

Lecture Notes in Electrical Engineering 483

Limin Jia

Yong Qin

Jianguo Suo

Jianghua Feng

Lijun Diao

Min An

Editors

Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017

Transportation

Lecture Notes in Electrical Engineering

Volume 483

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Bijaya Ketan Panigrahi, New Delhi, India
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Shanben Chen, Shanghai, China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Haibin Duan, Beijing, China
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Limin Jia, Beijing, China
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Qilian Liang, Arlington, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Chandra Mukhopadhyay, Palmerston North, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Kyoto, Japan
Federica Pascucci, Roma, Italy
Yong Qin, Beijing, China
Gan Woon Seng, Singapore, Singapore
Germano Veiga, Porto, Portugal
Haitao Wu, Beijing, China
Junjie James Zhang, Charlotte, USA

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

Lecture Notes in Electrical Engineering (LNEE) is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering
- Engineering

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

For general information about this series, comments or suggestions, please use the contact address under "service for this series".

To submit a proposal or request further information, please contact the appropriate Springer Publishing Editors:

Asia:

China, *Jessie Guo, Assistant Editor* (jessie.guo@springer.com) (Engineering)

India, *Swati Meherishi, Senior Editor* (swati.meherishi@springer.com) (Engineering)

Japan, *Takeyuki Yonezawa, Editorial Director* (takeyuki.yonezawa@springer.com)
(Physical Sciences & Engineering)

South Korea, *Smith (Ahram) Chae, Associate Editor* (smith.chae@springer.com)
(Physical Sciences & Engineering)

Southeast Asia, *Ramesh Premnath, Editor* (ramesh.premnath@springer.com)
(Electrical Engineering)

South Asia, *Aninda Bose, Editor* (aninda.bose@springer.com) (Electrical Engineering)

Europe:

Leontina Di Cecco, Editor (Leontina.dicecco@springer.com)
(Applied Sciences and Engineering; Bio-Inspired Robotics, Medical Robotics, Bioengineering; Computational Methods & Models in Science, Medicine and Technology; Soft Computing; Philosophy of Modern Science and Technologies; Mechanical Engineering; Ocean and Naval Engineering; Water Management & Technology)

(christoph.baumann@springer.com)
(Heat and Mass Transfer, Signal Processing and Telecommunications, and Solid and Fluid Mechanics, and Engineering Materials)

North America:

Michael Luby, Editor (michael.luby@springer.com) (Mechanics; Materials)

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

Limin Jia · Yong Qin · Jianguo Suo
Jianghua Feng · Lijun Diao
Min An
Editors

Proceedings of the 3rd
International Conference
on Electrical and Information
Technologies for Rail
Transportation (EITRT) 2017

Transportation

Editors

Limin Jia
School of Traffic and Transportation
Beijing Jiaotong University
Beijing
China

Jianghua Feng
CRRC Zhuzhou Institute Co., Ltd.
Zhuzhou
China

Yong Qin
Beijing Jiaotong University
Beijing
China

Lijun Diao
Beijing Jiaotong University
Beijing
China

Jianguo Suo
CRRC Zhuzhou Locomotive Co., Ltd.
Zhuzhou
China

Min An
University of Salford
Manchester
UK

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-10-7988-7 ISBN 978-981-10-7989-4 (eBook)
<https://doi.org/10.1007/978-981-10-7989-4>

Library of Congress Control Number: 2017963532

© Springer Nature Singapore Pte Ltd. 2018

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

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

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

Printed on acid-free paper

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

Committees

Honorary Chairs

Youmei Liu, Academician, China
Rongjun Ding, Academician, China
Qingquan Qian, Academician, China
Zhongheng Shi, Academician, China
Xinning Zhang, Professor, China
Sone Satoru, Professor, Tokyo University, Japan

General Chairs

Prof. Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China
Prof. Ming Gong, CRRC Institute Co. Ltd., China

Program Committee Chairs

Prof. Zhigang Liu, Beijing Jiaotong University, China
Prof. Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China
Prof. Yaohua Li, Institute of Electrical Engineering, Chinese Academy of Sciences, China
Mr. Zhenchen Chang, CRRC Changchun Railway Vehicles Co. Ltd., China
Mr. Bangcheng Sun, CRRC Institute Co. Ltd., China
Mrs. Jianying Liang, National Engineering Laboratory for High-speed Train, CRRC Qingdao Sifang Co. Ltd., China

Mr. Tianyun Shi, The Center of National Railway Intelligent Transportation System Engineering and Technology, China Academy of Railway Sciences, China
Prof. Tefang Chen, National Engineering Laboratory for High-Speed Railway Construction, China
Prof. An Min, University of Salford, UK
Prof. Dr.-Ing Li Zhong, Fern Universität in Hagen, Germany
Prof. Dr.-Ing. Holger Hirsch, Universität Duisburg-Essen, Germany
Prof. Lothar H. Fickert, Vienna University of Technology, Austria
Prof. Suleiman M. Sharkh, University of Southampton, UK

Organizing Committee Chairs

Mr. Jianguo Suo, The State Key Laboratory of Heavy Duty AC Drive Electric Locomotive Systems Integration, CRRC Zhuzhou Locomotive Co. Ltd., China
Prof. Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China
Mr. Changqing Liu, National Engineering Laboratory for System Integration of High-speed Train, CRRC Changchun Railway Vehicles Co. Ltd., China
Mr. Xiaojun Deng, National Engineering Laboratory for High-speed Train, CRRC Qingdao Sifang Co. Ltd., China
Mr. Yunxin Fan, The State Key Laboratory of Heavy Duty AC Drive Electric Locomotive Systems Integration, CRRC Zhuzhou Locomotive Co. Ltd., China
Mr. Jianghua Feng, National Engineering Research Center of Converting Technology, China
Mr. Tianyun Shi, The Center of National Railway Intelligent Transportation System Engineering and Technology, China Academy of Railway Sciences, China
Prof. Weirong Chen, State Key Laboratory of Traction Power, Southwest Jiaotong University, China
Mr. Bangcheng Sun, CRRC Institute Co. Ltd., China
Prof. An Min, University of Salford, UK
Prof. Lijun Diao, Beijing Jiaotong University, China
Prof. Mingjian Zuo, University of Electronic Science and Technology of China, China
Prof. Xiaochun Ren, State Key Laboratory of Rail Transit Engineering Informatization, China
Prof. Zhiwu Yu, National Engineering Laboratory for High-Speed Railway Construction, China

Technical Program Committee Members

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

Prof. Li Yaohua, Institute of Electrical Engineering, Chinese Academy of Sciences, China

Prof. Shibin Gao, Southwest Jiaotong University, China

Prof. Jianyun Chai, Tsinghua University, China

Prof. Youtong Fang, Zhejiang University, China

Ms. Minghua Zhao, Changchun Bombardier Railway Vehicles Co. Ltd., China

Mr. Jun Li, CRRC Changchun Railway Vehicles Co. Ltd., China

Prof. Litian Wang, China Railway Electrification Survey and Design Institute Co. Ltd., China

