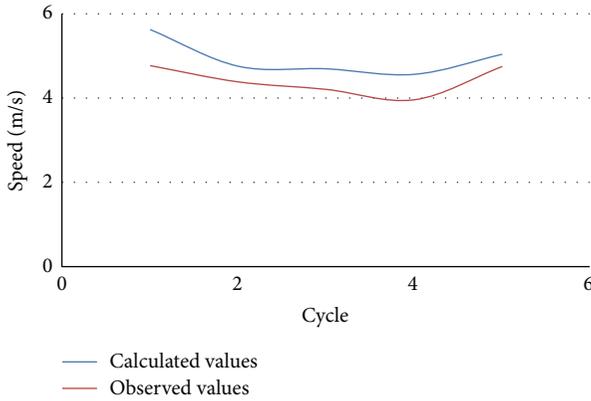


FIGURE 13: Comparison of the calculated and observed values of traffic wave speed.

the vehicles through observation is 0.52 m/s^2 , the average distance of queued vehicles is 2 meters, and the intersection distance ahead is $L_3 = 62 \text{ m} = L_1$, it can be found that the average space headway of the queued vehicles at the intersection \bar{h}_d and the jam density k_j are as follows:

$$\begin{aligned} \bar{h}_d &= 4.8 + 2 = 6.8 \text{ m}, \\ k_j &= \frac{1}{\bar{h}_d} = \frac{1}{6.8} \approx 0.147 \text{ veh/m}. \end{aligned} \quad (34)$$

Based on (6), (15), (16), (18), and (26), along with the survey data, relevant data relating to the departure process of traffic flow at the intersection is worked out as in Table 4.

Through the analysis of the data in Tables 3 and 4, it can be found from the result of the video survey that, during the five cycles, the average spreading speed at the intersection is 4.42 m/s , the number of the average passing vehicles in green light time is 20.2 , and the average time headway and space headway when the traffic flow is departure steadily are 1.89 s and 15.87 m , respectively. While through calculation, the relative data are 4.94 m/s , 23.4 , and 13.77 s , respectively. Comparing Figures 13 with 14, it can be found that the calculated value is higher than the observed value of the wave spreading speed during traffic flow departure and the passing vehicles during green light time. For example, the maximum error and the average error of traffic wave spreading speed are

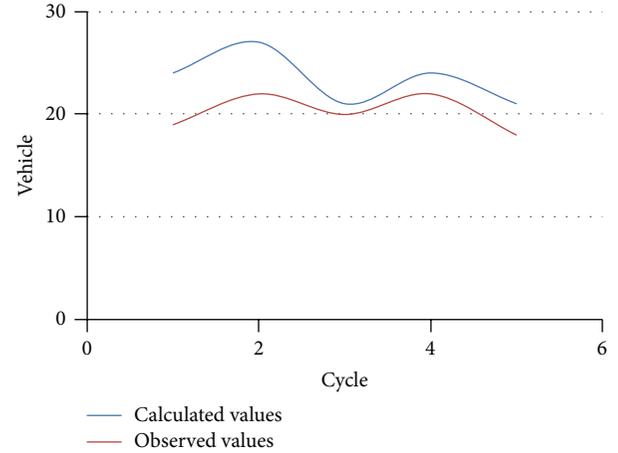


FIGURE 14: Comparison of the calculated and observed values of the passing vehicles in green time.

17.5% and 11.7% , respectively, while the relative numbers of the passing vehicles in green light time are 26.3% and 15.9% , respectively. The main cause for the deviation is that the calculation is based on the conditions that all the vehicles are small-sized cars, and they are moving with secure distance, while the fact is that there are large and medium vehicles and the drivers are extremely different in their driving behaviors.

3. Conclusions

In this paper, we applied kinematical equation and traffic wave theory to conduct qualitative analysis for the motion characteristic of the leading vehicle, the motion characteristic of the following vehicles, and the motion characteristic of the whole traffic flow at the intersection during the dispersion process at the signalized intersection. It further built spatiotemporal model of kinematics in the departure process at the intersection and traffic wave model in the dispersion process at the intersection to determine the changing situations of the leading vehicle at the departure process, the time for the following vehicles to reach the stable speed, and the relationship between the green light time at the intersection and the departure vehicle number at the intersection. In addition, combined with the vehicle operation data by video acquisition, we can know that the error is mostly less than

20% between the model calculation and the actual statistical analysis; therefore, the error is acceptable, which verified effectiveness of the model.

Acknowledgments

This work is supported by the National Science Foundation of China (nos. 51278220, 51278454, 51278520, and 51108208) and the Fundamental Research Funds for the Central Universities of China (no. 201103146). the Postdoctoral Science Foundation Funded Project of China (nos. 20110491307, 2013T60330), the Science and Technology Development Project of Jilin Province (no. 20130522121JH).

References

- [1] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Operations Research*, vol. 7, no. 4, pp. 499–505, 1959.
- [2] D. C. Gazis, R. Herman, and R. W. Rothery, "Follow-the-leader models of traffic flow," *Operations Research*, vol. 9, no. 4, pp. 545–567, 1961.
- [3] P. K. Yang, W. Z. Huang, and P. M. Che, "Traffic flow model in consideration of vehicle platoon dispersion on urban roads," *Journal of Tongji University*, vol. 22, no. 3, pp. 294–299, 1994.
- [4] W. Guo, W. Wang, G. Wets et al., "Influence of stretching-segment storage length on urban traffic flow in signalized intersection," *International Journal of Computational Intelligence Systems*, vol. 4, no. 6, pp. 1401–1406, 2011.
- [5] L. S. Brian, B. H. Russell, and B. B. Park, "Travel time estimation for urban freeway performance measurement: understanding and improving upon the extrapolation method," in *Proceedings of the 83rd Annual Meeting of the Transportation Research Board*, 2004.
- [6] W. Wang, Y. Mao, J. Jin et al., "Driver's various information process and multi-ruled decision-making mechanism: a fundamental of intelligent driving shaping model," *International Journal of Computational Intelligence Systems*, vol. 4, no. 3, pp. 297–305, 2011.
- [7] W. Wang and K. Bengler, "Computational intelligence for transportation: driving safety and assistance," *International Journal of Computational Intelligence Systems*, vol. 4, no. 3, p. 286, 2011.
- [8] S. Jin, *Modeling of Car Following Behavior Considering Visual Attention Performance*, Jilin university, JiLin, China, 2010.
- [9] P. F. Tao, *Modeling of Driving Behavior Based on the Psychology Field Theory*, Jilin university, JiLin, China, 2012.
- [10] R. W. Stokes, "Capacities of triple left-turn lanes," Tech. Rep. Committee 5P-5A, Institute of Transportation Engineers Technical Council, Washington, DC, USA, 1995.
- [11] R. I. Perez-Cartagena and A. P. Tarko, "Calibration of capacity parameters for signalized intersections in Indiana," *Journal of Transportation Engineering*, vol. 131, no. 12, pp. 904–911, 2005.
- [12] A. D. May, *Traffic Flow Fundamentals*, Prentice-Hall, Upper Saddle River, NJ, USA, 1990.
- [13] E. I. Vlahogianni, C. L. Webber Jr., N. Geroliminis, and A. Skabardonis, "Statistical characteristics of transitional queue conditions in signalized arterials," *Transportation Research C*, vol. 15, no. 6, pp. 392–403, 2007.
- [14] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Physical Review E*, vol. 51, no. 2, pp. 1035–1042, 1995.
- [15] B.-H. Wang, Y.-R. Kwong, and P.-M. Hui, "Statistical mechanical approach to cellular automaton models of highway traffic flow," *Acta Physica Sinica*, vol. 47, no. 6, pp. 914–915, 1998.
- [16] D.-H. Wang, C.-G. Jing, and Z.-W. Qu, "Application of traffic-wave theory in intersections traffic flow analysis," *China Journal of Highway and Transport*, vol. 15, no. 1, p. 93, 2002.
- [17] S. H. Yang, D. H. Wang, B. Dong, and Y. P. Wang, "Modification of start-wave model at signalized intersection," *Journal of Highway and Transportation Research and Development*, vol. 23, no. 1, pp. 130–134, 2006.
- [18] D. F. Ma, D. H. Wang, Y. M. Bie, and D. Sun, "A method of signal timing optimization for spillover dissipation in urban street networks," *Mathematical Problems in Engineering*, vol. 2013, Article ID 580546, 9 pages, 2013.
- [19] G. Y. Guo, "Based on distribution wave principle analysis of traffic congestion—dissipation process," *East China Highway*, vol. 5, pp. 77–86, 1984, <http://www.cnki.com.cn/Article/CJFDTotal-HDGL198405008.htm>.
- [20] L. Zongping B, Wei, and G. Mi, "Research on intersection signal switching model under emergency situation," *Mathematical Problems in Engineering*, vol. 2012, Article ID 186383, 14 pages, 2012.

Research Article

Study of Railway Track Irregularity Standard Deviation Time Series Based on Data Mining and Linear Model

Jia Chaolong,¹ Xu Weixiang,² Wei Lili,³ and Wang Hanning¹

¹ State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

² School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

³ Chongqing Public Security Bureau, Chongqing 401147, China

Correspondence should be addressed to Xu Weixiang; wxxu@bjtu.edu.cn

Received 13 July 2013; Revised 1 September 2013; Accepted 2 September 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Jia Chaolong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Good track geometry state ensures the safe operation of the railway passenger service and freight service. Railway transportation plays an important role in the Chinese economic and social development. This paper studies track irregularity standard deviation time series data and focuses on the characteristics and trend changes of track state by applying clustering analysis. Linear recursive model and linear-ARMA model based on wavelet decomposition reconstruction are proposed, and all they offer supports for the safe management of railway transportation.

1. Introduction

Temporal data and temporal data mining which reflect the dynamic nature of data are one of the focuses of academic community in recent years. Time series is an important temporal data. Time series similarity has been widely used in speech recognition. Euclidean distance and dynamic time warping are two classic methods. Euclidean distance is most frequently used in the time series, but Euclidean distance is a very brittle distance measure [1]. There is an obvious defect of Euclidean distance: sometimes when the sequence is very similar but the distance is great. Then DTW is proposed. DTW is an algorithm for measuring the similarity between two sequences which may vary in time or speed. A well-known application has been automatic speech recognition [2–4], to cope with different speaking speeds. Since Agrawal et al. first proposed overall matching algorithm of time series similarity search in 1993 [5], more and more scholars began to focus on temporal data mining study.

In the past four decades, scholars have proposed a variety of classic time series forecast methods, including Autoregressive Model (AR) [6], Moving Average Model (MA) [7], Autoregressive Moving Average Model (ARMA) [8–10], and Autoregressive Integrated Moving Average (ARIMA) [11].

There are a variety of methods and models in the forecast area, such as determining function method, statistical regression analysis, time series analysis, Markov model [12], state-space model [13], Bayesian forecasting model [14], as well as a variety of methods combined with theories, techniques, and methods, such as the hybrid of fuzzy theory and linear regression [15], the hybrid ARIMA and support vector machines model [16], the hybrid of time series forecast using neural networks, fuzzy logic, fractal theory [17], and the hybrid Markov model and neural networks [18]. At the same time, the number of subjects is divided into univariate and multivariate. In time series analysis and forecasting methods, the most commonly used method is based on the time domain and frequency domain. Linear and stochastic linear models and nonlinear models are two main types of study models. Random process cannot be expressed by definite function. Typically, there are mainly two types of methods to analyze random process: one is a probabilistic method, and the other is the analysis method, and the two methods are often used simultaneously in practical study. Each of AR model, MA model, ARMA model, the Markov Forecast [19], and Kalman filter model [20, 21], can be used to study the stochastic process. The characteristics of the first three models are linear forecast model and are relatively simple in terms of elements taken into consideration.

In the recent 20 years, fuzzy system is one of the most frequently used methods [22–25]. Fuzzy system has been used for a variety of optimization algorithms. In many studies, the design of fuzzy system forecast has been proposed.

In the study of time series forecast, the forecast accuracy is considered as top priority in the selection of forecast methods. The neural network is the most representative time series forecasting method, which has drawn more and more attention. In the past ten years, the neural network model is used in the study of time series forecast. The discovery of the neural network is considered to be a competitor to a variety of traditional time series models [26–28]. Because of flexible computing framework and general approximation, artificial neural network [29] is widely used in the field of predictive analytics and has higher precision.

However, an important and difficult task is faced by decision makers in many areas to improve the accuracy of the time series forecast. Using a hybrid model by combining several models has become a common practice to improve forecast accuracy. This study field has been significantly expanded [30–34].

The study has shown that forecasts of certain hybrid models are often better than those obtained from a single model forecast. Hybrid model forecast is considered to be a more accurate forecast. Its main problem can be described as follows: suppose there are n kinds of forecasts, such as $\hat{y}_1(t), \hat{y}_2(t), \dots, \hat{y}_n(t)$. The general form of this hybrid forecasting model can be defined as weighted sums of $\hat{y}_i(t)$, and the sum of all weights stack up is 1. The biggest difficulty is to determine the weights of every single forecast.

According to the principles recognized by the scientific community, simple theories are more reliable than complex ones under the same circumstances. The best scientific theory should be the simplest. railway transportation is different from automobile traffic. In automobile traffic, the driver is the major safety factor [35, 36], but in Railway transportation, track irregularity state is the decisive factor. This paper tries to look for a simple, reliable model to analyze the track irregularity change trend and to explore knowledge.

The paper is organized as follows. Section 2 introduces forecast models of time series. Data mining of track irregularity time series is described in Section 3. Section 4 presents analysis and forecast of track irregularity time series based on linear recursive forecast model. Section 5 proposes linear-ARMA model based on wavelet decomposition reconstruction. Finally, Section 6 concludes the paper with a summary.

2. Methods and Processes of Time Series Data Study

2.1. Forecast Methods of Time Series Data. Typically, there are two methods to forecast time series data: qualitative forecast and quantitative forecast. Qualitative forecast method focuses on the forecast of the nature of the development, trends, direction, and major turning points in the road of development and mainly depends on human experience and analytical ability. Quantitative forecast focuses on the analysis of quantitative aspects of development, attaches importance

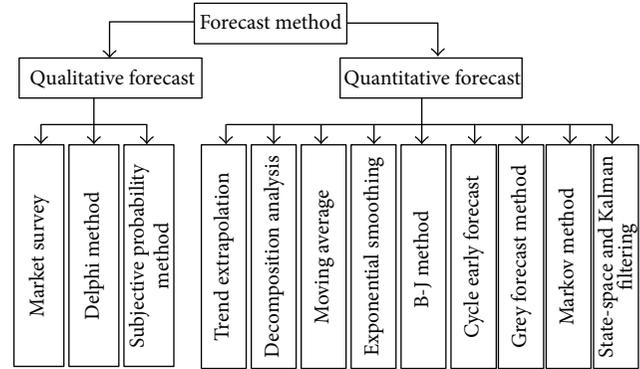


FIGURE 1: Main forecast methods of time series.

to the quantitative description of the degree of development and changes, is based more on historical statistics, and is less affected by subjective factors. Qualitative forecast and quantitative forecast are not mutually exclusive but can be complementary to each other and should be combined correctly in the actual forecast process.

Two types of time series forecasting methods and their methods are shown in Figure 1.

2.2. Study Processes of Time Series Data. The study processes of time series data include three steps: data acquisition, correlation analysis, and model identification. Specific study processes are shown in Figure 2.

3. Data Mining of Track Irregularity Time Series

3.1. Theoretical Analysis of Data Mining. Data mining is the process of mining interested information from databases, data warehouses, or other data repositories. It is an iterative process, whose study steps are shown in Figure 3.

The main method, purpose, and contents of time series data mining include: time series segmentation studying underlying mechanisms change of time series and representation at high level; similarity search looking for similar sequences; clustering analysis on similarity measure, clustering algorithms and results, and gathering similar time sequence variation into one class; classification and sequence analysis on time series and the time points in the entire time series; anomaly detection which finds the abnormality of sequence, points, and mode; analysis on the law and trends of time series changing over time, forecast of the future data and trends based on the historical and current data; using graphics technology, virtual reality technology, and data mining technology to display complex time series in a way which is easy, intuitive, and graphical for people to understand so that we can realize visualized and practical study of time series data.

3.2. Time Series Similarity. Time series similarity is the basis of time series analysis and time series data mining. The similarity is achieved by calculating the distance between the time series.

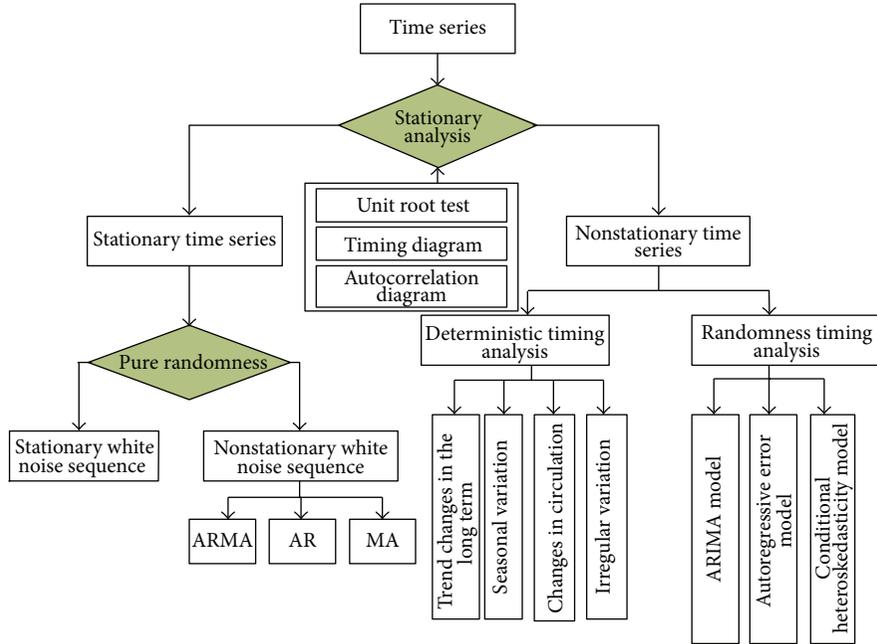


FIGURE 2: Classification of time series analysis.

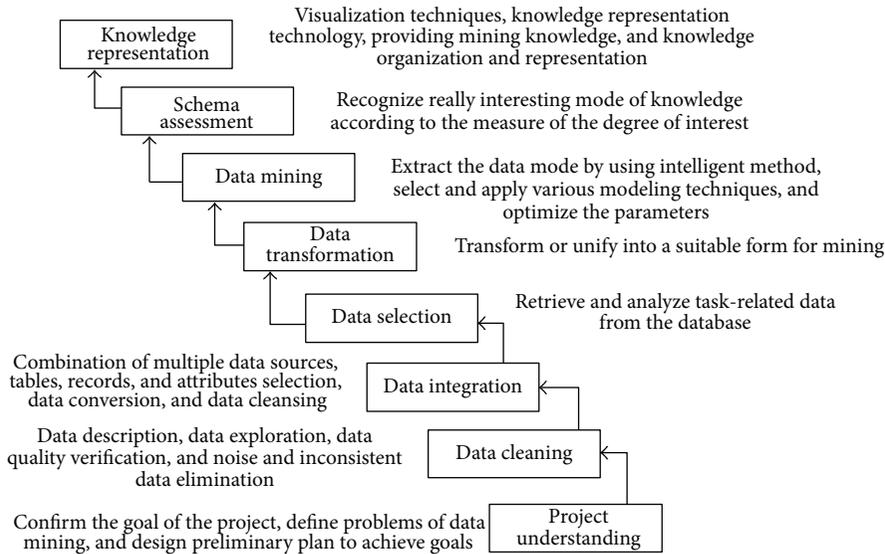


FIGURE 3: Steps of data mining process.

The definition of distance needs to meet the following four properties:

- (1) $d_{ij} = 0 \Leftrightarrow x_i = x_j$,
- (2) $\forall x_i, x_j, d_{ij} \geq 0$,
- (3) $\forall x_i, x_j, d_{ij} = d_{ji}$,
- (4) $\forall x_i, x_j, x_k, d_{ij} \leq d_{ik} + d_{kj}$.

d_{ij} is the distance between x_i and x_j in a group of objects $D = \{x_1, x_2, \dots, x_n\}$ in m -dimensional space.

Minkowski distance is the common formula to calculate distance, and the expression is

$$d_{ij} = \sum_{k=1}^m (|x_{ik} - y_{jk}|^p)^{1/p}, \quad (p > 0). \quad (1)$$

Many distances are formed by changing the parameters of the Minkowski distance. Distance formula commonly used includes Manhattan distance, Euclidean distance, Chebyshev distance, custom distance, Mahalanobis distance, Minkowski distance, variance weighted distance, Canberra Distance,

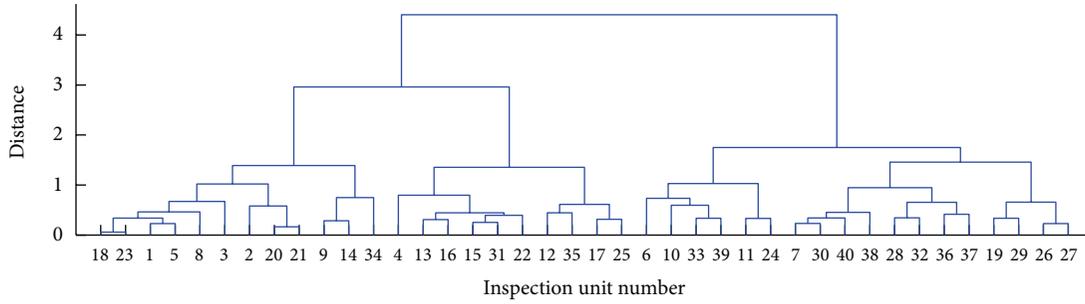


FIGURE 4: Visual cluster tree.