Mr. Changjun Cai, Guangzhou Metro Corporation, China

Prof. Ming Gong, CRRC Institute Co. Ltd., China

Prof. Baoming Liu, China CNR Qingdao Sifang Locomotive & Rolling Stock Research Institute, China

Prof. Zhigang Liu, Beijing Jiaotong University, China

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

Prof. Ping Li, The Center of National Railway Intelligent Transportation System Engineering and Technology, China Academy of Railway Sciences, China

Prof. Xiaochun Ren, State Key Laboratory of Rail Transit Engineering Informatization, China

Prof. Jiuchun Jiang, Beijing Jiaotong University, China

Prof. Zhongping Yang, Beijing Jiaotong University, China

Prof. Buchheit Karlheinz, Experts of Siemens, Germany

Prof. Clave Roberts, University of Birmingham, UK

Prof. Ing. Kyandoghere Kyamakya, Universität Klagenfurt, Germany

Prof. Jianqiao Ye, Mechanical Engineering Department of Engineering, Lancaster University, UK

Prof. Mark Hooper, Faculty of Engineering and Computing, Coventry University, UK

Prof. Dr.-Ing Li Zhong, Fern Universität in Hagen, Germany

Prof. Dr.-Ing. Holger Hirsch, Universität Duisburg-Essen, Germany

Prof. Rui Chen, Loughborough University, UK

Prof. Satoru Sone, Tokyo University, Japan

Prof. Simon Wang, School of Aeronautical and Automotive Engineering, Loughborough University, UK

Prof. Tung-Chai Ling, University of Birmingham, UK

Prof. Wolfgang A. Halang, Fern Universität in Hagen, Germany

Prof. Su Zhongqing, Hong Kong Polytechnic University, Hong Kong

Dr. Tatsuhiko Fujihira, Fuji Electric, Japan

Dr. Paramjit Singh, Bombardier (Singapore) Pte Ltd, Singapore

Organizing Committee Members

Mr. Jiangnong Zhao, State Key Laboratory of Heavy Duty AC Drive Electric Locomotive Systems Integration, CRRC Zhuzhou Locomotive Co. Ltd., China

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

Prof. Min An, University of Salford, UK

Prof. Lijun Diao, Beijing Jiaotong University, China

Prof. Chunmei Xu, Beijing Jiaotong University, China

Dr. Xiaoqing Chen, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China

Dr. Zhengyu Xie, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, China

Mr. Zhe Xie, National Engineering Research Center of Converting Technology, China

Prof. Dr. Wolfgang A. Halang, Fern Universität in Hagen, Germany

Prof. Dr.-Ing Li Zhong, Fern Universität in Hagen, Germany

Contents

Transient Overvoltage Analysis of Traction Power Supply System with Neutral Sections in China High-Speed Railway Using a State-Space Model	1
Xiaoqin Lv, Xiaoru Wang and Jiwei Zheng	
Fault Diagnosis of Train Wheels Based on Empirical Mode Decomposition Generalized Energy	15
Yejian Chen, Xiaolong Wang, Yong Zhang and Zongyi Xing	
Detection of Squats Length Based on Axle Box Vibration Analysis	27
Zhi Yang, Zongyi Xing and Jie Jiang	
Profile Calibration of Dynamically Measuring Rail Wear Using LIS	37
Chao Wang, Yanfu Li, Dawei Liu and Hongli Liu	
Design and Implementation of Online Monitoring System of Contact System Tensioner	47
Wanjiao Han, Mingguang Liu and Hui Yu	
A Method of Detecting Pantograph Slide’s Abrasion Based on Image Processing	55
Saiyan Yu, Shuang Chen and Zongyi Xing	
Subsystem Characteristics-Based Modeling Method for the Simulation of Electromagnetic Compatibility of Rail Transit Vehicles	65
Dafa Jiang and Zhongcheng Jiang	
Failure Analysis and Discussion of Bogie Temperature Sensor for Electric Multiple Unit	77
Yantong Liu, Wenjun Wang, Jintian Wang, Wanxiu Teng and Zhaoyu Ma	
Study on Fault Diagnosis for Bearing Based on VMD-SVD and Extreme Learning Machine	87
Qiang Zhou, Yong Qin, Zhipeng Wang and Limin Jia	

Optimization of Wheel Re-profiling Strategy Based on a Statistical Wear Model	99
Bingkui Li, Zhi Yang, Zongyi Xing and Xudong Gao	
Research on Bearing Fault Diagnosis Based on Cyclic Statistics	109
Dong Yan and Xiukun Wei	
Research on Performance Degradation Prediction of Urban Rail Vehicle Suspension System	119
Tengteng Wang, Xiukun Wei and Limin Jia	
Safe Location of High-Speed Maglev Train	129
Jianfeng Liu, Hongze Xu and Gegerile	
Dynamic Performance Evaluation of Beijing Subway Airport Line Vehicle with Domestic Primary Suspension and Secondary Suspension	137
Xi Li and Yuan Zhang	
Analysis of Traffic Impedance Effects of Highway Rescue for Passenger Train Accidents	149
Bing Wang, Xiexing Shi, Wenzhe Wang, Huiping Sun and Adila	
Application of Multi-resolution State Domain Method in State Identification of Train Motor Rolling Bearings	157
Xinan Chen, Limin Jia, Yunxiao Fu and Yong Qin	
Analysis of EMUs Vehicle Body Voltage Caused by Pantograph-Catenary Off-Line Arc	165
Ying Wang, Fengyang Gao, Xiuqing Mu, Qiyao Li, Xiaoqiang Chen and Shiwen Liang	
Safety Analysis of Train Control System Based on Colored Petri Nets and System-Theoretic Process Analysis	175
Shaoqiang Hu, Daohua Wu and Huashen Wang	
Research on Reliability Evaluation of Traction Power Supply Equipment Considering Human Factors	185
Qianqian Yang, Sheng Lin and Ding Feng	
Bayesian Network-Based Reliability Analysis of High-Speed Train Bogie System	197
Shuting Zheng, Yong Qin, Li Wang, Limin Jia and Jinchu Zheng	
A Rail Corrugation Detection Method Based on Wavelet Packet Energy Entropy	205
Kuan Lang, Zongyi Xing, Wei Dong and Xudong Gao	

Pantograph Slide Plate Abrasion Detection Based on Deep Learning Network 215
 Yan Li and Xiukun Wei

Track Surface Defect Detection Based on Image Processing 225
 Yuxin Liu and Xiukun Wei

A Hybrid Temporal-Spatio Fusion Algorithm for Moving Pedestrian Detection in Traffic Scenes 233
 Aili Wang and Yuanyun Sun