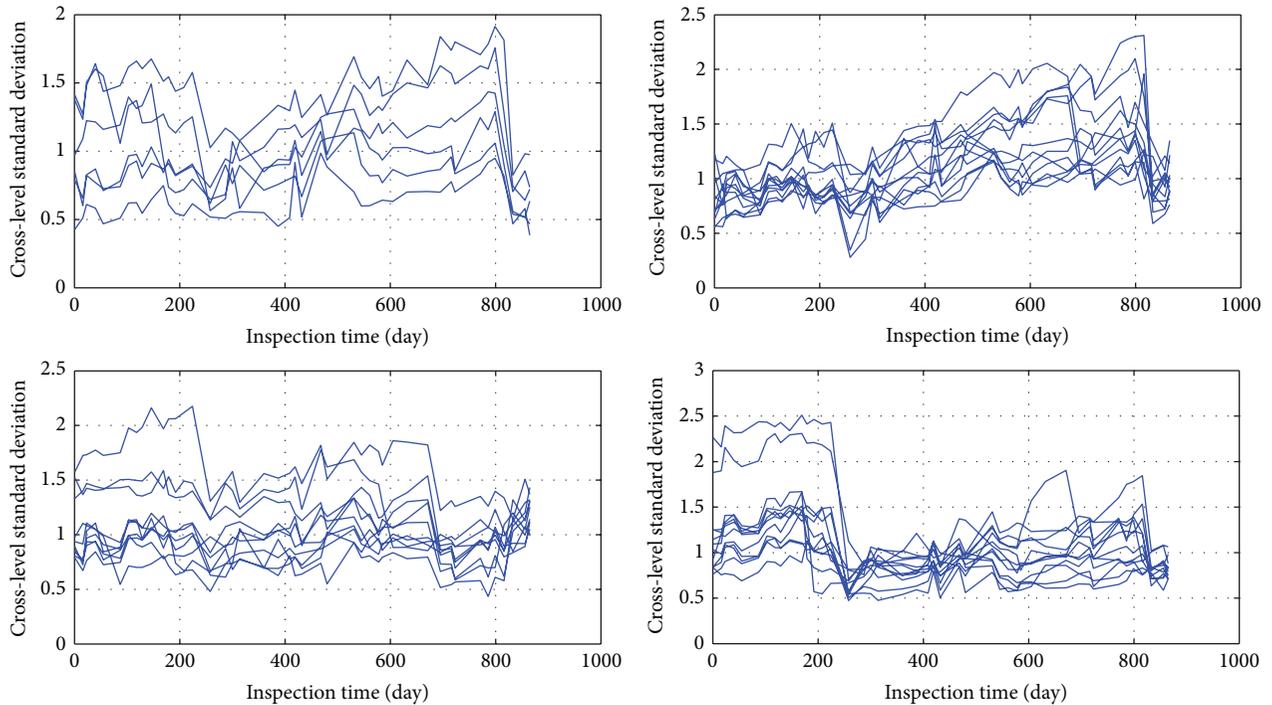


FIGURE 5: Clustering results.

dynamic time warping distance, the cross-correlation distance, KL distance, and so forth.

3.3. Clustering Analysis of Track Irregularity. Standard deviation between the data is used to reflect the degree of variation. When the two sets of data are under circumstances of the same units and similar means, the greater the standard deviation is, the greater the degree of variation between the data will be. When the data around the mean of the distribution is more discrete, representation of the mean is weaker. On the contrary, the smaller the standard deviation is, the little the variance between the data is indicated, and the distribution of data around the mean is more intensive, and representation of the mean is better. Therefore, standard deviation of track irregularity data is selected in the study of discrete distribution of track irregularity data over time to evaluate the development trends of track irregularity. In this study, track irregularity data is provided by State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University.

We take K449-K450 km section, Beijing-Kowloon line, with 44th cross-level inspection data as study data. The section is divided into 40 units, and the standard deviation of each unit is calculated; then we get 40 cross-level standard deviation time-series data. This standard deviation data is the object of cluster analysis. In this paper, k -mean algorithm is used in clustering algorithm, cross correlation distance is used in the distance between objects in the matrix, and the minimum variance algorithm is used in the connection between the variables. Clustering results are shown in Figures 4 and 5.

The above track irregularity standard deviation time series which have similar clustering methods, changing trends, and characteristics are clustered into one category.

4. Linear Recursive Model

4.1. Core Ideas. Because of the inertia characteristics of the track state changes, track state has a memory effect. The latest

track state and the nearest previous state shares similarity, and the inspection state of the adjacent time points has a similar trend. From the macroperspective, the track state presents nonlinear changes throughout the whole life cycle of the track [37, 38], but from the microperspective, in a short time, track state changes at adjacent time will be close to linear features.

Based on the above assumptions, this paper proposes linear model of track state changes as follows:

$$s_{i,j} = k_i t_{i,j} + b_i. \quad (2)$$

In the formula, s_{ij} is the track state of the j th day between the i th to the $(i + 1)$ th inspection; k_i is slope value of track state linear change between the i th to the $(i + 1)$ th inspection; b_i is intercept; $t_{i,j}$ is the j th day between the i th to the $(i + 1)$ th inspection.

According to the least square method, the model employs the vector form. Least squares estimation of the model parameters accords to the following conditions:

$$\hat{\alpha} = (B^T B)^{-1} B^T Y. \quad (3)$$

In the formula:

$$\hat{\alpha} = \begin{pmatrix} k_i \\ b_i \end{pmatrix},$$

$$Y = \begin{pmatrix} s_{i+1,0} \\ s_{i+2,0} \\ \vdots \\ s_{i+n,0} \end{pmatrix}, \quad (4)$$

$$B = \begin{pmatrix} t_{i+1} & 1 \\ t_{i+2} & 1 \\ \vdots & \vdots \\ t_{i+n} & 1 \end{pmatrix}.$$

In the formula, t_{i+n} is the length of time between the i th to the $(i + 1)$ th inspection; $t_{i+n} = T_{i+n} - T_{i+n-1}$, T_{i+n} is the time of state value $s_{i+n,0}$ at the $(i + n)$ th inspection; $s_{i+n,0}$ is state value of the $(i + n)$ th inspection.

The form of recursive forecast of the model is as follows:

$$\begin{aligned} s_{i+m,j} &= k_{i+m} t_{i+m,j} + b_{i+m}, \\ s_{i+m+1,j} &= k_{i+m+1} t_{i+m+1,j} + b_{i+m+1}, \\ s_{i+m+2,j} &= k_{i+m+2} t_{i+m+2,j} + b_{i+m+2} \\ &\vdots \\ s_{i+T,j} &= k_{i+T} t_{i+T,j} + b_{i+T}. \end{aligned} \quad (5)$$

In the formula, $m > n$, and $T > m + n$.

4.2. Model Analysis. Three time intervals data of the first four historical inspection data and the second to fourth inspection data in cross-level standard deviation time series data are considered as an input; the slope and intercept of track

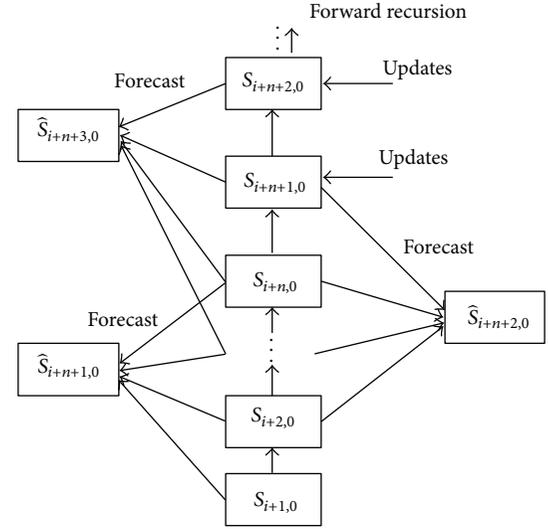


FIGURE 6: Forecast process of linear recursive model.

state linear changes in time interval between the fourth and fifth inspection are considered as output. When the forecast finishes, the fifth forecast data is updated by the fifth standard deviation of the actual inspection data, then three-time interval between the second and the fifth inspection data and the third to fifth inspection data are used to forecast the slope and intercept of track state linear changes in time interval between the fifth and sixth inspections. As the data is updated gradually, the model pushes and forecasts towards ahead.

Meanwhile, the linear forecast model proposed in this paper does not have to consider equal interval requirements of data samples like general time series models. In the track irregularity time series data, especially data of track inspection car, due to maintenance work, passenger and cargo plan changes, and the track inspection car scheduling and other reasons, inspection data is unequal interval, namely, in the model $t_{i+n,0} \neq \text{const}$. Linear model has the following advantages: simple model based on a linear model and least squares solution; without regard to unequal interval characteristics. Length of unequal interval is seen as a parameter directly in the model.

Forecast process of linear recursive model is shown in Figure 6.

4.3. Residual Correction. In order to further improve the forecast accuracy of the model, the residual needs to be corrected. Model residuals sequence data is a cyclical variation with concussive time series. In this paper, the Fourier transforming idea is used on residuals analysis.

Model residuals sequence data is as follows:

$$e_{i+k} = \hat{s}_{i+k} - s_{i+k}. \quad (6)$$

The Fourier transforming residuals expression $e(t_i)$ is

$$\hat{e}_{i+k} = \frac{1}{2} h_0 + \sum_{j=1}^m [h_j \cos(\omega_j \cdot T_{i+k}) + l_j \sin(\omega_j \cdot T_{i+k})]. \quad (7)$$

In the formula,

$$\begin{aligned}
 i &= 1, 2, \dots, n, \\
 h_0 &= \frac{1}{n} \sum_{k=1}^n |e_{i+k}|, \\
 h_j &= \frac{2}{n} \sum_{k=1}^n e_{i+k} \cdot \cos(\omega_j \cdot T_{i+k}), \\
 l_j &= \frac{2}{n} \sum_{k=1}^n e_{i+k} \cdot \sin(\omega_j \cdot T_{i+k}), \\
 \omega_j &= \frac{2j\pi}{\sum_{k=1}^n t_{i+k}}.
 \end{aligned} \tag{8}$$

By looking for the appropriate value of m , fitting errors between model residuals e_{i+k} and periodic sequence \hat{e}_{i+k} achieved through Fourier transforming can get to a minimum value; namely,

$$\sum_{k=1}^n \frac{(\hat{e}_{i+k} - e_{i+k})^2}{e_{i+k}} = \min. \tag{9}$$

After getting the best value of m , the final forecast model is

$$\begin{aligned}
 s_{i+k} &= \hat{s}_{i+k} + \hat{e}_{i+k} \\
 &= \hat{s}_{i+k} + \frac{1}{2}h_0 \\
 &\quad + \sum_{j=1}^m [h_j \cos(\omega \cdot T_{i+k}) + l_j \sin(\omega \cdot T_{i+k})], \\
 i &= 1, 2, \dots, n.
 \end{aligned} \tag{10}$$

4.4. A Case Study. Change trend and periodic characteristic of track irregularity standard deviation data can be classified by cluster analysis. Since the cycle of track maintenance work is determined by track irregularity change, so it is uncertain. Thus, data containing multiple cycle cannot be used for the study. The data within a period that best represents the development trend of track irregularity characteristics is selected as data object. Change trends of track irregularities within each period can be forecasted by studying the law of track irregularity standard deviation changes in this cycle. So, 25 of track cross-level standard deviation data of Beijing-Kowloon line, K449 + 800 – K449 + 825 mileage unit, from November 13, 2008 to May 18, 2010, are taken as the study object to forecast 5–25th standard deviation data of track cross-level.

By solving the least squares estimation, we get recursion $\{(k_i, b_i)\}$ as follows: $\{(k_i, b_i)\} = \{(0.004012877, 1.052496244), (0.004880568, 1.072606649), (-0.000814172, 1.257042149), (0.003246744, 1.247075988), (-0.003644461, 1.423421705), (0.001636605, 1.30965592), (-0.001246396, 1.374786532), (0.009376521, 1.168171032), (0.001440352, 1.442494856), (-0.006255225, 1.647403617), (0.002275014, 1.502381251),$

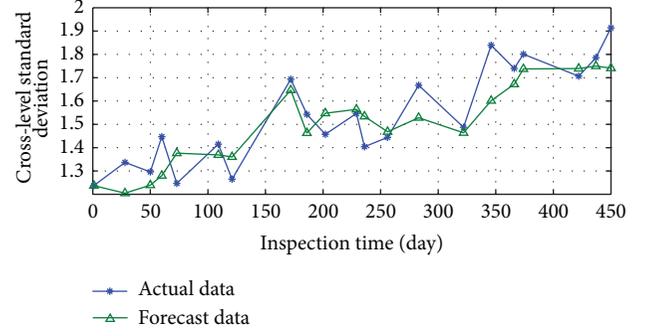


FIGURE 7: Forecast curve and the actual curve of cross-level standard deviation (K449 + 800 to K449 + 825 unit segment).

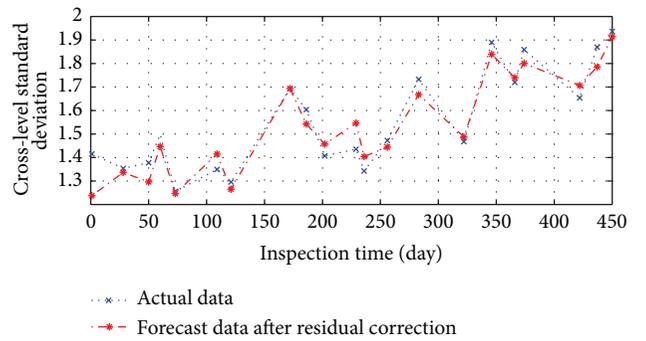


FIGURE 8: Forecast curve after residual revised and the actual curve of cross-level standard deviation.

$(-0.002533806, 1.551368672), (-0.000951367, 1.486815078), (0.007406862, 1.328357308), (-0.002678251, 1.568176542), (0.0066245, 1.442194457), (0.002152384, 1.629003922), (-0.001214053, 1.747250648), (-0.001427253, 1.807423072), (0.000826365, 1.737111301), (0.004952398, 1.676475089)\}$.

When the parameters $\{(k_i, b_i)\}$ are substituted into the short time track state changes linearly formula, the daily track state changes between two inspection time points can be estimated, and the forecast curve and the actual curve are shown in Figure 7.

Forecast results after residual correction are shown in Figure 8.

If we put several indicators together, we can fully measure the forecast accuracy of the model. Here, we use mean square error (MSE) and mean absolute percentage error (MAPE).

MSE is the expectation value of square of difference between the estimate value and true value of the parameter, and it indicates the degree of changes of data. The smaller the MSE value is, the more accurate the forecast model is in describing the experimental data. The expression of MSE is as follows:

$$\text{MSE} = \sqrt{\frac{\sum \varepsilon^2}{n}}. \tag{11}$$

TABLE 1: Analysis of accuracy of the cross level standard deviation.

Units	MSE	MAPE	MSE after residual correction	MAPE after residual correction
K449 + 000 – K449 + 025	0.11	12%	0.08	7.4%
K449 + 100 – K449 + 125	0.14	9.8%	0.07	5.0%
K449 + 200 – K449 + 225	0.09	6.4%	0.07	4.2%
K449 + 325 – K449 + 350	0.11	7.6%	0.06	3.8%
K449 + 425 – K449 + 450	0.08	6.0%	0.06	4.0%
K449 + 475 – K449 + 500	0.12	12%	0.07	7.5%
K449 + 575 – K449 + 600	0.10	7.7%	0.07	5.6%
K449 + 675 – K449 + 700	0.09	4.9%	0.05	2.3%
K449 + 800 – K449 + 825	0.10	5.5%	0.07	3.6%
K449 + 875 – K449 + 900	0.07	4.4%	0.04	2.0%
K449 + 975 – K450 + 000	0.09	7.4%	0.05	3.7%

Expression of MAPE is as follows:

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^n \left| \frac{z_t - \hat{z}_t}{z_t} \right|. \quad (12)$$

Forecast accuracy of MAPE can be divided into four indicators: high-precision forecast (MAPE < 10%), good forecast (10% < MAPE < 20%), feasible forecast (20% < MAPE < 50%), and error forecast (MAPE > 50%).

Forecast accuracy of the model in some units segment is shown in Table 1.

Through the comparison of forecast accuracy indicators MSE and MAPE, we can find that the values of MSE and MAPE of the model after residuals correction are significantly reduced, and the forecast accuracy is generally improved by 30–40%, which belongs to high-precision forecast.

5. Linear-ARMA Model Based on Wavelet Decomposition Reconstruction

Forecast models after residual correction generally enjoy higher forecast accuracy than the original forecast models, but this also increases the computation and complexity of the model. Therefore, this paper will propose linear-ARMA model based on wavelet decomposition reconstruction.

5.1. Wavelet Transform. The wavelet transform [39–45] is a new field developing rapidly in applied mathematics and engineering. It is a new branch of mathematics and perfectly combines functional analysis, Fourier analysis, sample transfer analysis, and numerical analysis. It is based on certain special functions; it converts data process or data sequence into series in order to find the similar spectrum characteristics and finally achieves data processing. The wavelet transform is the local transformation of space (time) and frequency and can extract information effectively from the signal and do multiscale detailed analysis to functions or signals by dilation and translation and other computing functions.

As the name implies, “wavelet” means a waveform with a small area, limited length, and zero mean value. “Small” refers to the decay of the wavelet; while the “wave” refers to its volatility, its amplitude shocks in alternate positive and

negative forms. Compared to the Fourier transform, wavelet transform is the localized analysis of time (spatial) frequency; it eventually reaches breakdown of time at a high frequency; and subdivision of frequency at the low frequency, can automatically adapt to the requirements of time-frequency signal analysis, and can focus on any detail of the signal, solving the difficulties of Fourier transforming. Thus, it becomes a major breakthrough in the scientific method after the Fourier transforming. So, the wavelet transform is even called “mathematical microscope.”

In summary, the method that divides functions into a series of simple basic functions is of theoretical and practical significance. In this paper, Daubechies wavelet is used in track irregularity time series data decomposition. Daubechies is the general name of a series of binary wavelet proposed by the French scholar Daubechies, which can do multiscale wavelet decomposition to signal.

5.2. Core Ideas. According to the idea of wavelet decomposition and reconstruction, the standard deviation of track irregularity sequence data waveform signal is decomposed into detail waveform signal ($D1, D2, D3$) and approximate waveform signal ($A3$), in which detail waveform signal is stationary series with zero mean. We can use the random linear model to study it. We use ARMA model in this paper. Approximate waveform signals are generally nonstationary, smooth sequence curves. According to its smooth characteristics, the linear recursive model is used in trend analysis. The modeling idea is shown in Figure 9.

In the figure, $A3_i$ is the actual value and $\hat{A}3_i, \hat{A}3_{i+1}$ are forecast values. According to the geometric relationship, the values $\hat{A}3_i, \hat{A}3_{i+1}$ are as follows:

$$\begin{aligned} \hat{A}3_i &= A3_{i-1} + (A3_{i-1} - A3_{i-2}), \\ \hat{A}3_{i+1} &= A3_i + (A3_i - A3_{i-1}). \end{aligned} \quad (13)$$

Finally, the two partial results are combined, and all the forecast sequence data are added up by weight 1. The formula is shown as follows:

$$\hat{S} = \hat{D}1 + \hat{D}2 + \hat{D}3 + \hat{A}3. \quad (14)$$

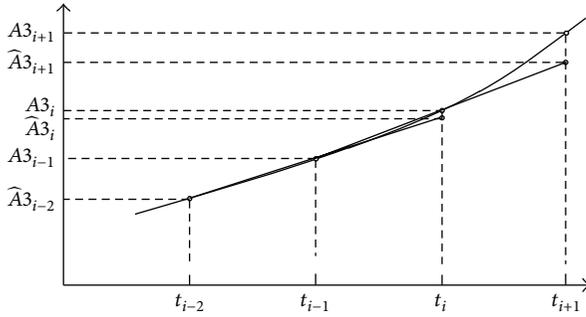


FIGURE 9: Linear recursive model of approximate waveform signal (LF).

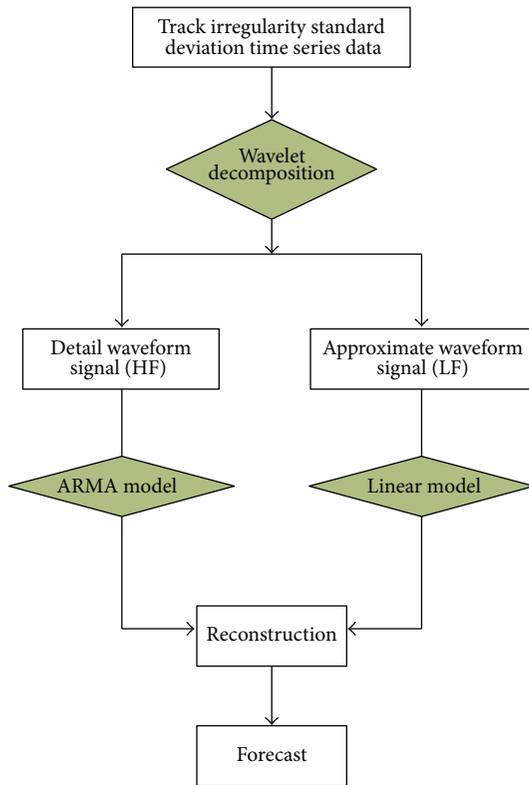


FIGURE 10: Modeling process of linear-ARMA model based on wavelet decomposition reconstruction.