Research of Online Nondestructive Monitoring of Composite Components Using FBG 247
 Xianfeng Wang, Zhongcheng Jiang, Xiaobo Liu and Jixiong Jiang

Study on the Adhesion Optimization Control Strategy and Its Application Based on Trend Forecasting 259
 Yunxin Fan, Biao Ye, Zhongcheng Jiang and Jixiong Jiang

Fault Identification for High-Speed Vehicle Suspension System Using Nonlinear Filtering 271
 Xiaozhong Zhang, Xiukun Wei and Limin Jia

Research on Gearbox Fault Isolation Based on VPMCD 281
 Xiukun Wei and Dong Yan

A Feature Extraction Method of Rolling Bearing Fault Signal Based on the Singular Spectrum Analysis and Linear Autoregressive Model 291
 Kun Xu, Gui Wang and Zongyi Xing

Passenger Flow Assignment of Evacuation Path in the Station Based on Time Reliability 301
 Chunwei Sun, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan and Jian Zhang

Simulation of Aging Characteristics of Power Switch Device Based on Saber 311
 Yujia Guo, Mengzhu Wang, Lei Wang, Ruichang Qiu and Guofu Chen

Evaluation of Emergency Evacuation Capacity of Subway Station Based on M/G/c/c 321
 Yacui He, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan and Jian Zhang

Research on the Abrasion Detection Technology of the Pantograph Slipper of Urban Rail Train 333
 Xuebing Hu, Yejian Chen, Yong Zhang and Zongyi Xing

Research on Fault Diagnosis of Vehicle Equipment for High-speed Railway Based on Case-Based Reasoning	343
Lixuan Chen, Dongxiu Ou and Hongjing Yao	
A Bayesian-MCMC Model to Assess Metro Train Collector Shoes Slider Degradation Under Different Materials	355
Yue Pan, Guoqiang Cai and Xi Li	
Design of High Power Factor Charger for Transportation Electrification	363
Fucun Li, Liang Guo, Hao Li and Guodong Qu	
Comparative Study of Fault Detection Algorithm Based on Multivariate Statistical Analysis	371
Shuyu Zhang	
Research on Cycling Energy Saving Based on Improved Violence Search	379
Liangliu Bai, Yong Zhang, Zongyi Xing, Xinrong Liu and Xuejin Wang	
Analysis of Running Stability of High-Power Locomotive Under Harmonic Wear Wheel	391
Qian Xiao, Zhixiang Luo and Jifeng Zheng	
Detection of Wheel Tread Wear Based on Laser Displacement Sensor	399
Jianyu Zhang, Zhi Yang, Yong Zhang and Zongyi Xing	
A Hybrid Filtering Algorithm for Pantograph Image Denoising	409
Weiwei Pei, Zongyi Xing and Zhuang Chen	
An Effective Railway Dynamic Invader Detection and Tracking Method	419
Qun Wang, Xiaofeng Li and Limin Jia	
Fault Diagnosis and Countermeasure for the Pantograph Automatic Lowering Under the Condition of Dual-Locomotive Running	429
Xiao Feng Zhong, De Chang Liu, Ping Luo and Hai Yi Zhou	
Simulation and Scheme Study on Braking Process of Urban Rail Transit	437
Xian Zhang, Yanhui Wang and Boyan Hao	
Study of the Destruction Resistance of Urban Rail Network Based on Random and Intentional Attack	447
Boyan Hao, Yanhui Wang, Xian Zhang and Huiru Zhang	

Method of Health State Assessment for Satellite Power System Based on Fuzzy Judgment and Variable Weight Layered Evaluation 455
 Pei Zhen Wang, Lei Wang, Ruichang Qiu, Lijun Diao and Yujia Guo

Intrusion Detection for High-Speed Railway Perimeter Obstacle 465
 Qinghong Liu, Yong Qin, Zhengyu Xie, Tangwen Yang and Gaoyun An

Research into Thermal Response Time Fault Diagnosis of Temperature Sensor in High-Speed EMU 475
 Xuguang Lin, Wenjun Wang, Bin Li, Jianwei Miao and Yantong Liu

Super-Gaussian Random Vibration Test Inductive Method Based on Measured Data 483
 Peng Wang, Hua Deng, Liuqing Xiao and Guotao Liu

RAM Assessment of CTCS-1 ATP Based on Enhanced Bayesian Network 497
 Yue Xu and Zhongtian Liu

A Novel Online Measuring System for Wheel Size Based on Laser Displacement Method 507
 Yong Zhang and Changgeng Wang

Sub-problems Interaction Analysis-Based Three-Level Decomposition Algorithm for Real-Time Train Scheduling and Routing Problems in Railway Stations 517
 Lijie Bai, Thomas Bourdeaud’huy, Emmanuel Castelain, Qi Zhang and Ziyuan Liu

Train Operation Intervals Optimization Based on a Group Dynamics Model 525
 Xuelei Meng, Bin Zhang and Limin Jia

Study on Energy Saving of Multi-vehicle Operation Based on Genetic Optimization Algorithm 533
 Xuejin Wang, Xiangxiang Zhou, Yong Zhang and Zongyi Xing

A Simulation Research for Fare Collection Equipments on Tiyuchang Station in Xi’an Subway 543
 Guang Li and Chaosheng Wu

An Integer Programming Model of Reinvestment Strategy Based Project Portfolio Selection and Scheduling with Constrained Resource 551
 Baolong Wang

Track Assignment Adjustment Problem in Complex Railway Passenger Stations 561
 Yinggui Zhang, Min An and Li Wang

Study on Dynamic Characteristics of Metro Stray Current Based on CDEGS 571
Mingjie Liu, Sheng Lin, Liping Zhao, Xiaohong Lin and Liang Chen

Analysis of Influencing Factors of Energy Saving and Emission Reduction for Railway Transportation of China 579
Liang Sun, Huiting Guo, Li Wang, Liang Chen, Jianhua Chen, Mei Liu and Wei Bao

Study on the Method of Passenger Transport Organization in Peak-Hour of Railway Station 587
Yuyang Nie and Xiaoning Zhu

Bayesian Network Inference on Departure Time Choice Behavior 599
Xian Li, Haiying Li, Linqiao Qin and Xinyue Xu

Metro Station Facility Layout Optimization Model Based on the Level of Service Reliability 609
Qingwen Jia, Jianrui Miao and Xu Chen

Monte Carlo Simulation Method Used in Reliability Evaluation of Railway Overhead Contact Line 619
Jianfeng Xu, Yuan Zhong and Shibin Gao

Optimization Study on Operation Plan of Intercity Trains Based on Minimization of Total Social Cost in Intercity Passenger Corridor 629
Yu-ling Ye, Yun-fei Zhou and Lu-qi Yang

High-Speed Railway Timetable Rescheduling Method: A Bi-level Integrated Programming Approach 639
Huiru Zhang, Limin Jia, Li Wang, Yong Qin and Min An

Characteristic Analysis of High-Speed Railway Network in China 649
Dan Guo, Yong Qin, Li Wang and Min An

Coordination of Transport Organization and Building Maintenance for Jixi Marshalling Station 661
Jianxiang Li and Xiaoning Zhu

A Method to Forecast and Estimate Traffic Demand of Airport Moving Sidewalk 671
Yuchun Ren and Xiaoning Zhu

Dynamic Multi-objective Optimization Problem of Container Intermodal Transport: An Empirical Analysis on the Belt and Road Initiative of China 679
Jue Hou

The Impact of Carbon Abatement Policies on Port Intermodal Freight Transportation Routing and Cost 689
 Shenghua Wang, Qi Zhang and Wenyan Wang

Improved Genetic Algorithm Based Passenger Flow Control in Subway 701
 Man Jiang

Study on Protection from Stray Current in the Metro System 709
 Hongbo Cheng, Zikang Xiao, Xun Wang and Nannan Sun

Energy Consumption Analysis of High-Speed Maglev Train 717
 Xiaochun Zhang, Siyuan Mu and Jinsong Kang

Risk Theory-Based Safety Evaluation of Passengers in Rail Transit Station 725
 Yunxiao Zheng, Yong Qin, Jianyuan Guo, Limin Jia and Jianghua Gao

Level of Service for Passage in Urban Rail Transit Station 735
 Sai Chen, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan and Jian Zhang

Acceptance Check Scheme of Integration Test and Commissioning for Urban Rail Transit 747
 Mao Tian

Forecast of Short-Term Passenger Flow of Urban Railway Stations Based on Seasonal ARIMA Model 759
 Zhirui Guang, Jun Yang and Jian Li

Research on the Route Choice Behavior of Subway Passengers Based on AFC Data 769
 Liping Xie, Haiying Li and Xinyue Xu

A Model of High-Density Passenger Boarding and Alighting in Urban Rail Transit Station 779
 Ning Jia and Yanhui Wang

A Multi-objective Programming Model for Subway Last Train Timetabling 791
 Wanxiao Xiang and Yanhui Wang

The Research of the Synergic Passenger Flow Control of Urban Rail Transit Line Operation 801
 Yiru Cui and Yanhui Wang

Application of Standardization Training of Operational Services in Rail Transit Enterprises 813
 Bin Xu and Yiming Shao

Research on Urban Rail Transit Line Collaborative Safety Operation	821
Xiaowei Shi and Yanhui Wang	
Research on Train Operation Daily Schedule Based on Balanced Use	829
Bo Wu, Zongyi Xing and Yao Zang	
Evaluating the Efficiency of Urban Public Transit Enterprises Based on DEA Approach with Preference	841
Rong Yu and Jiaqi Sun	
Research on BIM Application in High-Speed Railway OSC Operation Maintenance Management	853
Fei Meng, Tianyun Shi and Beisheng Liu	
Integrated Optimization Model on Maintenance Time Window and Train Timetabling	861
Lingyun Meng, Ce Mu, Xin Hong, Ran Chen, Xiaojie Luan and Tao Ma	
Controllability of Urban Rail Transit Network	873
Lu Zeng, Yong Qin, Jun Liu and Li Wang	
A Parallel Passenger Flow Management System for Outside Subway Stations Based on ACP Approach	885
Huijuan Zhou, Meijie Jia, Yu Liu, Qiang Zhang and Manrong Yuan	
Application of Triboelectric Nanogenerator in the Railway System	895
Laixin Geng, Sen Bian, Teng Li, Zhao Si and Zijun Wei	
Some Research on Suspension Gap Sensor of High-Speed Maglev Train	905
Jun Li, Jun Wu, Junyuan Tang and Shengjun Huang	
Superiority of Unit Serial Method Used in CRTS III-Type Pretensioning Ballastless Track Slab Production	915
Baoqun Wang, Lei Pei, Qikai Ai, Hongbin Zheng and Wei Si	
The Test Plan Design of Corrosion Fault Repeat and Material Selection for a Type of PCB	921
Xu Wang, Shaohua Du and Yingying Yuan	
Multi-hop Communication Protocol Optimization for the Linear Wireless Monitoring Network	929
Xiaoping Ma, Honghui Dong, Limin Jia, Yong Qin and Ruhao Zhao	

Application of High Accelerated Life Testing for an Integrated Control Board 941
 Yingying Yuan, Xuhui Zhang, Shaohua Du, Zaiwu Peng, Xiaochun Xiao and Jiani Zuo

Remote Monitoring System of Electrical Equipment Based on GPRS 951
 Yu Wang and Xuedong Jiang

YQuery: A Novel Privacy- and Integrity-Preserving Range Queries in Two-Tiered Sensor Networks 959
 Lvyou Yu, Jingwen Yu, Lei Wang and Linai Kuang

Simulation Study on High-Speed Milling Performance of Rail U71Mn Material 967
 Chao Pan, Jianqiang Zhang, Yuelei He, Haitao Xia and Baosheng Wang

Adaptive Control with Asymptotic Stability Guarantees for High-Speed Train Systems with Uncertain Input Nonlinearities 975
 Dongyue Yang and Lin Niu

Research on Key Technology of Platform Screen Door Control System for High-Speed Railway 985
 Zhifei Wang

Information Platform Framework for the Integrated Management of Intercity Arterial Roads 995
 Haiyang Wang

Study on a Mn–Cr–V–Ni Bainitic Forging Steel for Railway Transportation 1005
 Yunlei Lin and Qingyue Zhou

Design of and Research on an Attendance System Based on RFID and WSN Technologies for the Rail Transportation Industry 1013
 Chongjun Liu, Kuangang Fan, Yi Pan and Yindong Ren

Transient Overvoltage Analysis of Traction Power Supply System with Neutral Sections in China High-Speed Railway Using a State-Space Model

Xiaoqin Lv, Xiaoru Wang and Jiwei Zheng

Abstract This study proposes a complete state-space model of 2×25 kV traction power supply system with neutral sections in China high-speed railway where the catenary system is divided into several sections insulated with each other and is connected by neutral sections. First, the system is modularized by a neutral section module, a source module, autotransformer (AT) modules, series modules, and parallel modules which are represented with differential equations. Then the state-space model is built by means of slicing catenary system and combining with those modules. By calculating this proposed model, the transient processes of the electric train auto-passing through the neutral section are analyzed. Transient overvoltages occurred due to high-frequency oscillations as the circuit topology changing in each process.