In the formula, $\widehat{D}1$, $\widehat{D}2$, and $\widehat{D}3$ are high-frequency detail signal sequence forecasted by the ARMA model, $\widehat{A}3$ is sequence of the low-frequency approximation by linear model forecast, and \widehat{S} is forecast value of the final state.

After decomposition-reconstruction process, the final forecast result appears, and the modeling process is shown in Figure 10.

5.3. Case Study. The Daubechies wavelet is selected to decompose the track irregularity standard deviation time series data signal; with the decomposition depth being 3, we use Mallat tower algorithm to do decomposition and

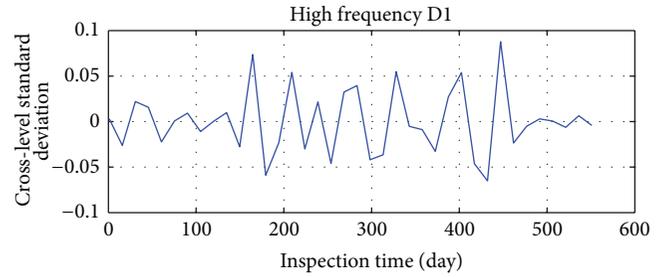


FIGURE 11: Level 1 detail waveform signal (HF) of cross-level standard deviation.

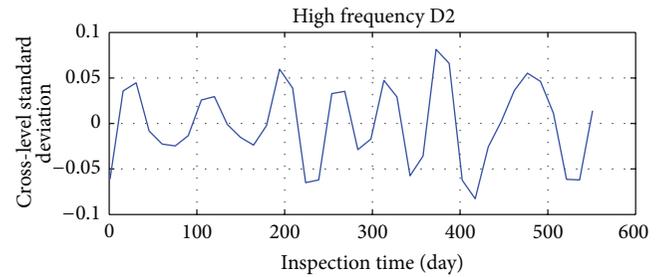


FIGURE 12: Level 2 detail waveform signal (HF) of cross-level standard deviation.

reconstruction of track irregularity standard deviation time series.

K449-450 km section, the Beijing-Kowloon line, upgoing, at 44th cross-level inspection data is selected as the study data. The section is divided into 40 units, and the standard deviation of each unit is calculated, and we get 40 cross-level standard deviation time-series data. After equal time intervals conversion of standard deviation time-series, we use linear-ARMA model based on wavelet decomposition and reconstruction for analysis.

Take K449 + 800 – K449 + 825 unit sections, from November 13, 2008, to May 18, 2010, with 38 cross-level standard deviation data as study data. The Linear-ARMA model process based on wavelet decomposition and reconstruction is as follows.

- (1) Wavelet decomposition of track cross-level standard deviation data: wavelet decomposition process of cross-level standard deviation time series is the process that divides cross-level standard deviation time series data into high frequency detail waveform signal and low frequency approximate waveform signal. The decomposed waveform signals are shown in Figures 11, 12, 13, and 14.
- (2) ARMA model forecast of high-frequency signals: high-frequency signal is stationary time series. We use ARMA model to forecast, and the results are shown in Figures 15, 16, and 17.
- (3) Low-frequency signal linear model forecast: low-frequency signal is a smooth curve, and its linear forecast result is shown in Figure 18.

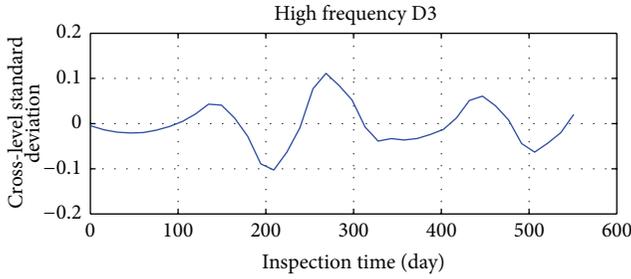


FIGURE 13: Level 3 detail waveform signal (HF) of cross-level standard deviation.

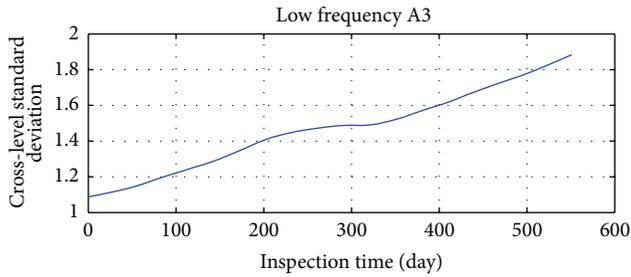


FIGURE 14: Layer 3 approximate series waveform signal (LF) of cross-level standard deviation.

(4) Model reconstruction and accuracy analysis: all forecast sequence data is added up by weight 1; comparison of forecast values and original values are shown in Figure 19.

Forecast error of the model is shown in Figure 20.

5.4. Accuracy Analysis. MSE and MAPE are used to measure accuracy of the model, and the forecast the accuracy of some units is shown in Table 2.

According to the forecast results of MSE, MAPE value in Table 2, the model has higher forecast accuracy, and so there is no need to correct the residual like linear recursive model. This has shown that linear-ARMA model based on wavelet decomposition reconstruction is an effective way to forecast the trend of track state changes.

6. Conclusions

In this paper, data mining and time series theory are used to study track irregularity standard deviation time series data. The main purpose of this study is to forecast track irregularity state in future. By using clustering analysis theory in data mining, different patterns and characteristics of track irregularity change can be found. Through a systematic study of time series data classification and time series forecast model, this paper puts forward linear recursive models and linear-ARMA model based on wavelet decomposition reconstruction to forecast the changing trends of track irregularity standard deviation time series. Simulation results show that the models have higher accuracy.

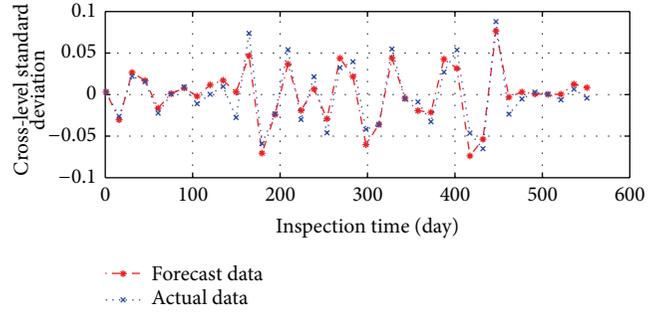


FIGURE 15: Level 1 forecast value of detail waveform signal (HF) and original data cross-level standard deviation.

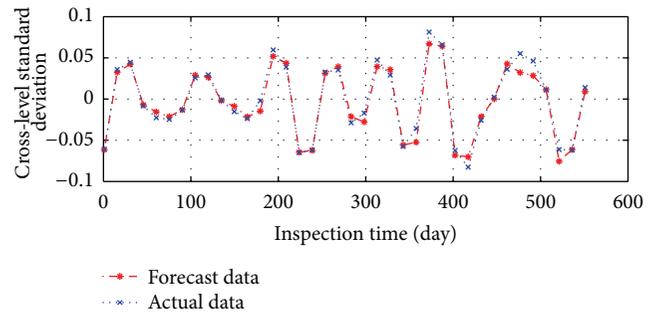


FIGURE 16: Level 2 forecast value of detail waveform signal (HF) and original data cross-level standard deviation.

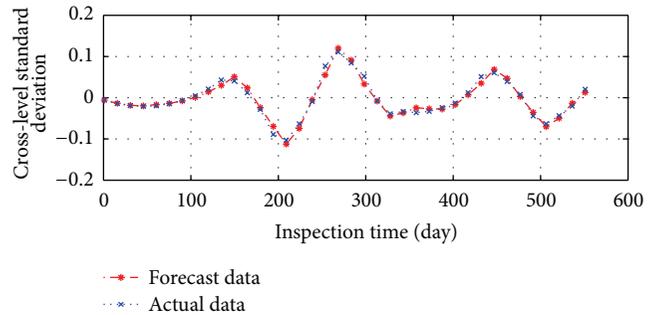


FIGURE 17: Level 3 forecast value of detail waveform signal (HF) and original data cross-level standard deviation.

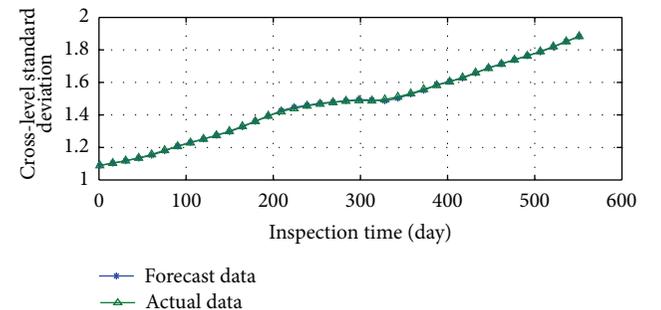


FIGURE 18: Layer 3 forecast value of approximation waveform signal (LF) and original data of cross-level standard deviation.

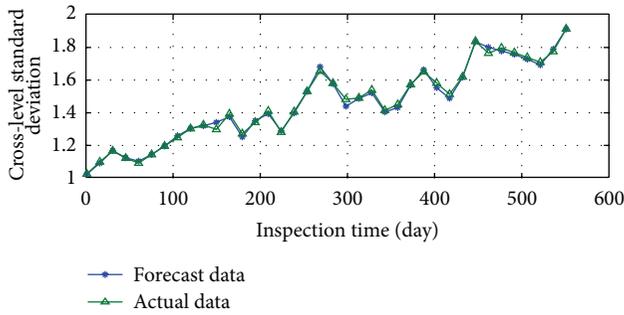


FIGURE 19: Comparison of forecast values and original values.

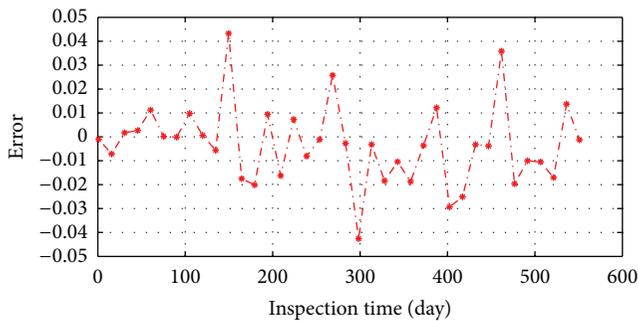


FIGURE 20: Error analysis.

TABLE 2: Model accuracy.

Units	MSE	MAPE
K449 + 000 – K449 + 025	0.03	3.5%
K449 + 100 – K449 + 125	0.02	2.0%
K449 + 200 – K449 + 225	0.03	2.1%
K449 + 325 – K449 + 350	0.02	2.0%
K449 + 425 – K449 + 450	0.02	1.8%
K449 + 475 – K449 + 500	0.03	2.8%
K449 + 575 – K449 + 600	0.03	1.9%
K449 + 675 – K449 + 700	0.03	1.1%
K449 + 800 – K449 + 825	0.02	0.8%
K449 + 875 – K449 + 900	0.02	1.3%
K449 + 975 – K450 + 000	0.02	1.7%

The change of railway track state is a complex process. It is affected by various aspects of the situation. Although it is an extremely difficult task to explore laws of its development and changes, the significance of the study is far reaching. Because track state is inspected by section, it is not carried out for each fixed measuring points, although there is mileage data information of fixed measuring points to be studied in the inspection data. However, it is unavoidable that there is mileage deviation between the actual measuring point's mileage data information and measuring point's mileage data information in inspection data, and the mileage deviation will be a few or dozens of meters. It is the basic idea in the study of track state changes that puts the data after mileage relative calibration as the subject or takes the section as a whole as

the subject. Due to track maintenance and repair cycle, we only study track state changes in a cycle. However, track state change trend between each of the maintenance and repair cycle is also worth to be studied.

Conflict of Interests

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence their work; there is no professional or other personal interests of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, this paper.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 61272029), National Key Technology R&D Program (Grant no. 2009BAG12A10), independent subject of State Key Laboratory of Rail Traffic Control and Safety, and Beijing Jiaotong University (Contract no. RCS2009ZT007) and partially supported by the MOE key Laboratory for Transportation Complex Systems Theory and Technology School of Traffic and Transportation, Beijing Jiaotong University.

References

- [1] E. Keogh and C. A. Ratanamahatana, "Exact indexing of dynamic time warping," *Knowledge and Information Systems*, vol. 7, no. 3, pp. 358–386, 2005.
- [2] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 23, no. 1, pp. 67–72, 1975.
- [3] J. B. Kruskal and M. Liberman, *The Symmetric Time Warping Algorithm: From Continuous to Discrete, Time Warps*, Addison-Wesley, Reading, Mass, USA, 1983.
- [4] C. Myers, L. R. Rabiner, and A. E. Rosenberg, "Performance tradeoffs in dynamic time warping algorithms for isolated word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 6, pp. 623–635, 1980.
- [5] R. Agrawal, C. Faloutsos, and A. Swami, "Efficient similarity search in sequence databases," in *Proceedings of the 4th International Conference on Foundations of Data Organization and Algorithms*, pp. 69–84, Chicago, USA, 1993.
- [6] G. U. Yule, "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers," *Philosophical Transactions of the Royal Society of London A*, vol. 226, pp. 267–298, 1927.
- [7] G. Walker, "On periodicity in series of related terms," *Proceedings of the Royal Society of London A*, vol. 131, no. 818, pp. 518–532, 1931.
- [8] E. J. Hannan, *Multiple Time Series. Wiley Series in Probability and Mathematical Statistics*, John Wiley and Sons, New York, NY, USA, 1970.
- [9] P. Whittle, *Prediction and Regulation by Linear Least-Square Methods*, University of Minnesota Press, Minneapolis, Minn, USA, 1983.

- [10] E. J. Hannan, M. Deistler, and John Wiley and Sons, *Statistical Theory of Linear Systems. Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons, New York, NY, USA, 1988.
- [11] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, NC, USA, 1970.
- [12] L. Rabiner and B. Juang, "An introduction to hidden Markov models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, 1986.
- [13] J. Durbin and S. Koopman, *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford, UK, 2001.
- [14] P. J. Harrison and C. F. Stevens, "Bayesian forecasting," *Journal of the Royal Statistical Society B*, vol. 38, no. 3, pp. 205–247, 1976.
- [15] P.-T. Chang and E. S. Lee, "Fuzzy linear regression with spreads unrestricted in sign," *Computers and Mathematics with Applications*, vol. 28, no. 4, pp. 61–70, 1994.
- [16] P. Pai and C. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497–505, 2005.
- [17] O. Castillo and P. Melin, "Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic, and fractal theory," *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1395–1408, 2002.
- [18] Y. H. Tay, M. Khalid, R. Yusof, and C. Viard-Gaudin, "Offline cursive handwriting recognition system based on hybrid Markov model and neural networks," in *Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation*, vol. 3, pp. 1190–1195, 2003.
- [19] R. Krzysztofowicz, "A Bayesian Markov model of the flood forecast process," *Water Resources Research*, vol. 19, no. 6, pp. 1455–1465, 1983.
- [20] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [21] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–107, 1961.
- [22] I. Sugianto and S. Natarajan, "Parameter estimation using least square method for MIMO Takagi-Sugeno neuro-fuzzy in time series forecasting," *Jurnal Teknik Elektro*, vol. 7, no. 2, pp. 82–87, 2007.
- [23] M. Zounemat-Kermani and M. Teshnehlab, "Using adaptive neuro-fuzzy inference system for hydrological time series prediction," *Applied Soft Computing Journal*, vol. 8, no. 2, pp. 928–936, 2008.
- [24] H. J. Rong, N. Sundararajan, G. B. Huang, and P. Saratchandran, "Sequential Adaptive Fuzzy Inference System (SAFIS) for nonlinear system identification and prediction," *Fuzzy Sets and Systems*, vol. 157, no. 9, pp. 1260–1275, 2006.
- [25] S. Paul and S. Kumar, "Subsethood-product fuzzy neural inference system (SuPFuNIS)," *IEEE Transactions on Neural Networks*, vol. 13, no. 3, pp. 578–599, 2002.
- [26] S. D. Balkin and J. K. Ord, "Automatic neural network modeling for univariate time series," *International Journal of Forecasting*, vol. 16, no. 4, pp. 509–515, 2000.
- [27] Y. Chen, B. Yang, J. Dong, and A. Abraham, "Time-series forecasting using flexible neural tree model," *Information Sciences*, vol. 174, no. 3–4, pp. 219–235, 2005.
- [28] F. Giordano, M. La Rocca, and C. Perna, "Forecasting nonlinear time series with neural network sieve bootstrap," *Computational Statistics and Data Analysis*, vol. 51, no. 8, pp. 3871–3884, 2007.
- [29] R. Adhikari and R. K. Agrawal, "Forecasting strong seasonal time series with artificial neural networks," *Journal of Scientific & Industrial Research*, vol. 71, no. 10, pp. 657–666, 2012.
- [30] G. Armano, M. Marchesi, and A. Murru, "A hybrid genetic-neural architecture for stock indexes forecasting," *Information Sciences*, vol. 170, no. 1, pp. 3–33, 2005.
- [31] S. M. Chen and N. Y. Chung, "Forecasting enrollments using high-order fuzzy time series and genetic algorithms," *International Journal of Intelligent Systems*, vol. 21, no. 5, pp. 485–501, 2006.
- [32] Z. J. Zhou and C. H. Hu, "An effective hybrid approach based on grey and ARMA for forecasting gyro drift," *Chaos, Solitons and Fractals*, vol. 35, no. 3, pp. 525–529, 2008.
- [33] L. Yu, S. Wang, and K. K. Lai, "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates," *Computers and Operations Research*, vol. 32, no. 10, pp. 2523–2541, 2005.
- [34] H. Kim and K. Shin, "A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets," *Applied Soft Computing Journal*, vol. 7, no. 2, pp. 569–576, 2007.
- [35] H. Guo, W. Wang, W. Guo, X. Jiang, and H. Bubb, "Reliability analysis of pedestrian safety crossing in urban traffic environment," *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.
- [36] W. Wang, Y. Mao, J. Jin et al., "Driver's various information process and multi-ruled decision-making mechanism: a fundamental of intelligent driving shaping model," *International Journal of Computational Intelligence Systems*, vol. 4, no. 3, pp. 297–305, 2011.
- [37] C. Jia, W. Xu, F. Wang, and H. Wang, "Track irregularity time series analysis and trend forecasting," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 387857, 15 pages, 2012.
- [38] H. Wang, W. Xu, F. Wang, and C. Jia, "A cloud-computing-based data placement strategy in high-speed railway," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 396387, 15 pages, 2012.
- [39] C. Cattani and A. Kudreyko, "Harmonic wavelet method towards solution of the Fredholm type integral equations of the second kind," *Applied Mathematics and Computation*, vol. 215, no. 12, pp. 4164–4171, 2010.
- [40] S. V. Narasimhan, M. Harish, A. R. HariPriya, and N. Basumallick, "Discrete cosine harmonic wavelet transform and its application to signal compression and subband spectral estimation using modified group delay," *Signal, Image and Video Processing*, vol. 3, no. 1, pp. 85–99, 2009.
- [41] H. A. Yousef, M. E. Elkhatib, and O. A. Sebakhy, "Wavelet network-based motion control of DC motors," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1522–1527, 2010.
- [42] M. Seonguk and K. Min, "Detection of combustion start in the controlled auto ignition engine by wavelet transform of the engine block vibration signal," *Measurement Science and Technology*, vol. 19, no. 8, Article ID 085407, pp. 6686–6696, 2008.
- [43] H. Janicke, M. Bottinger, U. Mikolajewicz, and G. Scheuermann, "Visual exploration of climate variability changes using wavelet analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1375–1382, 2009.
- [44] N. Kazantzis, K. T. Chong, J. H. Park, and A. G. Parlos, "Control-relevant discretization of nonlinear systems with time-delay using Taylor-Lie series," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 127, no. 1, pp. 153–159, 2005.

- [45] E. Hossaini-asl and M. Shahbazian, "Nonlinear dynamic system control using wavelet neural network based on sampling theory," in *Proceedings IEEE International Conference on Systems, Man and Cybernetics (SMC '09)*, pp. 4502–4507, October 2009.

Research Article

Rational Formations of a Metro Train Improve Its Efficiencies of Both Traction Energy Utilization and Passenger Transport

Xuesong Feng,¹ Haidong Liu,¹ Hanxiao Zhang,¹ Baoshan Wang,¹ and Qipeng Sun²

¹ MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, No. 3 Shangyuancun, Haidian District, Beijing 100044, China

² School of Economics and Management, Chang'an University, Nan Er Huan Zhong Duan, Xi'an 710064, China

Correspondence should be addressed to Xuesong Feng; charles.x.feng@gmail.com

Received 4 July 2013; Revised 31 July 2013; Accepted 1 August 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Xuesong Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on simulations of passenger transports of two representative types of metro trains in China, this study analyzes efficiencies of energy consumption and passenger transport of a metro train in the effect of its target speed, formation scale (FS) (i.e., length and mass of the formation), relative traction capacity (RTC) (i.e., ratio of the motoring cars to all its cars), and so forth. It is found that increasing energy cost efficiency of a metro train with decreasing its target speed is evidently accelerated with reducing its RTC below 0.50 at the expense of obviously lowering its passenger transport efficiency. Moreover, if the passenger capacity of the train is sufficiently utilized, increasing its FS for the same RTC is easy to have its passenger transport efficiency improved significantly even for a meanwhile much decreased target speed with consuming energy less intensively. Therefore, metro trains in peak hours may take comparatively big FSs, relatively high target speeds, and RTCs over 0.50 to meet usually urgent and large-scale travel demands in such time. In contrast, trains in nonpeak hours ought to have small FSs, relatively low target speeds, and RTCs smaller than 0.50 for mainly avoiding energy waste.

1. Introduction

Urban traffic congestion commonly happens in China today. The urban traffic networks are characterized by slow travel speeds, long trip time, and increased vehicle queuing, which adversely affects urban mobility [1]. As a result, the urban rail transit (URT) networks have been rapidly developed in different cities of China. Till the end of 2011, the total length of the URT lines in operation has been increased from only about 500.00 kilometers (km) in 2006 to approximately 1,714.00 km in China, and by 2050 it is going to exceed 13,000.00 km [2]. With such a rapid development, issues for both energy saving and transport efficiency improvement arise on, for example, how to reasonably decide the target speed, formation Scale (FS) (i.e., length and mass of the formation), relative traction capacity (RTC) (i.e., proportion of the motoring cars to all the cars), stop frequency, and so forth of a URT train.

Many studies have been made by scholars and practitioners to improve the passenger transport efficiencies of

trains as well as to decrease their energy consumption. One of the main categories of the efforts is the optimization work through determining driving control point(s) for the transport of a train. For instance, Liu and Golovitcher [3] develop an algorithm by utilizing the optimal control theory to tell the control change points of a train in its transport process for the most energy-efficient trip in the required travel time. Wong and Ho [4] compare the capacities of different methods for searching the most suitable coasting point(s) to regulate the train service and find that transport distance between neighboring stops plays a decisive role on their convergence speeds. Kim and Chien [5] take advantage of the simulated annealing algorithm to find the optimal control patterns of a train under various track alignments for the minimum energy cost of its transport adhering to the time schedule. Ignacio and Alberto [6] mathematically proved that increasing the speed of a train on downward slopes has positive effect upon not only decreasing travel time but also reducing energy consumption for the whole trip. In contrast, another kind of the achievements tries to

interpret the effect of controlling some specific factors of a train on the improvement of its energy utilization efficiency in consideration of passenger transport time. Such factors include driving strategy [7], total mass of the cars [8], stop frequency [9], or, in other words, transport distance between neighboring stops [10], acceleration and braking performance [11], mass distribution [12], power distribution between independent engines [13], passenger boarding rate [14], target speed at each time [15], and formation scale [16, 17]. Furthermore, research on incorporating optimal driving controls and timetable adherence of multitrains to minimize their total energy consumption has been attached much importance in recent years. For example, with the aid of an event-based model, a dynamic programming method is applied by Wong and Ho [18] to devise an optimal set of dwell times and run times of trains for their energy saving and service regulation. Acikbas and Soylemez [19] propose an approach with the utilization of artificial neural networks and genetic algorithms to optimize the coasting points of trains to minimize the energy consumption for their target travel times. A genetic algorithm integrated with simulation is designed by Yang et al. [20] to seek the approximate optimal coasting control strategies of trains for the optimization of energy cost and travel time on a rail network. Cucala et al. [21] use a fuzzy linear programming model to establish the optimal schedule with the goal of energy minimization for trains in view of their uncertain delays and the behavioral responses of drivers. Li et al. [22] propose a multiobjective fuzzy scheduling model to minimize the energy consumption, carbon emission, and passenger time of multitrains.

Despite the valuable research findings of the previous studies, the question on how to decide the target speed, FS, and RTC of a train from a comprehensive perspective for its traction energy saving under the premise of meeting the passenger transport demand in each time period of the daily operation of a URT line has not been answered in an adequately convictive manner. Based on computer-aided simulations, this study analyzes the changes of the energy cost and time usage of each transport of two representative types of metro trains in China with the improvements of their target speeds by taking into account the impacts of their FSs, RTCs, and so forth to provide a key to this question from the quantificational viewpoint of train traction calculation.

The latter parts of this paper are organized as follows. Specifications of the studied trains and their rail lines are presented in Section 2. Next, the computer-aided simulation approach utilized to compute the traction energy cost (TEC) (i.e., the energy consumed by each operating condition, namely, halt, motoring, coasting, and braking) of a train for its passenger transport and the technical operation time (TOT) (i.e., the travel time spent from the startup of the train in the original station to its final stop in the terminal station) of the whole trip are explained in Section 3. Thereafter, Sections 4 and 5, respectively, analyze the changes of the TECs per 10,000 passenger-km (p-km) and TOTs per 10,000 p-km of different types of trains with improving their target speeds in the effect of various FSs and RTCs. Finally, Section 6

draws conclusions to answer the aforeexplained question and indicates some future research issues.

2. Trains and Rail Lines

The TECs and TOTs of the passenger transports of the metro trains called the MT-Type-A and the MT-Type-B on, respectively, the Line-No. 1 of City-A and Line-CP of City-B in China are comparatively studied in this work for different FSs, RTCs, and target speeds of these two types of trains. Line-No. 1 and Line-CP have their respective lengths of 17.97 km and 21.24 km, and the numbers of their stations are 16 and 7 correspondingly. As a result, the average stop spacing of Line-CP is nearly three times of the Line-No. 1's. The target speeds of the MT-Type-A and the MT-Type-B are both changed from 40.00 km per hour (km/h) to 100.00 km/h in the simulation analyses of this research in view of the usually adopted target speeds of most metro trains in China for different transport conditions. According to the most commonly applied FSs of the metro trains in China, each of these two types of metro trains consists of 4 cars, 6 cars, or 8 cars in this study and the number of their motoring cars changes from 1 to 4, 6, or 8 for corresponding FSs to have them obtain different RTCs. If a metro train is made of X motoring cars and Y trailers, the abbreviation of XMYT is used to account for the composition of this train. Some technical specifications of the cars of the TM-Type-A and the TM-Type-B are presented in Table 1. The standard passenger capacities of these cars are designed for the case that there are on average 6 passengers per square meter inside train. It is also assumed that the average mass of one passenger here is 60.00 kg in comparison to the 80.00 kg which is the average mass of one passenger with his/her hand baggage in an intercity train [23].

3. Simulation Approach

With referring to the studies of Chandra and Aqarwal [24] and Andrews [25], the computer-aided simulation approach shown in Figure 1 is applied in this work to calculate the TEC and TOT of the passenger transport a train. The transport process of the train from one stop to another is simulated for each of the successive calculation intervals which are set to be equal to 0.10 second(s) in this study. As explained by (1), the traction power and traction force of the train are assumed to be unchanged in one calculation interval. The train at a station is started up with its full traction power towards a target speed. Once initially achieving the target speed by consecutive accelerations as interpreted by (2), the train adjusts its traction power in view of its basic resistance force for the target speed according to (1) and (3) to make its speed stabilized as much as possible at the target speed. If the additional resistance force, for example, from a slope of the rail line as illuminated by (4), makes the deviation of the speed of the train from the target speed reach certain values, the train adopts the operating condition of more powerful motoring, coasting, or braking correspondingly. In order to ensure the train's safe passing through somewhere requiring a speed limit or safe stop in the next station, the train begins

TABLE I: Technical specifications of the metro cars.

Parameters	TM-Type-A		TM-Type-B	
	Motoring car	Trailer	Motoring car	Trailer
Mass	36.00 tons (t)	33.00 t	35.00 t	30.00 t
Length	22.10 meters (m)	23.69 m	19.00 m	19.50 m
Standard passenger capacity	306	318	240	250
Designed top speed	100.00 km/h		110.00 km/h	

to check whether brakes are necessary or not in a calculation interval when there is a certain distance away from there. This is determined according to the speed (v_1) of the train at present and the currently permitted maximum speed (v_2) which is computable based on the braking performance of the train and the transport distance from its current location to that rail site or the next stop. If $v_1 \geq v_2$, the train brakes to decrease its speed as soon as possible to a small value which is able to absolutely ensure the safety; if $v_1 < v_2$, the train coasts. Such a decision is made for each latter calculation interval till the train passes through the rail site in safety or stops in security in the next station. After the stop with the engine working, that is, keeping the operating condition of halt, for some time at this station, the train commences to repeat the aforeexplained process by starting up again with its target speed and full traction power till its final arrival at the terminal station.

Each type of trains has a certain unchanged traction force for its startup (which is a special operating condition of motoring) and also has its own set of operating handle positions which make its traction power change from 0 W (e.g., for the operating condition of coasting or braking) to its full traction power. The traction force of a train for its mobility after its startup is determined by both the speed and the traction power of this train. If a train adopts the operating condition of halt, its traction force is 0 N. The change of the traction force of a train from its startup at a station to its completion of the halt at the next station is interpreted by:

$$f_k^{\text{ch}} = \begin{cases} F_0, & \text{if } (k = 0), \\ \frac{P_k^{\text{ch}}}{v_{k-1}^{\text{ph}}}, & \text{if } (k = 1, 2, \dots, m), \\ 0, & \text{if } (k = m + 1, m + 2, \dots, n), \end{cases} \quad (1)$$

where f_k^{ch} is the traction force of the train positioning its operating handle at ch in the k th calculation interval, unit: N, F_0 is the traction force of the train for its startup, unit: N, P_k^{ch} is the traction power of the train positioning its operating handle at ch in the k th calculation interval, unit: W, and v_{k-1}^{ph} is the speed of the train at the end of the $(k - 1)$ th calculation interval in which the operating handle of the train is at the position of ph, unit: m/s.

The speed of the train at the end of a calculation interval is determined by its speed at the end of the previous calculation interval, the traction force of the train in this calculation interval, the resistance force (including the basic resistance force and the additional resistance force) of the train in

this calculation interval, and the mass of the train, which is explained by:

$$v_k^{\text{ch}} = v_{k-1}^{\text{ph}} + \frac{f_k^{\text{ch}} - f_k^b - f_k^a}{M} \times \Delta t, \quad (2)$$

where v_k^{ch} is the speed of the train at the end of the k th calculation interval in which the operating handle of the train is at the position of ch, unit: m/s, f_k^b is the basic resistance force in the k th calculation interval, unit: N, f_k^a is the additional resistance force in the k th calculation interval, unit: N, M is the mass of the train together with all of its passengers, unit: kg, and Δt is the equivalent length of the calculation intervals, that is, 0.10 s, in this study.

It is clearly illuminated by (3) that the basic resistance force of a train mainly because of the air and rail frictions is much concerned with its speed. When the speed of the train is 0 m/s, the basic resistance force for its startup is determined with its mass and its own resistance force intensity, but in contrast, the basic resistance force of the train after its startup is increased with its speed, which is able to be described by a quadratic equation:

$$f_k^b = \begin{cases} q \times M, & \text{if } (k = 0), \\ \alpha_0 + \alpha_1 \times (v_{k-1}^{\text{ph}}) + \alpha_2 \times (v_{k-1}^{\text{ph}})^2, & \text{if } (k = 1, 2, \dots, n), \end{cases} \quad (3)$$

where q is the resistance force intensity for the startup of the train, unit: N/kg, and α_0 , α_1 , and α_2 are the resistance coefficients for the mobility of the train (and their values are directly affected by the FSs of the train).

The additional resistance forces of a train are caused by many factors such as the ramps and curves of the rail line, the strong wind, and the very high or very low temperature. Due to the inadequate data support, only the additional resistance forces from the ramps and curves of the rail line are considered on the assumption that the mass of the train with all of its passengers is equally distributed as a thread in this research, as explained by:

$$f_k^{a,R} = m_k^R \times g \times \sin \theta, \quad (4)$$

$$f_k^{a,C} = \frac{0.60 \times g \times m_k^C}{Ra}, \quad (5)$$

where $f_k^{a,R}$ is the additional resistance force from the rail's ramp R in the k th calculation interval, unit: N, m_k^R is the mass of the car(s) with its/their passengers on the rail's ramp R in

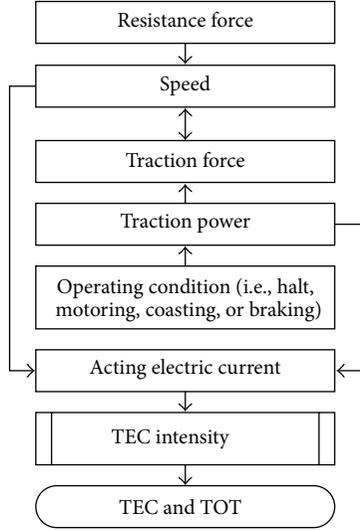


FIGURE 1: Simulation approach for calculations of TEC and TOT.

the k th calculation interval, unit: kg, g is the gravitational acceleration, unit: N/kg, θ is the angle between the level and the slope of the ramp R , unit: degree, $f_k^{a,C}$ is the additional resistance force from the rail's curve C in the k th calculation interval, unit: N, m_k^C is the mass of the car(s) with its/their passengers on the rail's curve C in the k th calculation interval, unit: kg, and R_a is the numerical value of the radius of the curve C , unit: m.

On the basis of both traction power and average speed of a train for a calculation interval, the average value of the acting electric current of the train is obtained in the consideration of the truly utilized electric current for train traction and the wasted electric current in this calculation interval. Therefore, the TEC intensity of a train for each calculation interval is determined by:

$$\text{TEC}_k^I = \bar{U}_k \times \bar{I}_k, \quad (6)$$

where TEC_k^I is the TEC intensity of the train for the k th calculation interval, unit: W, \bar{U}_k is the average voltage provided for the transport of the train in the k th calculation interval, unit: V, and \bar{I}_k is the average value of the acting electric current of the train in the k th calculation interval, unit: A.

As a result, the TECs of all of the calculation intervals from the startup of a train at its original station to its stop at the terminal station are computed by utilizing the TEC intensity of the train for each of the calculation intervals and summed into the TEC of the trip, as clarified by (7), and the TOT of the train between these two stops is computable by summation of all calculation intervals of the transport, as explained by (8):

$$\text{TEC}_{ij}^v = \frac{\sum_{k=0}^N (\text{TEC}_k^I \times \Delta t)}{1000 \times 3600}, \quad (7)$$

$$\text{TOT}_{ij}^v = \frac{\sum_{k=0}^N \Delta t}{3600} = \frac{(N+1) \times \Delta t}{3600}, \quad (8)$$

where TEC_{ij}^v is the TEC of the train with the target speed of v from station i to station j , unit: kilowatt hours (kWh), and TOT_{ij}^v is the TOT of the train with the target speed of v from station i to station j , unit: hours (h).

4. Energy Cost Efficiency Evaluation

The TEC per 10,000 p-km of a train with the target speed of v between different stops, as explained by (9), is utilized to comparatively evaluate the energy utilization efficiencies of this train with different target speeds, FSS, and RTCs to, respectively, complete the same transport mission:

$$e_{ij}^v = \frac{\text{TEC}_{ij}^v}{\sum_{s=0}^{j-i-1} P_{ij}^v \times R_{(i+s)(i+s+1)}^v \times D_{(i+s)(i+s+1)}^v}, \quad (9)$$

where e_{ij}^v is the TEC per 10,000 p-km of the train with the target speed of v from station i to station j , unit: kWh/10,000 p-km, P_{ij}^v is the standard passenger capacity of the train with the target speed of v from station i to station j , $R_{(i+s)(i+s+1)}^v$ is the utilization ratio of the standard passenger capacity of the train with the target speed of v from station $(i+s)$ to station $(i+s+1)$, and $D_{(i+s)(i+s+1)}^v$ is the transport distance of the train with the target speed of v from station $(i+s)$ to station $(i+s+1)$, unit: 10,000 km.

The travel demand of the passengers for each station along a URT line usually has little change in a certain daily time. In other words, the p-km completed by one trip of a train in one service time of the URT line is relatively fixed. In view of the extremely high travel demands of the passengers of Line-No. 1 and Line-CP in their busy times, the utilization ratio of the standard passenger capacities of the metro trains here is supposed to be unchanged as 100.00% for each stop spacing in peak hours. As for one transport of an MT-Type-A from the 1st station to the 16th station of Line-No. 1, the changes of its TECs per 10,000 p-km with the increase of the target speed for different FSS and RTCs are revealed in Figure 2. Such TEC changes of the MT-Type-B are displayed in Figure 3 for the transport from the 1st station to the 7th station of Line-CP. The time of the operating condition of halt of a train in each of the stations between the original and terminal stations of a metro line is set to be 0.10 s in the simulation studies to simplify the TEC and TOT calculations.

It is clearly shown in Figures 2 and 3 that the increase of the TEC per 10,000 p-km of a metro train with raising its target speed is accelerated with the improvement of its RTC for a certain utilization ratio of the standard passenger capacity of the train. When the RTC is smaller than 0.50, such accelerations are obvious. However, if the RTC exceeds 0.50, this accelerative effect becomes apparently small in comparison. Moreover, as for the same RTC, different FSS of a metro train have little impact on the change of its TEC per 10,000 p-km with increasing its target speed for the same utilization ratio of its standard passenger capacity. Now it is able to be confirmed that decreasing the RTC of a metro train below approximately 0.50 is able to evidently increase the efficiency of its traction energy utilization especially for a relatively high target speed. In addition, it is also proved that a

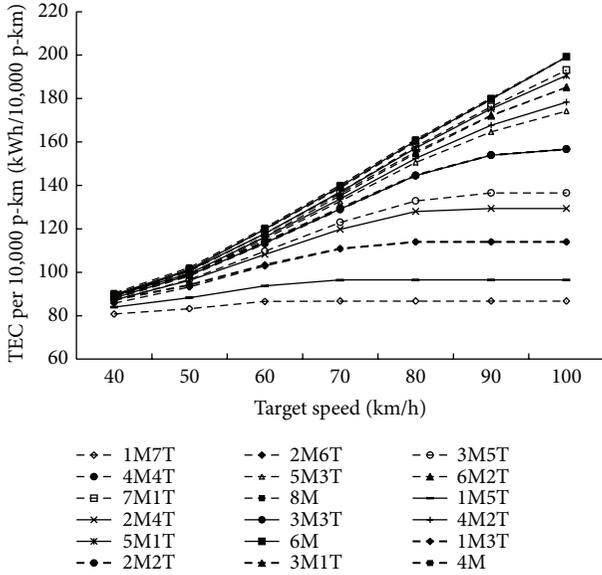


FIGURE 2: TECs of the transports of the MT-Type-A.

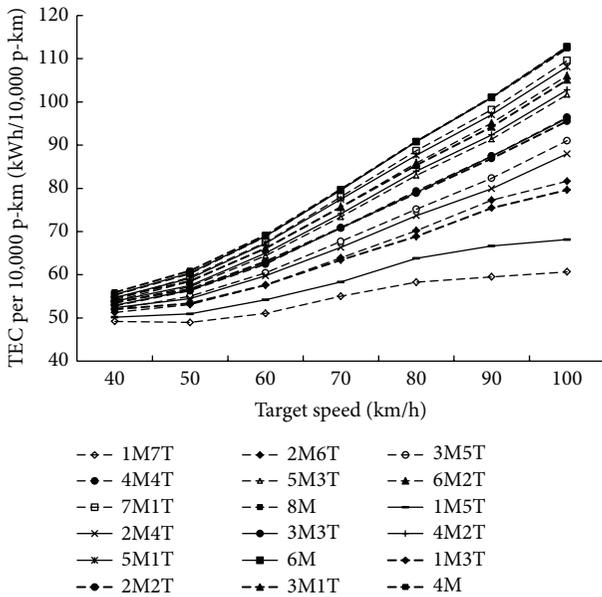


FIGURE 3: TECs of the transports of the MT-Type-B.

short stop spacing may make the actually reached maximum speed of a metro train extremely lower than a fairly high target speed [26] in particular for a comparatively small RTC, as reflected in Figure 2.

5. Passenger Transport Efficiency Analysis

The TOT per 10,000 p-km of a train with a target speed of v between different stops is defined by (10) and used to estimate the passenger transport efficiencies of this train with

various target speeds, FSs, and RTCs to accomplish the same transport task in a respective manner:

$$t_{ij}^v = \frac{\text{TOT}_{ij}^v}{\sum_{s=0}^{j-i-1} D_{ij}^v \times R_{(i+s)(i+s+1)}^v \times D_{(i+s)(i+s+1)}^v}, \quad (10)$$

where t_{ij}^v is the TOT per 10,000 p-km of the train with the target speed of v from station i to station j , unit: h/10,000 p-km.