Keywords Transient voltage · Neutral section · State-Space model
Traction power supply system · AT

1 Introduction

In recent years, high-speed railway (HSR) has developed quickly in China. 2×25 kV 50 Hz autotransformer (AT) traction power supply system (TPSS) is commonly adopted in China high-speed railways (HSRs). With more and more HSRs being put into operation, some problems such as harmonic resonance, catenary transient overvoltage occurred. In order to analyze the steady-state or transient problems, it is necessary to construct an accurate mathematical model of TPSS. The catenary system modeled as multiconductor transmission lines (MTL) is comprised of 14 conductors overhead or buried with a complex geometry, which can be regarded as groups [1–4]. TPSS has been represented with a nodal

X. Lv · X. Wang (✉) · J. Zheng
School of Electrical Engineering, Southwest Jiaotong University,
Chengdu 610031, Sichuan Province, China
e-mail: xrwang@home.swjtu.edu.cn

admittance matrix [3], which is the model for steady-state analysis both at power frequency and harmonic frequencies. In [5], the transient mathematical model of TPSS considering its real structure represented as the state-space equation has been built for dynamic analysis. However, all of these mathematical models for TPSS steady-state or transient analysis do not relate to a neutral section which connect two-side sections of the catenary system. As trains passing through neutral sections, there will be transient overvoltages which can harm the equipments of trains. Some studies for this transient overvoltage phenomenon have been performed. In [6], the time-domain simulations were carried out. In [7] and [8], the state-space equations were built for calculating the overvoltage values. However, these studies in which the time-domain simulation model or the state-space model was built did not relate to the real structures of TPSS and the neutral sections, only modeled TPSS as an equivalent impedance and the neutral section as one conductor.

This whole system involving an electric supply substation (ESS), ATs, the catenary system can be represented by modularized units, combining in parallel or in series [5]. The neutral section of double traction network has not only six grouped conductors, but also two additional neutral lines. Therefore, the mathematical model of the neutral section is different compared with the power supply catenary. In this paper, the modularized mathematical model of neutral sections is deduced, and the state-space model of TPSS involving a neutral section for China high-speed railway TPSS is built. According to TPSS configuration and system operation from ESS to sectioning post (SP), the state-space model can be easily established by those modularized differential equations.

With the proposed state-space model of TPSS, the pantograph transient voltages are analyzed in this paper as electric train auto-passing trough the neutral section. The circuit topology changes due to the train arriving at different position of the neutral section. Therefore, with the equivalent circuit of the train connecting to the system, the state-space model can be a little different. The mathematical models used for analyzing transient overvoltages are built in this paper, and the calculating results are obtained and analyzed.

2 Mathematical Model of TPSS

China TPSS consists of the ESS equipped with V/x transformers, which primary windings connect to 220 kV power grid, and each secondary winding presents three terminals. The three terminals of an AT with 2:1 ratio along lines has two ends, which are connected to contact wires and feeders, and one center, which connects to the rails. The contact wires, rails, feeders in up, and down tracks are respectively in parallel. A neutral section is located between ESS right and left sections or at the end of each section. Figure 1 shows a typical AT traction power supply system, where lines can be supposed to slice by a neutral section, AT parallel posts, electric train positions, fault locations, or SPs into some multiport networks with different length l_i . The phase β and phase α denote the two different phases of voltages respectively feeding to ESS right and left section.

2.1 Mathematical Model of TPSS Without the Neutral Section

Modularized one-side section consists of a source module, AT modules, parallel modules, series modules, and the equivalent circuit is shown in Fig. 2. According to conductor distances and physical parameters, the series impedance matrix \mathbf{Z} per unit can be calculated by Carson equations, the potential coefficient matrix \mathbf{P} can be calculated by electromagnetic field theory [9].

As in [5], the differential equations of each module and the state-space model of TPSS without a neutral section can be represented as follows.

The differential equation for a source module and a AT module are given by

$$\frac{di_s}{dt} = -\frac{R_{ST}}{L_{ST}}i_s + \mathbf{M}_1\mathbf{u}_{sj} + \frac{1}{L_{ST}}e_\alpha \quad (1)$$

$$\frac{di_z}{dt} = -\frac{R_T}{L_T}i_z + \mathbf{M}_2\mathbf{u}_{zj}, \quad (2)$$

where the source current is i_s , the leakage current is i_z , \mathbf{u}_{sj} and \mathbf{u}_{zj} are the node voltage matrices, $\mathbf{M}_1 = \begin{bmatrix} -\frac{1}{L_{ST}} & 0 & \frac{1}{L_{ST}} & 0 & 0 & 0 \end{bmatrix}$, $\mathbf{M}_2 = \begin{bmatrix} -\frac{1}{2L_T} & \frac{1}{L_T} & -\frac{1}{2L_T} & 0 & 0 & 0 \end{bmatrix}$.

The equation of a series module is given by

$$\frac{d}{dt}\mathbf{i}_{jh} = -\mathbf{L}_j^{-1}\mathbf{R}_j\mathbf{i}_{jh} + \mathbf{L}_j^{-1}\mathbf{u}_j - \mathbf{L}_j^{-1}\mathbf{u}_h, \quad (3)$$

where the inductance current matrix is \mathbf{i}_{jh} , \mathbf{u}_j and \mathbf{u}_h are node voltage matrices respectively forming by slicing j and slicing h .