The decreases of the TOTs per 10,000 p-km of the afore-explained transports by the MT-Type-A and the MT-Type-B on, respectively, Line-No. 1 and Line-CP with the increases of the target speeds of these two types of metro trains are presented in Figures 4 and 5 in a corresponding way. It is indicated in each of these two figures that decreasing the RTC of a metro train below 0.50 costs obviously additional TOT per 10,000 p-km for the unchanged total number of its cars with a fixed average utilization ratio of their standard passenger capacities, which becomes more manifesting for a comparatively high target speed. That is to say the aforeillustrated distinct improvement of the traction energy utilization efficiency of a metro train by decreasing its RTC below 0.50 is at the conspicuous expense of the decrease of its passenger transport efficiency especially for a relatively high target speed. In contrast, the decrease of the RTC over 0.50 does not increase much of the TOT per 10,000 p-km of the train. Furthermore, it is also demonstrated in both Figures 4 and 5 that a long formation of a metro train with a certain RTC has an obviously lower TOT per 10,000 p-km than the one of a short formation of this train for the same utilization ratio of the standard passenger capacities of different formations, even if the long formation of the train has a much lower target speed than its short formation's. This is because of the little difference of the transport time and the big difference of the completed p-km of different FSs of the train for the same trip. Now it is clarified that, in comparison to the efficiencies of traction energy utilization and passenger transport of a metro train with a short formation and meanwhile a very high target speed, a long formation of this train with the same RTC and a relatively much lower target speed for the same transport action is able to easily have less TOT and TEC per unit transport for the same utilization ratios of the passenger capacities of different FSs.

6. Conclusions

It is found that the increase of the traction energy cost intensity of a metro train with improving its target speed for a certain utilization ratio of its standard passenger capacity is obviously accelerated with raising its RTC below 0.50 in comparison to the increasing acceleration of its energy cost intensity when the RTC exceeds 0.50. At the same time, different FSs of the train have little influence upon its TEC per unit transport for the same RTC. However, such an apparently accelerated increase of the energy cost intensity of the train for the same number of its cars evidently promotes the improvement of its passenger transport efficiency with the increase of its target speed. Moreover, as for the same

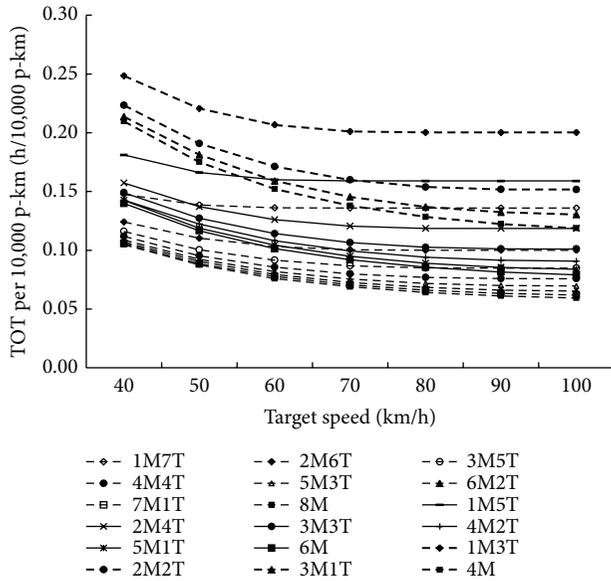


FIGURE 4: TOTs of the transports of the MT-Type-A.

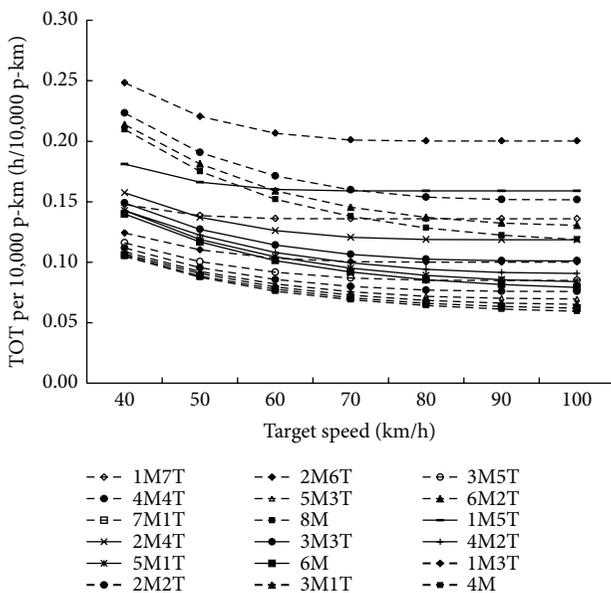


FIGURE 5: TOTs of the transports of the MT-Type-B.

RTC, a long formation of the train is able to easily transport passengers efficiently owing to its fairly utilized big passenger capacity even for a comparatively very low target speed which also results in less TEC per unit transport.

Therefore, on the condition of a rational general cost including labor cost and rail car purchase cost, if a metro line has its car depot(s) capable of changing the FSs and RTCs of its trains efficiently in technique and/or providing sufficient space for enough trains with various FSs and RTCs, trains with different FSs and RTCs should be applied flexibly according to the changeable travel demands of this line for different time periods in its daily operation. For instance,

trains running on the metro line in rush hours ought to have long formations and their RTCs may be bigger than 0.50 for a relatively high target speed to meet the usually urgent and ordinarily big travel demands in such time with the minimum energy cost as much as possible. In nonpeak hours, most of the trains should take short formations for the improvements of the utilization ratios of their passenger capacities. Meanwhile, relatively low target speeds and the RTCs smaller than 0.50 ought to be adopted to increase their traction energy cost efficiencies to the greatest extent at the premise of ensuring necessary efficiencies of their transport services.

In this research, only the transports of two representative types of the metro trains, respectively, on two metro lines in China have been analyzed. More transport operations of other kinds of trains on various URT lines are necessary to be studied to further validate the conclusions of this work. In addition, the impacts of more factors such as aerodynamic resistance, regenerative braking, and interaction between multitrains in their tracking operations, on the efficiencies of energy utilization and passenger transport also need to be systematically explored in future research, in view of the instability of a system when its behavior changes seriously during the system operation [27].

Acknowledgments

This research is supported by the National Natural Science Foundation of China (71201006, 71131001), the National Basic Research Program of China (2012CB725406), and the Fundamental Research Funds for the Central Universities (2011JBM248).

References

- [1] W. Wang, W. Zhang, H. Guo, H. Bubb, and K. Ikeuchi, "A safety-based approaching behavioural model with various driving characteristics," *Transportation Research C*, vol. 19, no. 6, pp. 1202–1214, 2011.
- [2] Research Group of the Annual Report of China Urban Mass Transit, *2011 Annual Report of China Urban Mass Transit*, Beijing Jiaotong University Press, Beijing, China, 2012.
- [3] R. Liu and I. M. Golovitcher, "Energy-efficient operation of rail vehicles," *Transportation Research A*, vol. 37, no. 10, pp. 917–932, 2003.
- [4] K. K. Wong and T. K. Ho, "Coast control for mass rapid transit railways with searching methods," *IEE Proceedings-Electric Power Applications*, vol. 151, no. 3, pp. 365–375, 2004.
- [5] K. Kim and S. I.-J. Chien, "Simulation-based analysis of train controls under various track alignments," *Journal of Transportation Engineering*, vol. 136, no. 11, pp. 937–948, 2010.
- [6] G.-F. Ignacio and G.-A. Alberto, "Can high-speed trains run faster and reduce energy consumption?" *Procedia*, vol. 48, pp. 827–837, 2012.
- [7] P. Lukaszewicz, *Energy consumption and running time for trains: modelling of running resistance and driver behavior based on full scale testing [Ph.D. dissertation]*, Royal Institute of Technology, Stockholm, Sweden, 2001.

- [8] Institut für Energieund Umweltforschung Heidelberg GmbH, "Energy savings by light-weighting," Final Report, International Aluminium Institute, Heidelberg, Germany, 2003.
- [9] E. Lindgreen and S. C. Sorenson, *Simulation of Energy Consumption and Emissions from Rail Traffic Evaluation*, Technical University of Denmark, Lyngby, Denmark, 2005.
- [10] X. Feng, "Optimization of target speeds of high-speed railway trains for traction energy saving and transport efficiency improvement," *Energy Policy*, vol. 39, no. 12, pp. 7658–7665, 2011.
- [11] Y. V. Bocharnikov, A. M. Tobias, C. Roberts, S. Hillmans, and C. J. Goodman, "Optimal driving strategy for traction energy saving on DC suburban railways," *IET Electric Power Applications*, vol. 1, no. 5, pp. 675–682, 2007.
- [12] M. Chou and X. Xia, "Optimal cruise control of heavy-haul trains equipped with electronically controlled pneumatic brake systems," *Control Engineering Practice*, vol. 15, no. 5, pp. 511–519, 2007.
- [13] S. Lu, S. Hillmans, and C. Roberts, "A power-management strategy for multiple-unit railroad vehicles," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 406–420, 2011.
- [14] X. Feng, "Energy cost of high-speed railway trains in the effect of utilization ratios of passenger seats," in *Proceedings of International Conference on Information, Services and Management Engineering (ISME '11)*, vol. 1, pp. 133–136, Beijing, China, December 2011.
- [15] B.-R. Ke, C.-L. Lin, and C.-W. Lai, "Optimization of train-speed trajectory and control for mass rapid transit systems," *Control Engineering Practice*, vol. 19, no. 7, pp. 675–687, 2011.
- [16] K. Kim and S. I.-J. Chien, "Optimal train operation for minimum energy consumption considering track alignment, speed limit, and schedule adherence," *Journal of Transportation Engineering*, vol. 137, no. 9, pp. 665–674, 2011.
- [17] X. Feng, J. Feng, K. Wu, H. Liu, and Q. Sun, "Evaluating target speeds of passenger trains in China for energy saving in the effect of different formation scales and traction capacities," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 621–626, 2012.
- [18] K. K. Wong and T. K. Ho, "Dwell-time and run-time control for DC mass rapid transit railways," *IET Electric Power Applications*, vol. 1, no. 6, pp. 956–966, 2007.
- [19] S. Acikbas and M. T. Soylemez, "Coasting point optimisation for mass rail transit lines using artificial neural networks and genetic algorithms," *IET Electric Power Applications*, vol. 2, no. 3, pp. 172–182, 2008.
- [20] L. Yang, K. Li, Z. Gao, and X. Li, "Optimizing trains movement on a railway network," *Omega*, vol. 40, no. 5, pp. 619–633, 2012.
- [21] A. P. Cucala, A. Fernández, C. Sicre, and M. Domínguez, "Fuzzy optimal schedule of high speed train operation to minimize energy consumption with uncertain delays and driver's behavioral response," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 8, pp. 1548–1557, 2012.
- [22] X. Li, D. Wang, K. Li, and Z. Gao, "A green train scheduling model and fuzzy multi-objective optimization algorithm," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 2063–2073, 2013.
- [23] International Union of Railways, *UIC Code 556 OR: Loadings of Coach Bodies and Their Components*, International Union of Railways, Paris, France, 3rd edition, 1990.
- [24] C. Chandra and M. M. Aqarwal, *Railway Engineering*, Oxford Higher Education, New York, NY, USA, 2008.
- [25] H. I. Andrews, *Railway Traction: The Principles of Mechanical and Electrical Railway Traction*, vol. 5 of *Studies in Mechanical Engineering*, Elsevier Scientific, New York, NY, USA, 1986.
- [26] X. Feng, B. Mao, X. Feng, and J. Feng, "Study on the maximum operation speeds of metro trains for energy saving as well as transport efficiency improvement," *Energy*, vol. 36, no. 11, pp. 6577–6582, 2011.
- [27] W. Wang, X. Jiang, S. Xia, and Q. Cao, "Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation," *Safety Science*, vol. 48, no. 10, pp. 1248–1262, 2010.

Research Article

Crossing Reliability of Electric Bike Riders at Urban Intersections

Huan Mei, Yang Xiaobao, and Jia Bin

MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Yang Xiaobao; yangxb@bjtu.edu.cn

Received 13 June 2013; Accepted 31 July 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Huan Mei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a crossing reliability model of electric bike riders at urban intersections using survival analysis approach. Riders' crossing behavior was collected by video cameras. Waiting times in the red-light phase were modeled by reliability-based model that recognizes the covariate effects. Three parametric models by the exponential, Weibull, and log-logistic distributions were proposed to analyze when and why electric bike riders cross against the red light. The results indicate that movement information and situation factors have significant effects on riders' crossing reliability. The findings of this paper provide an important demonstration of method and an empirical basis to assess crossing reliability of electric bike riders at the intersection.

1. Introduction

Nonmotorized vehicles (i.e., mainly regular bicycles and electric bikes) are one of the most popular means of transportation in some Asian developing countries, such as Vietnam, Cambodia, and China. Even in developed countries, cycling travel is recognized as low energy consumption, healthy to the users and do not damage the health of others. For example, people who go to work by bike would obtain material reward from the government in Korea [1]. In London, in order to encourage people to travel by bike, the local government carried out the cycling revolution [2].

In recent years, the electric bike has entered people's life. Because of electric bike is labor-saving and speedy, it has emerged as a popular mode of transportation in many large cities in China [3]. The number of Chinese electric bikes was about 140 million in 2012 [4]. Electric bikes in China are defined as electric two wheelers with relatively low speeds and weights compared to a motorcycle. Both bicycle-style electric bikes (with functioning pedals) and scooter-style electric bikes (with many of the features of gasoline scooters) are classified as bicycles and are given access to bicycle infrastructure (see Figure 1).

However, the growing popularity of cycling traffic also entails safety concerns as observed in accident and injury

statistics. Traffic accident proportion of cyclists has always been high because cyclists are vulnerable groups in the traffic conflict. In 2010, 4616 regular bicyclists were killed and 14,283 were seriously injured in road accidents, representing 7.1% of all traffic fatalities and 5.6% of injuries [5]. With the rapidly increasing number of electric bikes, more and more people pay much concerned about traffic security problems involved electric bikes. In 2004, 589 electric bike riders were died and 5295 were seriously injured in road accidents [6]. In 2010, the corresponding figures increased to 4029 and 20,311, respectively, representing 6.2% of all traffic fatalities and 8.0% of injuries [5].

Accident analysis reveals that over 60% of cyclist fatal crashes were caused by violation of traffic rules [5]. Red-light crossing is a typical type of rule violation behavior. Because of the poor law enforcement and peoples' low safety awareness, violation behavior in the red-light period is rather prevalent and represents a substantial safety problem in Chinese urban intersections [7]. Particularly, electric bike riders with relatively high speed are much more likely to increase the risk of traffic incident.

So far, many scholars have studied red-light crossing behaviors, but many focused on motorized vehicles [8, 9] and pedestrians [10–13]. Unfortunately, only a few studies have investigated bicyclists' red-light crossing behavior, much less



FIGURE 1: Bicycle-style electric bike at the left and scooter-style electric bike at the right.

about electric bike riders. Johnson et al. analyzed the bicycle violation behavior at intersections in Melbourne [14]. They found out that the rate of red-light crossing before noon was 3% and 11% in the afternoon. Johnson et al. collected 4225 bicyclists who arrived at the intersections at red-light time, and their result indicated that 6.9% bicyclists violated traffic rules [15]. Wu et al. used logistic model to analyze how gender, age, and conformity behavior affected the cyclist's violation, and classify red-light crossing behavior to three types: obey the rules, risk taking, opportunistic [7]. Johnson et al. obtained the data through the Internet and used TPB (theory of planned behavior) to investigate the bicyclists' red-light crossing behavior and attitude and the influencing factors in Australia [16]. Their result showed that 37.3% of the riders violated traffic rules, and the violation rate of males was higher than females, the young was higher than the old, and a rider who was accident-free may violate them easier.

Most of the existing researches on cyclists' crossing behavior used logistic model to analyze the violating probability and its influence factors. And little effort was focused on the red-light crossing behavior of electric bike riders. In this paper, we use survival analysis method to study electric bike riders' crossing reliability at the signalized intersection. Survival analysis is a common topic in many areas including biomedical, engineering, and social sciences. In the transportation field, it has been applied to describe a number of time-related issues including activity duration [17, 18], traffic accidents [19], and vehicle travel time [20, 21]. Hazard-based duration models of survival analysis have an advantage in that it allows the explicit study of the relationship between duration time and the explanatory variables [22–24]. More importantly, survival analysis models can deal with not only uncensored data but also censored data. For example, the exact waiting duration reflecting cyclist endurance cannot be observed if cyclists wait until the permission of traffic rules. This is the very reason why survival analysis method is chosen to analyze riders' crossing behavior. The empirical data, which were obtained by video cameras, are modeled by three typical parametric hazard functions. Both crossing reliability and waiting times of electric bike riders under various conditions are calculated and the covariate effects are quantified. The finding of this paper can explain when and why electric bike riders violate traffic light at the intersection. It is hoped that the results can help to improve the planning and designing of signalized intersections in developing countries.

2. Method

Reliability is the probability that a system or component will perform its required function under stated conditions for a specified period of time. Accordingly, crossing reliability of electric bike riders can be defined as the probability that an electric bike rider obey the traffic light after a specified waiting time at the signalized intersection.

Let T denote a nonnegative random variable representing the failure time or time-to-failure. Mathematically, the reliability function $R(t)$ is the probability that a system will be successfully operating without failure in the interval from time 0 to time t ,

$$R(t) = P(T > t), \quad t > 0. \quad (1)$$

In this paper, it is assumed that an electric bike rider violate the traffic light when his/her waiting duration time is T . $R(t)$ is the crossing reliability of electric bike riders. It is the probability that a rider, who arrives at the intersection in the red-light period, will not violate the traffic light in his/her waiting duration time from 0 to t .

The failure probability, or unreliability of riders' crossing behavior, is then

$$F(t) = 1 - R(t) = P(T \leq t), \quad (2)$$

which is known as the distribution of T .

If the random variable T has a density function $f(t)$, then

$$R(t) = \int_t^{\infty} f(x) dx. \quad (3)$$

The density function can be mathematically described as $\lim_{\Delta t \rightarrow 0} P(t \leq T < t + \Delta t)$. This can be interpreted as the probability that the failure/violation time will occur between time t and the next interval of waiting time, $t + \Delta t$. The three functions, $R(t)$, $F(t)$, and $f(t)$ are closely related to one another. If any of them is known, all the others can be determined.

The waiting time of an electric bike rider in the red-light period can be regarded as the waiting duration that starts when a cyclist arrives at the intersection in the red period and ends when the rider starts to cross the intersection. Therefore, hazard-based duration model in survival analysis can be adopted to calculate the probability of rider violation

with different waiting durations. In survival analysis, $R(t)$ is also called survival function or survivor probability. The survival function is defined to be the probability that the waiting time of a rider in a red-light is longer than a specific time, t .

Another important function in survival analysis is the hazard function, $h(t)$. The hazard function in this paper is the instantaneous rate at which the waiting duration will end in an infinitesimally small time period, Δt , after time t , given that the duration time has lasted to time t

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t \mid T > t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t \times P(T > t)} = \frac{f(t)}{R(t)}. \end{aligned} \quad (4)$$

The importance of the hazard function is that it indicates the changing rate in the red-light violating behavior over the waiting period of a rider. For example, two riders may have the same crossing reliability at a specific point in time, but the hazard rate curves can be very different.

Note that crossing reliability of electric bike riders is influenced by various factors. The influential factors can be defined as a vector of explanatory variables, $\mathbf{x} = (x_1, x_2, \dots, x_m)'$. To accommodate the effects of these influential factors, a parametric hazard model or a semiparametric hazard model may be adopted. If little or no knowledge of the functional form of the hazard is available, a semiparametric approach for modeling the hazard function is convenient. If theoretical support for a parametric shape is available, using a particular parametric distributional form for the hazard function will be appropriate. Both semiparametric and fully parametric hazard-based models have been widely cited in the literature. For detailed discussion of different approaches for duration models see Bhat [17], Lee and Wang [25]. Here, the primary objective of this paper is to examine the crossing reliability and consider the application of the model. A parametric hazard approach is adopted because its hazard function can be chosen flexibly. Several typically parametric distributions are exponential, Weibull, and log-logistic. These common distributions are summarized below.

The exponential distribution is suitable for modeling data with constant hazard. The exponential hazard and survivor functions are

$$\begin{aligned} h(t) &= \lambda \text{ (constant hazard)}, \\ R(t) &= \exp(-\lambda t). \end{aligned} \quad (5)$$

The model is implemented by parameterizing $\lambda_j = \exp(-X_j \beta)$ when the influenced covariates are considered. Let $h_i(t, \lambda_i)$ and $h_j(t, \lambda_j)$ be the hazards of individuals i and j ; the hazard ratio of these two individuals is

$$\frac{h_i(t, \lambda_i)}{h_j(t, \lambda_j)} = \frac{\lambda_i}{\lambda_j} = \exp \left[-\sum_k \beta_k (x_{ki} - x_{kj}) \right]. \quad (6)$$

This ratio is dependent only on the differences of the covariates of the two individuals and the coefficients. It does not depend on the time t .

The Weibull distribution is suitable for modeling data with monotone hazard rates that either increase or decrease exponentially with time. The Weibull hazard and survivor functions are

$$\begin{aligned} h(t) &= \lambda \gamma (\lambda t)^{\gamma-1}, \\ R(t) &= \exp(-(\lambda t)^\gamma). \end{aligned} \quad (7)$$

The model is implemented by parameterizing $\lambda_j = \exp(-X_j \beta)$, where γ is an ancillary parameter to be estimated from the data. Note that the hazard increases with duration time if $\gamma > 1$, decreases if $\gamma < 1$, and reduces to exponential if $\gamma = 1$. Similar to the exponential distribution model, the hazard ratio of the Weibull distribution model is also not time dependent.

The log-logistic model assumed that the natural logarithm of time follows a logistic distribution. The log-logistic hazard and survivor functions are

$$\begin{aligned} h(t) &= \frac{\lambda \gamma (\lambda t)^{\gamma-1}}{\{1 + (\lambda t)^\gamma\}}, \\ R(t) &= \{1 + (\lambda t)^\gamma\}^{-1}. \end{aligned} \quad (8)$$

The model is implemented by parameterizing $\lambda_j = \exp(-X_j \beta)$ and treating the scale parameter γ as an ancillary parameter to be estimated from the data. Unlike the exponential and Weibull distributions, the log-logistic distribution is indicated for data exhibiting non-monotonic hazard rates. When $\gamma > 1$, the hazard initially increases and then decreases with duration time. When $0 < \gamma \leq 1$, the hazard decreases with duration time. The hazard ratio of the log-logistic distribution model is time dependent.

In order to keep the linear form of the covariates, three new terms can be defined:

$$y_i = \log t_i = \beta X + \varepsilon_i \sigma, \quad (9)$$

$$\varepsilon = \frac{(\log t - \beta X)}{\sigma}, \quad (10)$$

$$\gamma = \frac{1}{\sigma}. \quad (11)$$

The likelihood and log-likelihood functions for N observations on y_i and right censoring indicator δ_i are

$$L = \prod_i \left[\sigma^{-1} f(\varepsilon_i) \right]^{\delta_i} [R(\varepsilon_i)]^{1-\delta_i}, \quad (12)$$

$$\log L = \sum_i \left[\delta_i (-\log \sigma + \log f(\varepsilon_i)) + (1 - \delta_i) \log R(\varepsilon_i) \right]. \quad (13)$$

Log-likelihood functions can be maximized using any standard method (e.g., Newton; BFGS).

3. Data

The field observation approach was used in this study. It has been widely used to study pedestrian and cyclist behaviors

TABLE 1: Covariates selection and explanation.

Covariate	Type	Explanation
AGE (age group)	Continuous variable	1 if under 20, 2 if 20–29, 3 if 30–39, 4 if 40–49, 5 if 50–59, and 6 otherwise
GEN (gender)	Binary indicator	1 if male, 0 female
WN (waiting number)	Continuous variable	The number of other cyclists that are waiting for a green light when arrives
CN (crossing number)	Continuous variable	The number of other cyclists that are crossing against the red light when arrives
TC (twice crossing)	Binary indicator	1 if twice crossing behavior, 0 otherwise
MV (motor vehicle volume)	Continuous variable	Motor vehicle volume per lane per min in red-light phase when the rider arrives
LT (left turn phase)	Binary indicator	1 if an electric bike rider violates traffic rules in the left-turn period of motorized vehicles, 0 otherwise
YL (yellow light)	Binary indicator	1 if an electric bike rider violates traffic rules in the yellow light period, 0 otherwise

in actual traffic situations [7, 13, 26, 27]. For example, Tiwari et al. used video recording to analyze pedestrians' violating behavior in signalized intersections in India [13]. Yao et al. also used video cameras to study the behavior characteristics of pedestrian crowd weaving flow in transport terminal [27]. Wu et al. used video recordings to examine urban commuter cyclists' red-light running behaviors in Beijing [7]. Wang et al. used video data to study pedestrians' crossing behavior in red-light period in Beijing [26].

3.1. Site Characteristics. A cross-sectional observational study was conducted at five signalized intersections in Beijing. Three criteria were used to select the observational sites. First, the selected sites should represent the typical intersection design characteristics and traffic conditions of urban areas in Beijing. Second, the selected intersections should have similar characteristics involved geometrics, traffic conditions, traffic control, and the absence of pointsmen. In addition, there have to be a reasonably high number of electric bike traffic during the observation period.

3.2. Data Collection and Processing. Video cameras were used to collect data of electric bike riders' crossing behaviors at signalized intersections. The cameras were carefully placed so that the road users were unaware that they were being observed. The data collection was conducted on weekdays during daylight hours (i.e., 8:00 a.m. to 5:30 p.m.) in good weather conditions.

All road users who entered the intersection were recorded on video, but only the riders arriving in red-light phases were coded. In addition, only the straight-going riders crossing through the intersection were recorded. Left-turners and right-turners of electric bike were excluded because of the limited field of view of the cameras. The waiting duration was from the time a rider arrived at the stop line to the time he/she began to cross. It can be classified into two kinds: uncensored data and censored data. Uncensored data is defined as the waiting duration which ends within the red-light period (violating crossing). Otherwise, the waiting duration is called

censored data as long as it ends within the green light period (normal crossing). For censored data, it is unknown about his/her exact maximum endurance of waiting time.

Considering the previous researches and intuitive arguments regarding the influential factors of cyclist crossing behavior, three sets of variables were coded (see Table 1). The first set described the riders' individual characteristics, including gender and age. The second set of variables focused on the riders' movement information, including twice crossing behavior, the times of arrival at and departure from the stop line, the time when crossing is completed, the status of the traffic light at each of three times, and whether the violating behavior occurs in the left-turn period of motorized vehicles or the yellow-light period. The last set of variables of concern were situation factors, including the red-light cycle length, the number of riders waiting upon arrival, the number of riders crossing against the red-light, and crossing traffic volume (i.e., the number of motorized vehicles that crossed the intersection from either direction during the time of the red-light cycle when the rider arrives).

4. Results

A total of 8 h of video recordings (approximately 1.5 h for each site) were collected and 1946 crossing events of cyclists were observed. Demographics and behavioral data were coded only for electric bike riders during the red-light phases. Thus, a total of 312 (16%) valid observations were obtained (see Table 2).

4.1. Descriptive Statistics. The proportions of red-light crossing of electric bike riders in each sub-group were presented in Table 2. Of all the electric bike riders we observed, the number of males was larger than females (84.3% versus 15.7%). And most of the electric bike riders are young and middle-aged persons; the respective rates are 26.5% and 67.6%. The old riders only occupied 6.1%.

The overall proportion of electric bike riders who crossed against the red-light was 75.4%. Male electric bike riders

TABLE 2: Red-light crossing rates and waiting times by each subcategory.

	Violating crossing		Normal crossing		Overall	
	Rates	Waiting time (s)	Rates	Waiting time (s)	Rates	Waiting time (s)
Gender						
Male	76.9%	26.11	23.1%	64.88	84.3%	35.06
Female	67.3%	22.11	32.7%	69.14	15.7%	37.46
Age group						
Young (<30)	80.7%	24.83	19.3%	82.57	26.5%	35.96
Middle (30–50)	74.4%	26.67	25.6%	59.46	67.6%	35.12
Old (>50)	63.2%	13.78	36.8%	75.97	6.1%	36.69
Overall	75.4%	25.55	24.6%	65.76	100%	35.44

TABLE 3: Parameter regression values of the influence variables.

Parameter	Exponential model			Weibull model			Log-logistics model	
	Coef.	Haz. ratio	P value	Coef.	Haz. ratio	P value	Coef.	P value
Constant	2.960	0.052	<0.001	1.009	0.364	0.084	1.449	0.359
SITE	0.145	0.865	0.077	0.044	0.957	0.543	0.012	0.951
AGE	−0.065	1.067	0.424	−0.055	1.057	0.485	−0.220	0.329
GEN	0.012	0.988	0.954	0.040	0.961	0.841	0.197	0.737
WN	0.041	0.960	0.018	0.064	0.938	<0.001	0.266	<0.001
CN	−0.134	1.143	<0.001	−0.089	1.094	<0.001	−0.382	<0.001
TC	−0.206	1.229	0.254	−0.331	1.392	0.049	−1.151	0.017
MV	0.189	0.828	0.001	0.146	0.864	0.006	0.429	0.006
LT	−1.166	3.209	<0.001	−0.95	2.585	<0.001	−2.431	<0.001
YL	−0.737	2.090	0.002	−0.882	2.415	<0.001	−3.065	<0.001
γ	1.000			0.408			0.514	
$\log L$		−831.6			−656.6			−671.2

were found to be more likely to violate traffic rules at signal intersections than females (76.9% versus 67.3%). Red-light violation rates also differed by age group. Young- and middle-aged riders were more likely to run against a red-light than the old ones (80.7% and 74.4% versus 63.2%). The result is consistent with the research in pedestrian crossing behavior [13, 23, 26].

The average waiting time of all samples was 35.44 seconds. The average waiting time of the violating crossing was 25.55 seconds, while the average waiting time of the normal crossing is 65.76 seconds. The maximum waiting duration was 161 seconds while the minimum was 0 second. The latter means people cross the street without any waiting time. As shown in the last column of Table 2, the overall waiting times of electric bike riders vary from 35.06 to 37.46 seconds in different genders, and from 35.12 to 36.99 seconds in different age groups. Therefore, there is no significant difference of the overall waiting times in gender and age group.

This descriptive statistic cannot reflect the exact waiting behavior due to the neglect of the censored data. The estimation of the waiting duration with censored data will be discussed later.

4.2. Parameter Regression Analysis. Table 3 shows the estimated parameters of the parametric hazard models in predicting red-light crossing behavior of electric bike riders. Three models show perfect overall goodness-of-fit indicated by the log-likelihood values. Among them, the Weibull distribution is considered the best to describe riders' crossing behavior. The hazard ratios of both the exponential and Weibull models are not time-dependent. According to (6), the hazard ratio is $\exp(-\beta_i)$ times if the i th covariate increases by one unit. However, the risk of the log-logistic model cannot be calculated by this way, since it is time dependent.

From the estimated results, the site variable is insignificant, which means the selected intersections have similar traffic and geometric characteristics. Besides, gender and age group have low significant level. It is partly because of the low proportions of female (15.7%) and old people (6.1%) in the sample. Most of other covariates are statistically significant at the 0.10 level of significance. These covariates are significantly related to violation behavior.

For the specific parameter estimation in Table 2, based on (9), the positive parameters have a positive effect on waiting time of riders (e.g., WN and MV). Therefore, an increase

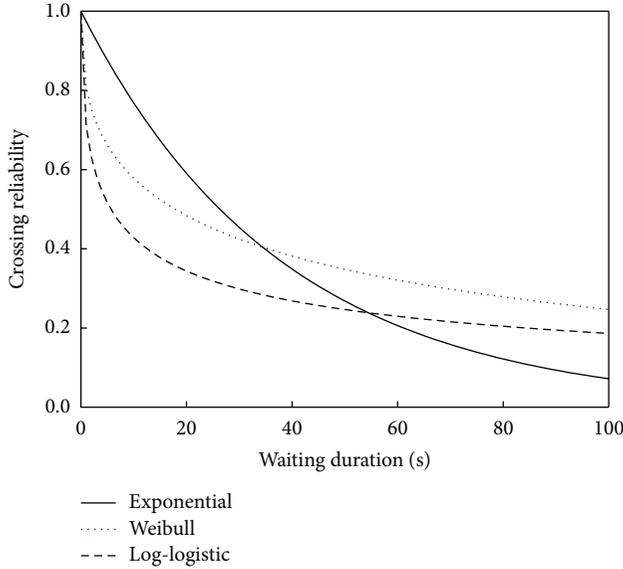


FIGURE 2: Crossing reliability with waiting duration in different parametric models.

in the corresponding covariates can increase the waiting duration or decrease the risk rate. For example, an electric bike rider is willing to wait longer time if many other cyclists are waiting when he/she arrives at the intersection. Also, a rider is likely to wait to reduce the collision risk in the presence of heavy traffic.

On the contrary, the negative effect means that an increase in the corresponding covariates can decrease the waiting duration or increase the risk rate (e.g., CN, TC, LT, and YL). For example, an electric bike rider may follow to violate rules when other cyclists are crossing against the red-light. A rider of twice crossing has higher hazard and shorter waiting time. Besides, a rider is more likely to violate rules during the left-turn period of motorized vehicles and the yellow-light period.

4.3. Crossing Reliability Analysis of Electric Bike Riders. Crossing reliability of electric bike riders calculated by three parametric distributions are shown in Figure 2. All of the probability curves present a monotonically decreasing trend, which means the crossing reliability decreases with the increasing waiting time. In addition, the curves of the Weibull and log-logistic distributions have a rapidly decreasing trend in the early part of the waiting time. It is consistent with the fact that about 28.1% of electric bike riders are at high risk of violation and low waiting time to cross against the red light in our sample. But, the exponential distribution has a low adaptability to describe the waiting time of electric bike riders.

Figure 3 gives the risk rate curve calculated by three parametric models. The constant risk in the exponential model reflects the fact that the risk of violating behavior would not change with the elapse of waiting time. Therefore, the exponential model is not appropriate for describing the crossing behavior of electric bike riders. On the other hand,

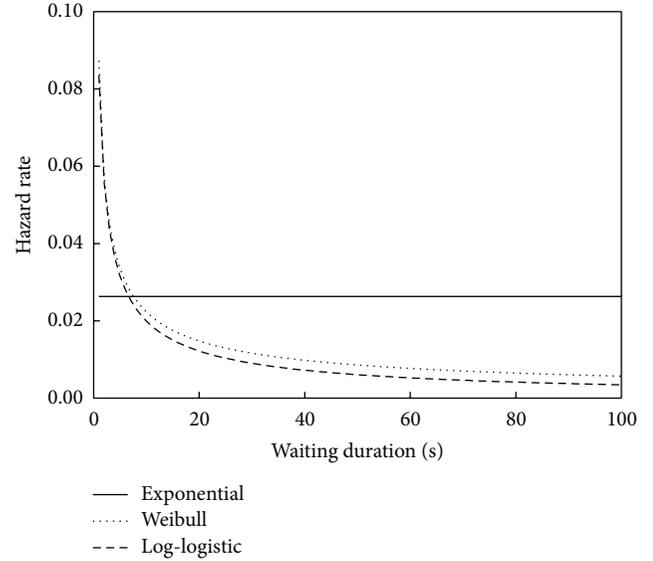


FIGURE 3: Hazard rate with waiting duration in different parametric models.

the curves of the Weibull and log-logistic distributions show a clear downward trend over waiting time, reflecting a negative dependence. It means that the likelihood of terminating the waiting time decreases with elapsed waiting time. It is consistent with the fact that a part of riders are generally nonrisk takers who can obey the traffic rules after waiting a long time.

4.4. Reliability Function of Crossing Behavior. Once electric bike riders violate traffic light, such a crossing behavior is defined as the occurrence of a risk event. Crossing reliability of electric bike riders is the probability that an electric bike rider obey the traffic light after a specified waiting time at the intersection. The Weibull distribution model gives the best description of riders' crossing behavior. Taken the Weibull distribution model, crossing reliability of electric bike riders using our sample data can be written as

$$\begin{aligned}
 R(t) &= \exp(-(\lambda t)^\gamma) = \exp\{-[t \exp(-\beta X)]^\gamma\} \\
 &= \exp\{-[t \exp(-1.009 - 0.064 * WN + 0.089 * CN \\
 &\quad + 0.331 * TC - 0.146 * MV \\
 &\quad + 0.95 * LT + 0.880 * YL)]^{0.408}\}. \tag{14}
 \end{aligned}$$

The proposed reliability model can capture the covariate effects of crossing behavior of electric bike riders at the intersection. Before the applications, however, it is noted that the model should be estimated using the specified field data. Additionally, the explanatory variables should be chosen flexibly according to the specified traffic circumstance.

5. Conclusions

The paper uses the concept of crossing reliability and the methodology of survival analysis to examine violating behavior of electric bike riders at the intersection. Three parametric duration models by the exponential, Weibull, and log-logistic distributions are proposed to analyze the covariate effects on crossing reliability of electric bike riders in red-light phase. The methodology uses a duration model structure that is able to predict crossing reliability at any given traffic conditions. More importantly, three sets of variables are selected as the covariates: individual characteristics, movement information, and situation factors. It is proved that the Weibull distribution is the most appropriate for describing crossing behavior of electric bike riders.

The paper provides several important insights into the determinants of crossing reliability of electric bike riders. Firstly, from the methodological standpoint, this study has provided the empirical evidence that hazard-based duration approach is appropriate for crossing reliability analysis. A statistical test of covariate effects on crossing reliability clearly indicates the goodness-of-fit of the used parametric hazard models. Secondly, the results indicate that movement information and situation factors are significantly related to riders' crossing reliability. Finally, various factors can modify crossing reliability of electric bike riders. The more riders waiting at the stop line, the less other riders crossing against the red-light, and the more motorized vehicle volume, electric bike riders would endure longer waiting times and have higher crossing reliability. While, electric bike riders have lower crossing reliability during the left-turn period of motorized vehicles and the yellow-light period.

In terms of the future work, research with more datasets is required. Also, more parameters under different situations should be taken into account. In addition, it is necessary to study the comparison between electric bikes and common bikes. It is expected that the findings will help traffic engineers, urban planners, and policy makers to understand electric bike riders' behavior at signalized intersections.

Acknowledgments

This work was supported by the National Basic Research Program of China (Grant no. 2012CB725401), the National Natural Science Foundation of China (Grant nos. 70901005, 71131001), and Fundamental Research Funds for the Central Universities (Grant no. 2013JBM043).

References

- [1] <http://zhaokuang541020.blog.163.com/blog/static/1279301062-01021931336293/>.
- [2] <http://www.nea.gov.cn/2012-06/20/c.131664583.htm>.
- [3] Q. Zhou, H. Lu, and W. Xu, "Comparison and analysis of urban resident trip characteristics," *Journal of Central South Highway Engineering*, vol. 32, no. 2, pp. 145–149, 2007 (Chinese).
- [4] <http://www.cebike.com/news/html/201204/2012042908414362.htm>.
- [5] CRTASR, *China Road Traffic Accidents Statistics Report*, Traffic Administration Bureau of China State Security Ministry, Beijing, China, 2010, (Chinese).
- [6] CRTASR, *China Road Traffic Accidents Statistics Report*, Traffic Administration Bureau of China State Security Ministry, Beijing, China, 2004, (Chinese).
- [7] C. X. Wu, L. Yao, and K. Zhang, "The red-light running behavior of electric bike riders and cyclists at urban intersections in China: an observational study," *Accident Analysis and Prevention*, vol. 49, no. 11, pp. 186–192, 2012.
- [8] K. L. H. Martinez and B. E. Porter, "Characterizing red light runners following implementation of a photo enforcement program," *Accident Analysis and Prevention*, vol. 38, no. 5, pp. 862–870, 2006.
- [9] B. E. Porter, K. L. Johnson, and J. F. Bland, "Turning off the cameras: red light running characteristics and rates after photo enforcement legislation expired," *Accident Analysis and Prevention*, vol. 50, pp. 1104–1111, 2013.
- [10] M. M. Ishaque and R. B. Noland, "Behavioural issues in pedestrian speed choice and street crossing behaviour: a review," *Transport Reviews*, vol. 28, no. 1, pp. 61–85, 2008.
- [11] E. Papadimitriou, G. Yannis, and J. Golias, "A critical assessment of pedestrian behaviour models," *Transportation Research F*, vol. 12, no. 3, pp. 242–255, 2009.
- [12] R. G. Zhou, W. J. Horrey, and R. F. Yu, "The effect of conformity tendency on pedestrians' road-crossing intentions in China: an application of the theory of planned behavior," *Accident Analysis and Prevention*, vol. 41, no. 3, pp. 491–497, 2009.
- [13] G. Tiwari, S. Bangdiwala, A. Saraswat, and S. Gaurav, "Survival analysis: pedestrian risk exposure at signalized intersections," *Transportation Research F*, vol. 10, no. 2, pp. 77–89, 2007.
- [14] M. Johnson, J. Charlton, and J. Oxley, "Cyclists and red light-a study of behaviour of commuter cyclists in Melbourne," in *Proceedings of the Australasian Road Safety Research, Policing and Education Conference*, Adelaide, Australia, November 2008.
- [15] M. Johnson, S. Newstead, J. Charlton, and J. Oxley, "Riding through red lights: the rate, characteristics and risk factors of non-compliant urban commuter cyclists," *Accident Analysis and Prevention*, vol. 43, no. 1, pp. 323–328, 2011.
- [16] M. Johnson, J. Charlton, J. Oxley, and S. Newstead, "Why do cyclists infringe at red lights? An investigation of Australian cyclists' reasons for red light infringement," *Accident Analysis and Prevention*, vol. 50, no. 1, pp. 840–847, 2013.
- [17] C. R. Bhat, "Duration modeling," in *Handbook of Transport Modelling*, D. A. Hensher, D. A. Hensher, and K. J. Button, Eds., pp. 91–111, Elsevier Science, Oxford, UK, 2000.
- [18] K. M. N. Habib, "Modeling commuting mode choice jointly with work start time and work duration," *Transportation Research A*, vol. 46, no. 1, pp. 33–47, 2012.
- [19] Y. Chung, "Development of an accident duration prediction model on the Korean freeway systems," *Accident Analysis and Prevention*, vol. 42, no. 1, pp. 282–289, 2010.
- [20] X. B. Yang, Z. Y. Gao, H. W. Guo, and M. Huan, "Survival analysis of car travel time near a bus stop in developing countries," *Science in China E*, vol. 55, no. 8, pp. 2355–2361, 2012.
- [21] X. B. Yang, M. Huan, H. W. Guo, and L. Gao, "Car travel time estimation near a bus stop with non-motorized vehicles," *International Journal of Computational Intelligence Systems*, vol. 6, no. 4, pp. 1350–1357, 2011.
- [22] H. W. Guo, W. H. Wang, W. W. Guo, and F. C. Zhao, "Modeling lane-keeping behavior of bicyclists using survival analysis

- approach,” *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 197518, 6 pages, 2013.
- [23] H. W. Guo, W. H. Wang, W. W. Guo, X. B. Jiang, and H. Bubb, “Reliability analysis of pedestrian safety crossing in urban traffic environment,” *Safety Science*, vol. 50, no. 4, pp. 968–973, 2012.
- [24] X. B. Yang, M. Huan, B. F. Si, L. Gao, and H. W. Guo, “Crossing at a red light: behavior of cyclists at urban Intersections,” *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 490810, 12 pages, 2012.
- [25] E. T. Lee and J. W. Wang, *Statistical Methods for Survival Data Analysis*, John Wiley & Sons, New York, NY, USA, 2003.
- [26] W. H. Wang, H. W. Guo, Z. Y. Gao, and H. Bubb, “Individual differences of pedestrian behaviour in midblock crosswalk and intersection,” *International Journal of Crashworthiness*, vol. 16, no. 1, pp. 1–9, 2011.
- [27] L. Yao, L. S. Sun, Z. Y. Zhang, S. W. Wang, and J. Rong, “Research on the behavior characteristics of pedestrian crowd weaving flow in transport terminal,” *Mathematical Problems in Engineering*, vol. 2012, Article ID 264295, 9 pages, 2012.