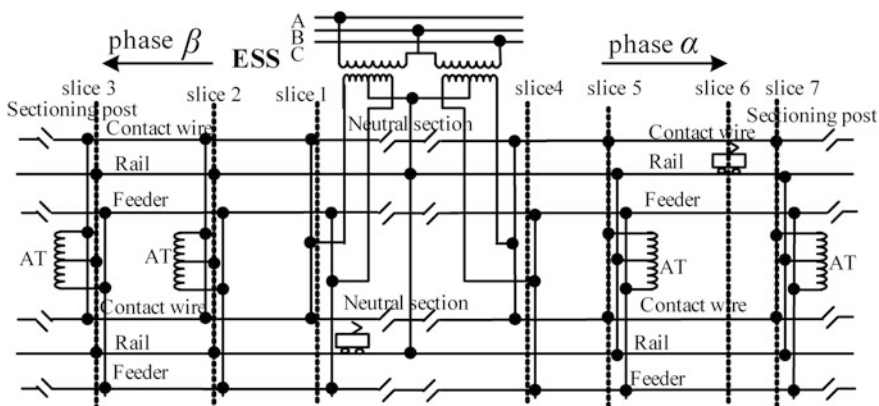


Fig. 1 AT traction power supply system

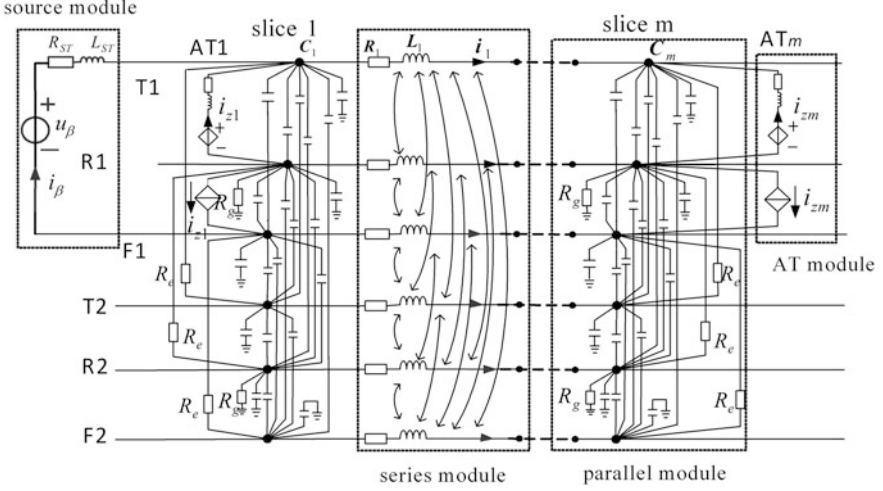


Fig. 2 Equivalent circuit of one-side section with ESS

The equation of a parallel module is given by

$$\left. \begin{aligned} \frac{du_j}{dt} &= C_j^{-1}(-G_e u_j - G_g u_j + i_{jc}) \\ i_{jc} &= N_s i_{js} + N_{at} i_{jz} + i_{j1} - i_{j2} + M_3 i_{LL} \end{aligned} \right\}, \quad (4)$$

where the source module, AT, two series modules and load injection current matrices are respectively i_{js} , i_{jz} , i_{j1} (i_{j2}), and i_{LL} . The components of total injection current i_{jc} relate to modules which connect to nodes j . And

$$N_s^T = [1 \ 0 \ -1 \ 0 \ 0 \ 0], N_{at}^T = [1 \ -2 \ 1 \ 0 \ 0 \ 0], g_e = 1/R_e, g_g = 1/R_g,$$

$$G_e = \begin{bmatrix} g_e & 0 & 0 & -g_e & 0 & 0 \\ 0 & g_e & 0 & 0 & -g_e & 0 \\ 0 & 0 & g_e & 0 & 0 & -g_e \\ -g_e & 0 & 0 & g_e & 0 & 0 \\ 0 & -g_e & 0 & 0 & g_e & 0 \\ 0 & 0 & -g_e & 0 & 0 & g_e \end{bmatrix}, G_g = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g_g & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g_g & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$M_3^T = [-1 \ 1 \ 0 \ 0 \ 0 \ 0].$$

Therefore, according to the real operation from substation to section post, combining with modularized Eqs. (1)–(4), the state-space model of TPSS without the neutral section can be obtained.

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \quad (5)$$

2.2 Neutral Section Module

The neutral section makes two-side sections weak interaction (see Fig. 1). A typical catenary structure of China HSR neutral section is shown in Fig. 3. A neutral line N with a distance of 0.5 m from contact wires and a length of approximately 300 m is located in this area. Taking insulator positions for four slicing points ($ns1$, $ns2$, $ns3$, $ns4$), this section is divided into two parts with the same length, one is the left side with neutral lines ($N1$ for up track and $N2$ for down track), the other is the right side with neutral lines, which can both be equivalent to π -circuit. Then, the equivalent circuit for a neutral section can be illustrated in Fig. 4.

These two π -circuit sections are respectively parallel with C_{nsl} and C_{nsr} , series with R_{nsl} , L_{nsl} and R_{nsr} , L_{nsr} , and $C_{nsl} = C_{nsr} = C_{ns}$, $R_{nsl} = R_{nsr} = R_{ns}$, $L_{nsl} = L_{nsr} = L_{ns}$.

C_1 denotes the capacitance matrix in parallel module of slicing 1 in Fig. 2. For the nodes in slicing $ns1$ (see Fig. 4), the integrated capacitance C_{ns1} combined C_{nsl} with C_1 is

$$C_{ns1} = \begin{bmatrix} E_2 & E_3 \\ E_3^T & E_1 \end{bmatrix}, \quad (6)$$

$$\text{where } E_3 = \begin{bmatrix} C_{N1T1} & C_{N1R1} & C_{N1F1} & C_{N1T2} & C_{N1R2} & C_{N1F2} \\ C_{N2T1} & C_{N2R1} & C_{N2F1} & C_{N2T2} & C_{N2R2} & C_{N2F2} \end{bmatrix}, E_2 = \begin{bmatrix} C_{N1} & C_{N1N2} \\ C_{N1N2} & C_{N2} \end{bmatrix},$$

$$E_1 = \begin{bmatrix} C_{T1} & \cdots & C_{T1R2} & C_{T1F2} \\ C_{T1R1} & \cdots & C_{R1R2} & C_{R1F2} \\ \vdots & \vdots & \vdots & \vdots \\ C_{T1F2} & \cdots & C_{F2R2} & C_{F2} \end{bmatrix} + C_1$$

The equations of nodal voltages in slicing $ns1$ are given by

Fig. 3 Neutral overlapping section of China HSR

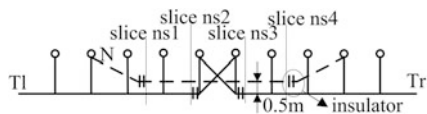
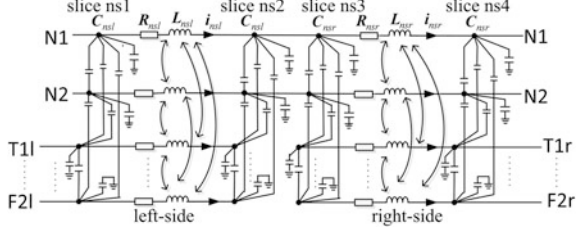


Fig. 4 Equivalent circuit of the neutral section



$$\left. \begin{aligned} \frac{d\mathbf{u}_{ns1}}{dt} &= \mathbf{C}_{ns1}^{-1} (-\mathbf{G}_{en}\mathbf{u}_{ns1} - \mathbf{G}_{gn}\mathbf{u}_{ns1} + \mathbf{i}) \\ \mathbf{i} &= \mathbf{N}_{sn}\mathbf{i}_s + \mathbf{N}_{atn}\mathbf{i}_{z1} - \mathbf{i}_{nsl} - \mathbf{F}_g\mathbf{i}_{l1} + \mathbf{M}_{3n}\mathbf{i}_{LL} \end{aligned} \right\}, \quad (7)$$

where \mathbf{u}_{ns1} is the nodal voltage matrix of eight conductors in the left side of neutral section, \mathbf{i}_s , \mathbf{i}_{z1} , \mathbf{i}_{LL} are injection currents from source module, AT module and load module, respectively, \mathbf{i}_{l1} , \mathbf{i}_{nsl} are the currents from the first series module of left-side catenary system and the series module of left-side neutral section respectively. Therefore,

$$\mathbf{u}_{ns1}^T = [u_{N1l} \ u_{N2l} \ u_{T1l} \ u_{R2l} \ u_{F1l} \ u_{T2l} \ u_{R2l} \ u_{T2l}], \mathbf{F}_g = \begin{bmatrix} \mathbf{F}_5 \\ \mathbf{F}_6 \end{bmatrix},$$

$\mathbf{N}_{sn} = \begin{bmatrix} \mathbf{F}_7 \\ \mathbf{N}_s \end{bmatrix}$, $\mathbf{N}_{atn} = \begin{bmatrix} \mathbf{F}_7 \\ \mathbf{N}_{at} \end{bmatrix}$, $\mathbf{M}_{3n} = \begin{bmatrix} \mathbf{F}_7 \\ \mathbf{M}_3 \end{bmatrix}$, $\mathbf{G}_{en} = \begin{bmatrix} \mathbf{F}_7 \\ \mathbf{G}_e \end{bmatrix}$, $\mathbf{G}_{gn} = \begin{bmatrix} \mathbf{F}_7 \\ \mathbf{G}_g \end{bmatrix}$, and \mathbf{F}_5 is a (2×6) zero matrix, \mathbf{F}_6 is a (6×6) unit matrix, \mathbf{F}_7 is a (2×1) unit matrix.