Research Article

Research on Quantitative Models of Electric Vehicle Charging Stations Based on Principle of Energy Equivalence

Zhenpo Wang,¹ Peng Liu,² Jia Cui,¹ Yue Xi,¹ and Lei Zhang¹

¹ School of Mechanical and Vehicular Engineering, Beijing Institute of Technology, Beijing 10081, China

² Key Laboratory of Power Electronics and Electric Drive, Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing 100190, China

Correspondence should be addressed to Zhenpo Wang; wangzhenpo@bit.edu.cn

Received 8 June 2013; Revised 25 July 2013; Accepted 28 July 2013

Academic Editor: Wuhong Wang

Copyright © 2013 Zhenpo Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to adapt the matching and planning requirements of charging station in the electric vehicle (EV) marketization application, with related layout theories of the gas stations, a location model of charging stations is established based on electricity consumption along the roads among cities. And a quantitative model of charging stations is presented based on the conversion of oil sales in a certain area. Both are combining the principle based on energy consuming equivalence substitution in process of replacing traditional vehicles with EVs. Defined data are adopted in the example analysis of two numerical case models and analyze the influence on charging station layout and quantity from the factors like the proportion of vehicle types and the EV energy consumption at the same time. The results show that the quantitative model of charging stations is reasonable and feasible. The number of EVs and the energy consumption of EVs bring more significant impact on the number of charging stations than that of vehicle type proportion, which provides a basis for decision making for charging stations construction layout in reality.

1. Introduction

In the transportation sector, the widespread use of the combustion engine vehicles has become an important factor of energy consumption and environmental pollution [1, 2], which leads the EV to get more attention owing to its distinguished features of genuine zero emission and high efficiency [3, 4]. In China, the EV has become an essential part of the new energy strategy and smart grid field [5, 6]. The State Council of China has defined the EV as a new strategic industry [7]. Thus, the electric vehicle will become the focus of Chinese automotive industry and energy industry in the future.

EV charging station is an important supportive infrastructure to boost the market acceptance and popularity of EVs [8, 9], because the accessibility and convenience of charging stations would impose significant influence on consumers' decisions due to the inherent demand of periodic charging [10, 11]. Currently, relative research has been a hot topic and received increasing interest and dedicated effort from researchers and practitioners. Hatton et al. [12] pointed out that the construction of charging facilities' network was of

great significance for the EV market expansion and described a variety of charging facility construction modes as well as their necessary requirements. Wang et al. [13] proposed a multiobjective planning model, which takes factors such as characters of charging station, characters of charging consumers, distribution of the charging demands, power grid, and municipal planning into account. Revelle and Eiselt [14] established a classical location model and adjusted the model by using different constraints. Wang et al. [15] researched on matching theory of EV charging stations. Johannes et al. [16] established a dynamic spatial model of the development of a charging infrastructure for EVs to allocate the need for charging stations in space. He et al. [17] established EV charging station's minimum comprehensive cost model. Feng et al. [18] presented a novel method for EV charging station planning based on the weighted Voronoi diagram. Ning et al. [19] analyzed the economy of battery-swapping charging mode, which is one of the factors on the charging stations construction layout. However, in the field of the location of EV charging stations, there was not enough analysis on quantitative model, and the layout

planning system of charging station has not been formed yet.

Since the charging time of electric vehicles is longer than conventional fuel vehicles, taking up a longer time for parking, the planning and layout of the charging stations are subject to geographical location conditions, including distribution capacity, available space, time of waiting for charging, peripheral facilities, and many other factors, especially the cities with larger population density and scarce land resources. If geographical conditions permit, theoretical research of charging station construction and planning and its scale and influencing factor analysis is a relative effective planning measure at the early stage of electric vehicles. The network planning of charging stations in Beijing in which the author participated is also done based on this idea.

In this paper, a quantitative model and a location model along intercity road are established by introducing some parameters such as regional coefficients of variation and attraction coefficient of charging stations, which are all based on the principle of energy equivalence. Aiming at designing the layout of urban charging stations, a quantitative model has been established based on the oil sales transaction. The correctness and effectiveness of models are further proved by case studies. The influence and sensitivity of the factors including the number of EVs, the proportion of vehicle type, and the electricity consumption of EVs are analyzed.

2. Quantitative Model of Electric Vehicle Charging Stations

2.1. Quantitative Model of Charging Stations along Intercity Roads. Quantitative model of charging stations along intercity roads is proposed based on the consumption of typical linear roads. That is, the road system is assumed to be composed of a main line and a bunch of forks which divided the region into several sections [20]. In order to simplify the model, forks of the road should not be too many, which indicates that the theory is not applicable to the layout planning of charging stations in urban area. However, it has a good applicability to typical intercity roads such as highways, national highways, and provincial highways.

2.1.1. Parameter Definitions

(1) *Length of Road (km, L).* The length of road is selected to establish the model.

(2) *Traffic Flow of Road (Vehicles/Day, V).* It denotes the number of vehicles that pass the road L in a certain time, coming from the result of statistics. The traffic flow of vehicle type i ($1 \leq i \leq N$) is expressed as V_i .

(3) *Power Consumption per 100 km of EVs (kWh/km, Q).* The statistics of average power consumption per 100 km of EVs is used to convert into average power consumption per kilometer which can be indicated by C_y .

(4) *Regional Coefficient of Variation (C_d).* The conditions of different roads used by the model are various. For instance,

the average power consumption per 100 km in mountainous regions and plains is different. Therefore, a regional coefficient of variation (C_d) is introduced to regulate the average power consumption per 100 km in the model. This parameter comes from statistics.

(5) *Distance between Charging Stations (km, L).* The distance between two adjacent charging stations is calculated to ensure the charging stations to make a certain profit. In order to make sure that at least charging stations are not at a loss, the minimum distance is defined as l_{\min} .

(6) *Zero Profit Daily Electricity Sales of a Charging Station (kWh).* The zero profit daily electricity sales of a charging station (W_0) is figured out based on the cost of the charging station (hardware costs, service costs, daily expenses, and other factors) and the electricity sales profit.

(7) *Attraction Coefficient of a Charging Station (C_y).* In practice, customers tend to choose better-equipped charging stations offering higher-quality service at a lower cost when their EVs need to be charged. If attraction coefficient of a charging station was not introduced, the calculated distance between two charging stations would be shorter, which could lead to low profits or losses, resulting in a waste of social resources. This parameter comes from statistics. It is the ratio of actual electricity sales of charging stations and the theoretical power consumption of EVs in a certain road.

(8) *Expected Daily Electricity Sales of a Charging Station (kWh, W_q).* Expected daily electricity of a charging station is the electricity that the charging station expects to sell every day. This parameter comes from statistics.

(9) *There Are M Forks of the Road.* These M forks divided the road into $M - 1$ short sections; each short section has no forks of the road, so that the traffic flow of this section is uniform. $V_{i,j}$ can be defined as the traffic volume of vehicle type (i) of the road (L_j). This parameter comes from statistics.

2.1.2. Assumptions

(1) *Zero Profit Electricity Sales in the Same Level of Charging Stations Are Equal.* The daily zero profit electricity in a charging station will be influenced by plenty of factors including the scale of the charging station, the number of workers, service levels, and equipment costs. However, in order to simplify the calculation, it is assumed that the zero profit electricity sales in the same level of charging stations are equal in the modeling process.

(2) EVs are recharged in accordance with the principle of proximity and needs, which means that the charging pattern accords with the condition that the EV needs to be recharged when the remaining energy reaches the alert value and the energy can be filled up each time, ignoring other factors.

Under this assumption, only the position factor is taken into consideration. Actually, when EVs need electricity supply, drivers will certainly choose the cheaper and larger

ones with higher quality services. This may contribute to drivers' violation towards the principle of proximity and needs when they want to recharge their EVs. This will inevitably lead to a shorter distance between the charging stations which are calculated based on the model. In order to make the model more practical, C_y (attraction coefficient of a charging station) is introduced when modeling, so that the electricity sales in the section L is equivalent to total energy consumption $\times C_y$.

(3) *The Charging Electricity Level of EVs on the Road Is Equal to the Energy Consumption on the Road \times Attraction Coefficient of a Charging Station.* Right now, road L is selected as an object of study. It will be verified from the simple case at the beginning that the selected road should not be too long, so that the EV will be recharged only once at most in a certain road.

For any chosen EV, w (kWh) is the control remaining electricity in the battery and W (kWh) is the fully charged battery level. When the EV runs on road L , battery storage of electricity ($W(x)$) is between w and W , $w < W(x) < W$. The power consumption of the EV in this section is expressed as $W(L)$, so the remaining battery level of this EV is $W(x) - W(L)$ when it leaves this section. Then, if $W(x) - W(L) < w$, the EV must be recharged on the road, and the complementary power is expressed as $W - w$; when leaving the road the remaining battery level of the EV is

$$W_s = W(x) - W(L) + W - w. \quad (1)$$

When $W(x) - W(L) > w$, the EV will not need to be charged at the charging station along this road. If $W(x) - W(L) < w$, it means $w < W(x) < w + W(L)$, and the EV must be charged in the charging station in this section. The complementary energy is $W - w$.

Assuming that reduce rate of the battery energy is proportional to the driving distance for EVs and because the selection of the section is arbitrary, the $W(x)$ will be distributed uniformly between $w - W$. Under this assumption, the probability of EV charged on the road is

$$P_{W(x)|w < W(x) < w + W(L)} = \frac{W(L)}{W - w}. \quad (2)$$

Thus, the charged electricity of the EV on the road is

$$(W - w) \times \frac{W(L)}{W - w} \times C_y = W(L) \times C_y. \quad (3)$$

The charging electricity of the EV on the road is equal to the consumption of energy \times attraction coefficient of charging station.

Then, charge electricity of all EVs passed through the road equals $\sum W(L) \times C_y$. Additionally, if the EVs are recharged more than once because of the overlong selected road, the road could be divided into many sections, the EV would be recharged only once in each section. It is proved right in every short section, so the assumption is still valid.

2.1.3. *Modeling.* (1) W_{Q_i} is the electricity consumption of the EV i in the road L_j . Taken the attraction coefficient of

a charging station and regional coefficient of variation and other factors into consideration, the result is shown in the following:

$$W_{Q_i} = Q_i \times C_d \times L_j \times V_{i,j} \times C_y. \quad (4)$$

(2) The electricity consumption (W_{Q_j}) of all the EVs in the road L_j is expressed as the summation of the electricity consumption of all types of EVs in this section, as shown in the following:

$$\begin{aligned} W_{Q_j} &= \sum_{i=1}^N W_{Q_i} \\ &= \sum_{i=1}^N (Q_i \times C_d \times L_j \times V_{i,j} \times C_y) \\ &= C_d \times L_j \times C_y \times \sum_{i=1}^N (Q_i \times V_{i,j}). \end{aligned} \quad (5)$$

From the above, the total electricity consumption of all EVs in the section L is defined as

$$\begin{aligned} W_Q &= \sum_{j=1}^M W_{Q_j} \\ &= \sum_{j=1}^M \left(C_d \times L_j \times C_y \times \sum_{i=1}^N (Q_i \times V_{i,j}) \right) \\ &= C_d \times C_y \times \sum_{j=1}^M \left(L_j \times \sum_{i=1}^N (Q_i \times V_{i,j}) \right). \end{aligned} \quad (6)$$

(3) According to assumption (3) in Section 2.1.2, electricity sales on the road L equal the electricity consumption \times the attraction coefficient of a charging station. Because it is assumed that, in Section 2.1.2 (1), the same level of charging stations is equal to the zero profit electricity sales, when the zero profit electricity sales of a charging station is W_0 , the maximum number of permitted charging stations is the ratio of the total electricity sales and zero profit electricity sales, as shown in the following:

$$\begin{aligned} z_{\max} &= \frac{W_Q}{W_0} \\ &= \frac{C_d \times C_y \times \sum_{j=1}^M (L_j \times \sum_{i=1}^N (Q_i \times V_{i,j}))}{W_0}. \end{aligned} \quad (7)$$

(4) Average distance between the charging stations is the ratio of the length and the number of charging stations in a section, as shown the following:

$$\begin{aligned} l_{\min} &= \frac{L}{z_{\max}} \\ &= \frac{L \times W_0}{C_d \times C_y \times \sum_{j=1}^M (L_j \times \sum_{i=1}^N (Q_i \times V_{i,j}))}. \end{aligned} \quad (8)$$

(5) Assuming that a charging station is at any point along the road L_j , W_q is defined as the expected sales electricity of charging stations, and l_j is expressed as the control distance from the upstream charging station, based on assumption (3) and formula (7), W_{Q_j} can be defined as

$$W_{Q_j} = C_d \times l_j \times C_y \times \sum_{i=1}^N (Q_i \times V_{i,j}), \quad (9)$$

$$W_{Q_j} = W_q.$$

Scilicet,

$$C_d \times l_j \times C_y \times \sum_{i=1}^N (Q_i \times V_{i,j}) = W_q, \quad (10)$$

$$l_j = \frac{W_q}{C_d \times C_y \times \sum_{i=1}^N (Q_i \times V_{i,j})}. \quad (11)$$

2.2. Quantitative Model of Charging Stations in a Certain Area. According to the number of gas stations and average data of daily oil sales in a certain area, the data of average daily oil sales of all gas stations are translated into average daily electricity sales of charging stations through introducing conversion coefficient. The average daily electricity sales will be the expected average daily electricity sales when the EVs are popularized in a certain area in the future. Based on the electricity sales level of charging stations, the quantitative model of charging stations in a certain area can be established.

2.2.1. Parameter Definitions

(1) *Daily Oil Sales of a Gas Station (L, O_i)*. This parameter is defined as the daily oil sales in a gas station, which comes from statistics. It is calculated from the data of yearly oil sales divided by the number of days of a year. By this way, the fluctuation of the data due to the changes of seasons, weather, traffic flow, and other random factors can be reduced.

(2) *Daily Electricity Sales of a Charging Station (kWh, E_i)*. This parameter is defined as the daily electricity sales of a charging station in the selected area, which can be figured out.

(3) *The Proportional Coefficient of the Corresponding Oil Sales and Electricity Sales ($kWh/L, I_j$)*. The proportional coefficient is the total oil sales and corresponding electricity sales which are converted out by calculation.

(4) *The Number of Gas Stations Is M ($1 \leq i \leq M$)*. This parameter is the number of gas stations in a certain area which will be calculated.

(5) *Average Oil Consumption per 100 km of Traditional Vehicles ($L/100 km, G_j$)*. The parameter is defined as an average oil consumption per 100 km of the main vehicles which fuel up in the gas station, ignoring the vehicles that seldom fuel up

in the gas station in a certain area. The numbers come from statistics.

(6) *Average Electricity Consumption per 100 km of EVs ($kWh/100 km, Q_j$)*. The parameter is defined as the electricity consumption per 100 km of EVs which is corresponded to the main types of traditional vehicles, which comes from statistics.

(7) *The Proportional Coefficient of Fuel Consumption and Electricity Consumption of Corresponding Types of Vehicles ($kWh/L, I_j$)*. The parameter is a proportional coefficient of the fuel consumption per 100 km of traditional vehicle and the electricity consumption of corresponding EVs. This parameter will be figured out.

(8) *The Number of Types of Vehicles Entering the Gas Station ($N, 1 \leq j \leq N$)*. The parameter is used for the classification of the main vehicles which are fuelled up in a certain area, ignoring the vehicles which are seldom fuelled up in the gas stations in the certain area.

(9) *The Proportion of Vehicle Types (D_j)*. This parameter is the proportion of vehicles of various types, which is from statistics.

(10) *The Proportion of Users Charging in the Charging Station ($C_{cd}, 0 < C_{cd} \leq 1$)*. Since the users may charge at home or residential parking, this parameter needs to be introduced. It is the proportion of the energy getting from the charging station.

(11) *Zero Profit Daily Electricity Sales of a Charging Station (kWh, W_0)*. This parameter is defined the same as the one in Section 2.2.1 (6).

(12) *Expected Daily Electricity Sales of a Charging Station (kWh, W_q)*. This parameter is defined the same as the one in Section 2.2.1 (8).

2.2.2. Modeling. (1) The proportional coefficient I_j is the ratio of electricity consumption and fuel consumption. It equals the ratio of electricity consumption per 100 km of EVs and fuel consumption per 100 km of refueling vehicles. It is shown in the following:

$$I_j = \frac{Q_j}{G_j}. \quad (12)$$

(2) The proportional coefficient I_i is the ratio of oil sales of gas stations and electricity sales of charging stations. It equals the summation of the product of the proportion of different EV types and the corresponding proportional coefficient I_j . It is shown in the following:

$$I_i = \sum_{j=1}^N (D_j I_j), \quad \sum_{j=1}^N D_j = 1, \quad I_i = \frac{E_i / C_{cd}}{O_i}. \quad (13)$$

(3) The daily electricity sales of a charging station (E_i) is a integrated result, which takes factors such as proportion of

TABLE 1: Data of traffic flow in each section.

Section	Length (km)	Traffic flow (vehicle/d)	Traffic flow of EVs (vehicle/d)	Traffic flow of passenger EVs (vehicle/d)	Traffic flow of commercial EVs (vehicle/d)
AB	7.1	34614	3462	2770	692
BC	2.36	32300	3230	2584	646
CD	10.95	27661	2766	2213	553
DE	2.9	29797	2980	2384	596

the user charging in charging stations and the oil sales into account:

$$E_i = C_{cd}O_iI_i = C_{cd}O_i \sum_{j=1}^N (D_j I_j) = C_{cd}O_i \sum_{j=1}^N \left(D_j \frac{Q_j}{G_j} \right). \quad (14)$$

(4) From the above, the sum of the daily electricity sales of charging stations in a certain area can be acquired. And it is defined as

$$\begin{aligned} E &= \sum_{i=1}^M E_i \\ &= \sum_{i=1}^M \left(C_{cd}O_i \sum_{j=1}^N \left(D_j \frac{Q_j}{G_j} \right) \right) \\ &= C_{cd} \sum_{i=1}^M \left(O_i \sum_{j=1}^N \left(D_j \frac{Q_j}{G_j} \right) \right). \end{aligned} \quad (15)$$

Then the maximum permitted number of charging stations in the certain area z_{\max} is defined as

$$z_{\max} = \frac{E}{W_0} = \frac{C_{cd} \sum_{i=1}^M \left(O_i \sum_{j=1}^N \left(D_j \left(\frac{Q_j}{G_j} \right) \right) \right)}{W_0}. \quad (16)$$

The proper amount of charging stations that will be built in the certain area is defined as

$$z_{jy} = \frac{E}{W_q} = \frac{C_{cd} \sum_{i=1}^M \left(O_i \sum_{j=1}^N \left(D_j \left(\frac{Q_j}{G_j} \right) \right) \right)}{W_q}. \quad (17)$$

It can be seen from the modeling process that the model is based on the proportion of the oil consumption of traditional vehicles and the energy consumption of corresponding EVs, and then the conversion coefficient of the oil sales and the electricity sales can be worked out by statistical method and weighting mathematical method. Finally, the potential demand for electricity of EVs in a certain area can be calculated according to the conversion coefficient.

3. Analysis on the Model Influencing Factors

To verify the validity of the model and analyze its influencing factors, some examples are computed and analyzed by using typical road traffic flow and the number of regional gas stations as basic parameters, and the sensitivity of the parameters needs to be analyzed and quantified.

3.1. Example of the Quantitative Model of Charging Stations along Intercity Road

3.1.1. Settings of Basic Parameters. For example [21], the total length of one section of highway is 23.3 km, there are five entrances or exits (points A, B, C, D, and E) dividing it into four sections, and the length and traffic flow data of each section are shown in Table 1. Assuming that the proportion of EVs could be 10% of the total number, the proportion of passenger EVs and commercial EVs is 4 : 1, and the traffic flow data of EVs is shown in Table 1, too.

According to the statistics of electricity consumption per 100 km of a EV in [22, 23], the electricity consumption per km of a passenger EV is calculated as $Q_1 = 0.25$ kWh/km and the electricity consumption per kilometer of a commercial EV is calculated as $Q_2 = 1.22$ kWh/km.

The regional coefficient of variation is defined as $C_d = 0.85$, and the attraction coefficient of charging stations is defined as $C_y = 0.8$.

Referring to “Technical specifications of electricity supply and assurance for electric vehicle: electric vehicle charging station” [24], the data of zero profit electricity sales of the tertiary charging station is introduced, which can provide service to 100 EVs per day. It is assumed that each EV can be charged for 100 kWh each time, so the daily maximum sales of a charging station is 10000 kWh. Assuming that the zero profit daily sales is 60% of the maximum daily sales and the expectation daily sales is 90%, W_0 is 6000 kWh and W_q is 9000 kWh.

3.1.2. Calculations of the Number of Charging Stations along the Intercity Road. (1) The electricity consumption in each section and the total consumption of the passenger EVs and commercial EVs are calculated, respectively, according to (4) and (5), and the results are shown in Table 2.

The total electricity sales in section AE are calculated as $W_Q = 21473.07$ kWh, and then the maximum permitted number of charging stations is figured out as $z_{\max} = 21473.07/6000 = 3.58$; it could be rounded up to an integer ($z_{\max} = 4$). The average length between the charging stations can be figured out as $l_{\min} = 23.3/4 = 5.825$ km.

(2) In order to obtain a satisfactory benefit, traffic flow and electricity consumption expectation sales should be taken into account when making the layout plan. It is assumed that there is a charging station in section BC, and the control distance with the upstream charging station could be calculated as shown in Table 3.

TABLE 2: Data of electricity consumption in each section.

Section	Electricity consumption of passenger EVs (kWh)	Electricity consumption of commercial EVs (kWh)	Total electricity consumption (kWh)
AB	3343.39	4075.99	7419.38
BC	1036.70	1264.78	2301.48
CD	4119.50	5023.52	9143.02
DE	1175.31	1433.88	2609.19

TABLE 3: Layout of charging stations.

No.	1	2
Control distance (km)	9.38	9.63
Control coordinate (km)	9.38	19.01

TABLE 4: Data of charging stations in different proportions of vehicle type.

Proportion of vehicle type	z_{\max}	l_{\min}
9:1	2.80	8.32
4:1	3.58	6.51
7:3	4.36	5.34

According to this result, defined amount of charging stations in this section is 2. If the number was more than two, the control coordinate would exceed the length of the road, which deviates from actuality. Therefore, this scheme is the best choice. This case study has shown that the quantitative model of charging stations along the intercity road is reasonable and feasible, which can help making layout plan for charging stations in a typical road.

3.1.3. *Analysis of Influencing Factors on the Number of the Charging Stations along Intercity Road.* (1) Proportion of vehicle type influence on the number of charging stations along intercity road.

Supposing that other parameters are constant values, the numbers of charging stations are calculated respectively when the proportions of passenger vehicles and commercial vehicles are 9:1 and 7:3. Table 4 shows the consequence which also demonstrates the corresponding data in Section 3.1.2.

The data in Table 4 shows that the amount of charging stations increases with the reduction of the proportion of commercial vehicles. With the proportion of the commercial vehicles increasing by 10% each time, the maximum number of permitted charging stations (z_{\max}) would become 3.58 and 4.36, increased by 27.6% and 21.7%, respectively.

(2) The number of EVs influences the number of charging stations along intercity road.

It is assumed that other parameters are constant values. Data of 15% and 20% of the total traffic flow are used, respectively, to represent the number of EVs in this calculation. Table 5 shows the consequence.

Table 5 shows that the maximum permitted number of charging stations increases significantly and the average

TABLE 5: Data of charging stations under the condition of different numbers of EVs.

Proportion of EVs	z_{\max}	l_{\min}
10%	3.58	6.51
15%	5.36	4.35
20%	7.13	3.27

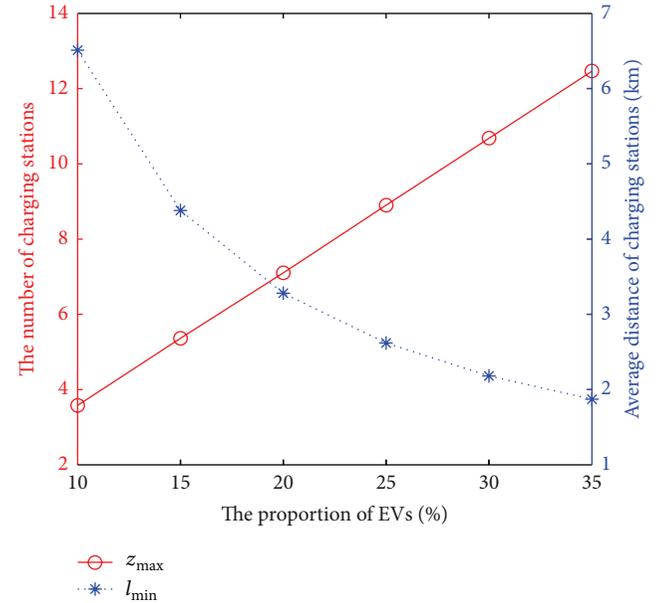


FIGURE 1: Data of charging stations in different number of EVs.

distance between the charging stations becomes shorter as the number of EVs increases. With the number of EVs increasing by 5%, the maximum number of charging stations would add 1.78 and 3.55 which is relative to the original data, respectively, which means a linear growth of 49% and 99%. Consequently, the number of EVs has remarkable influence on the number of charging stations in the case of a certain scale of charging stations. It means that the number of EVs is a sensitive factor on the charging stations planning.

With the development of EV market, the number of EVs will increase continuously. The conditions of different numbers of EVs (25%, 30%, and 35% of total number of vehicles) should be calculated and analyzed; Figure 1 shows the results.

It is clearly shown in Figure 1 that the number of charging stations increases linearly with the increase of the number of EVs, and the average distance between the charging stations becomes shorter with the increase of the number of EVs. Therefore, if the number of EVs is continually expanding, the layout of charging stations would be more and more intensive. So it is no more reasonable to meet the needs of intercity road by increasing the number of charging stations. This problem should be solved by taking some other measures such as expanding the scale of charging stations or increasing the serving efficiency.

(3) Electricity consumption of EVs influences the number of charging stations.

TABLE 6: Data of charging stations in different electricity consumptions.

Q_1	Q_2	z_{\max}	l_{\min}
0.25	1.22	3.58	6.51
0.225	1.098	3.22	7.24
0.2	0.976	2.86	8.15

With the continuous development and progress of battery technology and EV technology, an average of energy consumption per 100 km of EVs reduced by 10% or 20% energy saving ratio is possible. So the influence on the number of charging stations under the condition of different energy consumptions should be taken into consideration. It is assumed that other parameters are constant values, and the numbers of charging stations are calculated, respectively, when the electricity consumption of EVs is decreased by 10% and 20%. The result is shown in Table 6.

Table 6 indicates that the maximum permitted number of charging stations is reduced by 0.52 and 0.48 with the reduction of the electricity consumption of EVs. The analysis shows that the energy consumption of EVs decreasing by 10% may lead to the reduction of charging stations by 14.53% and 14.90%, respectively.

(4) Different grades of charging stations influences the number of charging stations along intercity road.

According to “Technical specifications of electricity supply and assurance for electric vehicle: electric vehicle charging station” [24], the charging stations are divided into four grades based on the distribution capacity and service capabilities. Different grades of charging stations have different daily electricity sales which will influence the layout of charging stations. Assuming that other parameters are constant values, the maximum permitted number of charging stations and the average distance between charging stations are calculated, respectively, under the condition of the second grade and the fourth grade of charging stations. The result is in Table 7, compared with the data under the condition of third grade of charging stations which has been shown in Section 3.1.2.

It can be seen that different grades of charging stations have different zero profit daily electricity sales. The number of the third grade of charging stations is 1.79 more than that of the second grade, and the number of the fourth grade of charging stations is 5.37 more than that of the third grade. At the same time it is shown that when the zero profit daily electricity sales decrease by 50% the number of the charging stations increases by 100%, and when the zero profit daily electricity sales decreases by 40%, the number of charging stations increases by 150%.

In conclusion, the proportion of EVs, number of EVs, electricity consumption of EVs, and different grades of charging stations influence the layout of charging stations. Among them, the number of charging stations is the most sensitive factor; if the number of EVs increased by 5% each time, the number of charging stations would increase by 49% and 99%, respectively. Therefore, the number of EVs should be taken into serious consideration when making the layout of charging stations.

TABLE 7: Data of charging stations in different grades of charging stations.

Grade of charging stations	W_0	z_{\max}	l_{\min}
II	12000	1.79	13.02
III	6000	3.58	6.51
V	2400	8.95	2.60

TABLE 8: Data of daily oil sales in gas stations.

No.	The number of gas stations	Daily oil sales per gas stations (L)
1	18	20000
2	90	15000
3	92	5000

TABLE 9: Data of other basic parameters.

Item	Parameter
Fuel consumption per 100 kilometers of passenger vehicle	9 L/100 km
Fuel consumption per 100 kilometers of commercial vehicle	45 L/100 km
Electricity consumption per 100 kilometers of passenger EV	25 kWh/100 km
Electricity consumption per 100 kilometers of commercial EV	122 kWh/100 km
The number of the vehicle type	2
Proportion of the passenger vehicle	0.8
Proportion of the commercial vehicle	0.2
The proportion of the user charging in charging stations	0.7
Zero profit daily electricity sales of a charging station	9000 kWh
Expected daily electricity sales of a charging station	12000 kWh

3.2. Example of the Layout of Charging Stations in a Certain Area

3.2.1. *Settings of Basic Parameters.* Suppose that there are 200 gas stations in an urban area [25]. The data of daily oil sales in gas stations are shown in Table 8, and the data of other basic parameters are shown in Table 9.

3.2.2. *Calculations of the Number of Charging Stations in a Certain Area.* According to (12), the scale factor of the oil consumption and the electricity consumption of passenger vehicles is calculated as $I_1 = 2.78$, and the scale factor of the oil consumption and the electricity consumption of commercial vehicles is figured out as $I_2 = 2.71$. The scale factor of the oil consumption and the electricity consumption in the corresponding gas stations and charging stations is figured out according to (13), and then all the data of daily oil sales in gas stations and the data of daily electricity sales in the charging stations can be worked out as Table 10 shows.

TABLE 10: Data of daily electricity sales in the charging stations.

No.	The number of gas STs	Daily oil sales per gas STs (L)	Corresponding daily electricity sales (kWh)	Corresponding total daily electricity sales (kWh)
1	18	20000	38780	698040
2	90	15000	29085	2617650
3	92	5000	9695	891940

The total daily electricity sales in this area is worked out as $E = 4207630$ kWh. The maximum number of charging stations is calculated as $z_{\max} = 4207630/9000 = 467.51$, and it could be rounded up to an integer ($z_{\max} = 468$). The defined amount of charging stations is calculated as $z_{jy} = 4207630/12000 = 350.64$, and it could be rounded up to an integer ($z_{jy} = 351$).

The result indicates that an ideal program of building urban charging stations can be provided according to the quantitative model of charging stations in a certain area.

3.2.3. Analysis of Influencing Factors on the Number of the Charging Stations in a Certain Area. (1) Vehicle type proportion influences the number of charging stations in a certain area.

Assuming that other parameters are constant values, the numbers of charging stations are calculated, respectively, when the proportions of passenger vehicles and commercial vehicles are 9:1 and 7:3. Table 11 shows the consequence which also demonstrates the corresponding data in Section 3.2.2.

Table 11 shows that when the proportion of commercial vehicles increases by 10%, the maximum number of charging stations decreases by 0 and 2, respectively, the same as the defined amount of charging stations. It is obvious that the proportion of vehicle type has little influence the number of charging stations.

(2) Electricity consumption of EVs influences the number of charging stations in a certain area. According to the schemes described in Section 3.1.3 (3), the numbers of charging stations are calculated, respectively, when the electricity consumption of EVs decreased by 10% and 20%. The result is shown in Table 12.

Table 12 indicates that the electricity consumption and the number of charging stations are positively correlated. With the electricity consumption decreasing by 10% each time, the maximum number of charging stations decreases by 47 (10% in proportion), and defined amount of charging stations deduced by 31 (11.2% in proportion).

(3) Different grades of charging stations influence the number of charging stations in a certain area.

Assuming that other parameters are constant values, the number of charging stations is calculated in different grades of charging stations. The data is shown in Table 13. Taking different construction costs between the same grades of urban charging stations and intercity charging stations and other factors into consideration, the data of zero profit daily electricity sales in Table 13 and that in Section 3.1.3 (4) is different.

TABLE 11: Data of charging stations in different proportions of vehicle type.

Proportion of vehicle type	Maximum permitted number of charging stations	Defined amount of charging stations
9:1	468	351
4:1	468	351
7:3	466	349

TABLE 12: Data of charging stations in different electricity consumptions.

The decreased proportion of electricity consumption	Maximum permitted number of charging stations	Defined amount of charging stations
0	468	351
10%	421	316
20%	374	285

TABLE 13: Data of charging stations in different grades of charging stations.

Grade of charging station	W_0	Maximum permitted number of charging stations	Defined amount of charging stations
II	18000	234	176
III	9000	468	351
V	3600	1168	878

It can be seen that when the zero profit daily electricity sales decreases by 50% both the maximum permitted number of charging stations and defined amount of charging stations increase by 100% and when the zero profit daily electricity sales decrease by 40% both the maximum permitted number of charging stations and defined amount of charging stations increase by 150%. This law is corresponded with the one in Section 3.1.3 (4). So the influence of the number of charging stations is significant due to the conditions of different grades of charging stations.

This part analyzes the influence of proportion of vehicle types, electricity consumption of EVs, and the grades of charging stations on the number of charging stations in a certain area. Among them, the change of proportion of vehicle types has little effect on the number of charging stations. If the electricity consumption is decreased by 10%, the number of charging stations would reduce by 10% or so. The sensitivity factor is the grades of charging stations. If the zero profit daily electricity sales decreased by 50%, the

number of charging stations could reduce by 100%. When future plans for the construction of charging stations are actually made, the grades of charging stations and the serving capabilities should be taken into account, and at the same time, the impact of energy consumption of EVs should not be ignored.

4. Conclusions

(1) With energy equivalent substitution as the basic principle, functions and energy supply features of the charging stations are comprehensively considered to conduct influencing factor analysis of quantity model of the linear roads and fixed area charging stations, and results show that the model is applicable to the optimization calculation of the charging stations of the electric vehicles.

(2) The analysis of model proportion, number of electric vehicles, electric vehicles power consumption, and other influencing factors of the number of charging stations show that the proportion change of the models of vehicles will not become the main influencing factor of the number of the charging stations in a fixed area. With the number of EVs increasing by 5%, the maximum number of charging stations would have a linear growth of about 50%. The energy consumption of EVs decrease by 10% may lead to the reduction of charging stations by at least 14%. And the zero profit daily electricity saleing are an important factor for the number of the charging stations. With the electricity consumption decreasing by 10% each time, the maximum number of charging stations decreases 10% in proportion.

(3) At present, the future development should be taken into account when incubating a charging stations plan. As different scales of charging stations have different serving capabilities, some measures should be taken to enhance their service capabilities, such as decreasing the service time per vehicle are improving the technology and the quality of services. The growing number of EVs and the expanded scale of charging stations should also be taken into consideration.

Acknowledgments

The project was supported by the National Natural Science Foundation of China (61004092), National High Technology Research and Development Program of China (2011AA11A251), and Research Foundation of National Engineering Laboratory for Electric Vehicles (2012-NELEV-03).

References

- [1] Q. Guo, J. Han, M. Yoon, and G. Jang, "A study of economic dispatch with emission constraint in smart grid including wind turbines and electric vehicles," in *Proceedings of the IEEE Vehicle Power and Propulsion Conference*, pp. 1002–1005, 2012.
- [2] M. Inoue, "Environment friendly new energy supply system," *IEEE Transactions on Electrical and Electronic Engineering*, vol. 2, no. 1, pp. 39–43, 2007.
- [3] M. Etezadi-Amoli, K. Choma, and J. Stefani, "Rapid-charge electric-vehicle stations," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1883–1887, 2010.
- [4] M. Yilmaz and P. T. Krein, "Review of battery charger topologies, charging power levels, and infrastructure for plug-in electric and hybrid vehicles," *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pp. 2151–2168, 2013.
- [5] Z. Liu, F. Wen, and G. Ledwich, "Optimal planning of electric-vehicle charging stations in distribution systems," *IEEE Transactions on Power Delivery*, vol. 28, no. 1, pp. 102–110, 2013.
- [6] M. Simonov, "Mastering cooperation: electric vehicle and smart grid," in *Proceedings of the 11th International Conference on ITS Telecommunications (ITST '11)*, pp. 480–485, August 2011.
- [7] General Office of the State Council of China, *Decision of the State Council on Accelerating the Development of Strategic Emerging Industries*, General Office of the State Council of China, Beijing, China, 2010.
- [8] A. M. Sharaf and B. Khaki, "Novel switched capacitor-filter compensator for smart grid-electric vehicle charging scheme," in *Proceedings of the IEEE International Conference on Smart Grid Engineering (SGE '12)*, pp. 1–5, 2012.
- [9] S. Shahid, K. R. Narumanchi, and D. Gurkan, "Plug-in electric vehicle battery sensor interface in smart grid network for electricity billing," in *Proceedings of the IEEE Energy Conversion Congress and Exposition (ECCE '12)*, pp. 3403–3410, February 2012.
- [10] M. Singh, P. Kumar, and I. Kar, "A model of electric vehicle charging station compatibles with vehicle to grid scenario," in *Proceedings of the IEEE International Electric Vehicle Conference (IEVC '12)*, pp. 1–7, 2012.
- [11] T. W. Ching, "Design of electric vehicle charging station in Macau," *Journal of Asian Electric Vehicles*, vol. 9, no. 1, pp. 1453–1458, 2011.
- [12] C. E. Hatton, S. K. Beella, J. C. Brezet et al., "Charging station for urban settings: the design of a product platform for electric vehicle infrastructure in Dutch cities," in *Proceedings of the International Battery Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition*, pp. 514–526, 2009.
- [13] H. S. Wang, Q. Huang, C. H. Zhang, and A. Xia, "A novel approach for the layout of electric vehicle charging station," in *Proceedings of the International Conference on Apperceiving Computing and Intelligence Analysis (ICACIA '10)*, pp. 64–70, December 2010.
- [14] C. S. Revelle and H. A. Eiselt, "Location analysis: a synthesis and survey," *European Journal of Operational Research*, vol. 165, no. 1, pp. 1–19, 2005.
- [15] Z. P. Wang, X. H. Sun, Y. Wang, and L. Y. Liu, "Research on matching theory of EV charging station," *Information A*, vol. 15, no. 11, pp. 4359–4366, 2012.
- [16] W. Johannes, L. Susanne, and K. Alois, "Modelling the development of a regional charging infrastructure for electric vehicles in time and space," *European Journal of Transport and Infrastructure Research*, vol. 12, no. 4, pp. 391–416, 2012.
- [17] J. He, B. X. Zhou, C. Feng et al., "Electric vehicle charging station planning based on multiple-population hybrid genetic algorithm," in *Proceedings of the International Conference on Control Engineering and Communication Technology (ICCECT '12)*, pp. 403–406, 2012.
- [18] L. Feng, S. Y. Ge, and H. Liu, "Electric vehicle charging station planning based on weighted voronoi diagram," in *Proceedings of the Asia-Pacific Power And Energy Engineering Conference (APPEEC '12)*, 2012.
- [19] G. B. Ning, Z. J. Zhen, P. Wang, Y. Li, and H. X. Yin, "Economic analysis on value chain of taxi fleet with battery-swapping mode

- using multiobjective genetic algorithm,” *Mathematical Problems in Engineering*, vol. 2012, Article ID 175912, 15 pages, 2012.
- [20] Y. Wang and C. Lin, “Locating road-vehicle refueling stations,” *Transportation Research E*, vol. 45, no. 5, pp. 821–829, 2009.
- [21] L. Zhang, “Application of data of traffic volume in highway station,” *Transportation Information Industry*, pp. 85–86, 2009.
- [22] W. Chen, J. Guo, K. You, and Z. Zhen, “Energy-saving & emission reduction effects prospecting of ten cites & thousand units demonstration operation,” *Automobile & Parts*, pp. 34–35, 2010.
- [23] W. J. Sweeting, A. R. Hutchinson, and S. D. Savage, “Factors affecting electric vehicle energy consumption,” *International Journal of Sustainable Engineering*, vol. 4, no. 3, pp. 192–201, 2011.
- [24] Beijing Municipal Administration of Quality and Technology Supervision, *DB11/Z 728. 2010, Technical Specifications of Electricity Supply and Assurance for Electric Vehicle: Electric Vehicle Charging Station*, Beijing Municipal Administration of Quality and Technology Supervision, Beijing, China, 2010.
- [25] C. Wang, “Exploration on principles and strategies of rational distribution of gas stations of national oil companies,” *Petroleum & Petrochemical Today*, vol. 16, no. 11, pp. 30–34, 2008.