Equations of \mathbf{u}_{ns4} in slicing $ns4$ can be represented by the same form.

Combining slicing $ns2$ with $ns3$, a (14×14) voltage matrix \mathbf{u}_{n0} can be obtained. Rearrange the elements in matrix \mathbf{C}_{ns1} in Eq. (6), a new (14×14) matrix \mathbf{E} is given by

$$\mathbf{E} = \begin{bmatrix} 2 \times \mathbf{E}_4 & \mathbf{E}_5^T & \mathbf{E}_5^T \\ \mathbf{E}_5 & \mathbf{E}_6 & 0 \\ \mathbf{E}_5 & 0 & \mathbf{E}_6 \end{bmatrix}, \quad (8)$$

where

$$\mathbf{E}_4 = \begin{bmatrix} C_{N1} & C_{N1N2} & C_{N1R1} & C_{N1R2} \\ C_{N1N2} & C_{N2} & C_{N2R1} & C_{N2R2} \\ C_{R1N1} & C_{R1N2} & C_{R1} & C_{R1R2} \\ C_{R2N1} & C_{R2N2} & C_{R1R2} & C_{R2} \end{bmatrix}, \mathbf{E}_5 = \begin{bmatrix} C_{T1N1} & C_{T1N2} & C_{T1R1} & C_{T1R2} \\ C_{F1N1} & C_{F1N2} & C_{F1R1} & C_{F1R2} \\ C_{T2N1} & C_{T2N2} & C_{T2R1} & C_{T2R2} \\ C_{F2N1} & C_{F2N2} & C_{F2R1} & C_{F2R2} \end{bmatrix},$$

$$\mathbf{E}_6 = \begin{bmatrix} C_{T1} & C_{T1F1} & C_{T1F2} & C_{T1F2} \\ C_{F1T1} & C_{F1} & C_{F1T2} & C_{F1F2} \\ C_{T2T1} & C_{T2F1} & C_{T2} & C_{T2F2} \\ C_{F2T1} & C_{F2F1} & C_{F1T2} & C_{F2} \end{bmatrix}.$$

Then, the equation of voltage \mathbf{u}_{n0} is

$$\frac{d}{dt}\mathbf{u}_{no} = \mathbf{E}^{-1}(\mathbf{F}_a\mathbf{i}_{nsl}) + \mathbf{E}^{-1}(\mathbf{F}_b\mathbf{i}_{nsr}), \quad (9)$$

where $\mathbf{u}_{n0}^T = [u_{N1}, u_{N2}, u_{R1}, u_{R2}, u_{T1l}, u_{F1l}, u_{T2l}, u_{F2l}, u_{T1r}, u_{F1r}, u_{T2r}, u_{F2r}]$, and

$$\mathbf{F}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{F}_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\mathbf{F}_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{F}_a = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \mathbf{F}_3 \\ \mathbf{F}_4 \end{bmatrix}, \mathbf{F}_b = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \mathbf{F}_4 \\ \mathbf{F}_3 \end{bmatrix},$$

$\mathbf{F}_4 = (4 \times 8)$ zero matrix.

The current equations of series modules, respectively in left-side and right-side neutral sections, are given by

$$\frac{d}{dt}\mathbf{i}_{nsl} = \mathbf{L}_{ns}^{-1}(\mathbf{u}_{ns1} - \mathbf{F}_e\mathbf{u}_{n0} - \mathbf{R}_{ns}\mathbf{i}_{nsl}) \quad (10)$$

$$\frac{d}{dt}\mathbf{i}_{nsr} = \mathbf{L}_{ns}^{-1}(\mathbf{u}_{ns4} - \mathbf{F}_d\mathbf{u}_{n0} - \mathbf{R}_{ns}\mathbf{i}_{nsr}), \quad (11)$$

where

$$\mathbf{i}_{nsl}^T = [i_{N1l} \quad i_{N2l} \quad i_{T1l} \quad i_{R1l} \quad i_{F1l} \quad i_{T2l} \quad i_{R2l} \quad i_{F2l}],$$

$$\mathbf{F}_e = [\mathbf{F}_1^T \quad \mathbf{F}_2^T \quad \mathbf{F}_3^T \quad \mathbf{F}_4^T]$$

$$\mathbf{i}_{nsr}^T = [i_{N1r} \quad i_{N2r} \quad i_{T1r} \quad i_{R1r} \quad i_{F1r} \quad i_{T2r} \quad i_{R2r} \quad i_{F2r}],$$

$$\mathbf{F}_d = [\mathbf{F}_1^T \quad \mathbf{F}_2^T \quad \mathbf{F}_4^T \quad \mathbf{F}_3^T].$$

When considering a neutral section connecting the two sides, the current equation of first series module connected to ESS with series \mathbf{L}_1 and \mathbf{R}_1 , taking left-side section as an example, becomes

$$\frac{d}{dt}\mathbf{i}_{12} = -\mathbf{L}_1^{-1}\mathbf{R}_1\mathbf{i}_{12} + \mathbf{L}_1^{-1}\mathbf{F}_g^T\mathbf{u}_{ns1} - \mathbf{L}_1^{-1}\mathbf{u}_2, \quad (12)$$

where \mathbf{u}_2 is the voltage matrix of nodes forming by slicing 2 (see Fig. 1).

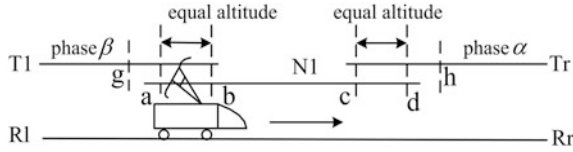


Fig. 5 Relative position of pantograph, contact wires, and neutral lines

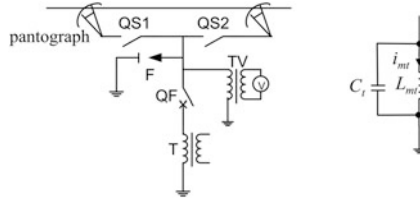
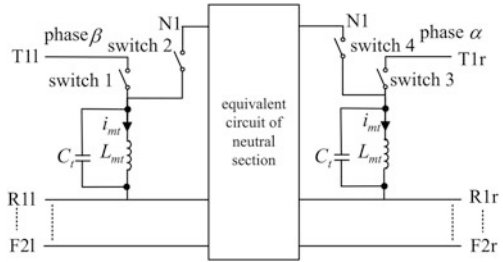


Fig. 6 (Left) Schematic diagram of electric train line-side circuit; QS: disconnector, QF: circuit breaker, F: arrester, TV: potential transformer, T: transformer; (right) Equivalent circuit with QF turned off

Fig. 7 Schematic diagram of transient processes



3 Transient Analyses of Auto-Passing Neutral Section

Automatic onboard neutral section passing is mostly applied in China HSR. When an electric train coming from phase β passes through neutral section to get to phase α , as shown in Fig. 5, the main circuit breaker (QF in Fig. 6) of the train will be turned off at position g, and be turned on at position h. The equivalent circuit in Fig. 6 represents the equivalence of an electric train with QF turned off. It consists of a magnetizing inductance L_{mi} referred to primary side of potential transformer, and a line to earth equivalent capacitance C_t of the train.

As illustrated in Fig. 5, the pantograph which contacts with the contact wire and the neutral line at the same time in equal altitude segments (a–b, c–d), and only contacts with the neutral line in b–c segment, will experience an electricity-neutral-electricity process. This process can be described as four states, as shown in Fig. 7 with electric train running on up track. (1) The train coming

from phase β (with switch 1 on, and switch 2, 3, 4 off), arrives at position a (turn on switch 2). (2) The train arrives at position b (turn off switch 1). The pantograph connects to the neutral line only. (3) The train arrives at position c (with switch 1, 2 off, and switch 4 on, turn on switch 3). The pantograph connects to the right-side contact wire $T1r$ and the neutral line. (4) The train arrives at position d (turn off switch 4), with the pantograph only connecting to the right-side contact wire $T1r$.

For state 1, additional variable i_{mt} associated with the equivalent circuit of Fig. 6 is put into the matrix x_n in Eq. (15), and the state vector becomes

$$\mathbf{x}_{n2} = [\mathbf{x}_l \quad \mathbf{x}_{ns} \quad \mathbf{x}_r \quad i_{mt}] \quad (17)$$

The additional equation with switch 1 on is

$$\frac{di_{mt}}{dt} = \frac{1}{L_{mt}} \mathbf{F}_5 \mathbf{u}_{ns1}, \quad (18)$$

where $\mathbf{F}_5 = [0 \quad 0 \quad 1 \quad -1 \quad 0 \quad 0 \quad 0 \quad 0]$.

An equivalent small resistance R_d can be substituted for switch 2 when switch 2 is on. The parallel module equation in slicing $ns1$ of Eq. (7) is substituted by the following formula,

$$\frac{d\mathbf{u}_{ns1}}{dt} = \mathbf{C}_{ns1}^{-1} (-\mathbf{G}_{en} \mathbf{u}_{ns1} - \mathbf{G}_{gn} \mathbf{u}_{ns1} + \mathbf{i} + \mathbf{i}_t) \quad (19)$$

$$\mathbf{i}_t = \mathbf{F}_5^T i_{mt} + k \frac{1}{R_d} \mathbf{F}_6 \mathbf{u}_{ns1}, \quad (20)$$

where k is equal to 1 when switch 2 is on, and equal to 0 when switch 2 is off, $\mathbf{F}_6^T = [\mathbf{F}_{61} \quad \mathbf{F}_{62} \quad -\mathbf{F}_{61} \quad \mathbf{F}_{63}]$, $\mathbf{F}_{61} = [-1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$, \mathbf{F}_{62} is a (1×8) zero matrix, \mathbf{F}_{63} is a (5×8) zero matrix. The element \mathbf{E}_1 of capacitance \mathbf{C}_{ns1} in Eq. (6) is replaced with \mathbf{E}'_1 , then we can get the capacitance matrix \mathbf{C}_{ns1-t} .

$$\mathbf{E}'_1 = \begin{bmatrix} C_{T1} + C_t & C_{T1R1} - C_t & \cdots & C_{T1F2} \\ C_{T1R1} - C_t & C_{R1} + C_t & \cdots & C_{R1F2} \\ \vdots & \vdots & \vdots & \vdots \\ C_{T1F2} & C_{F2R2} & \cdots & C_{F2} \end{bmatrix} + \mathbf{C}_1 \quad (21)$$

The definitions of the rest variables in Eq. (19) are the same as in Eq. (7).

Hence, referring to Eq. (14), the state matrix becomes as Eq. (22).

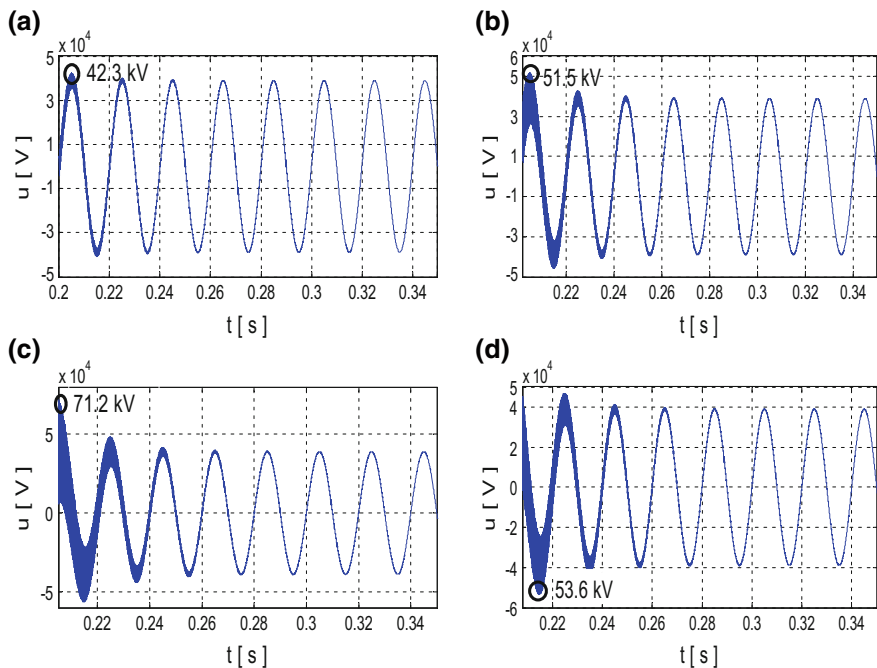


Fig. 8 Pantograph voltage in transient state 1. **a** Switching at 0.2 s. **b** Switching at 0.202 s. **c** Switching at 0.205 s. **d** Switching at 0.208 s

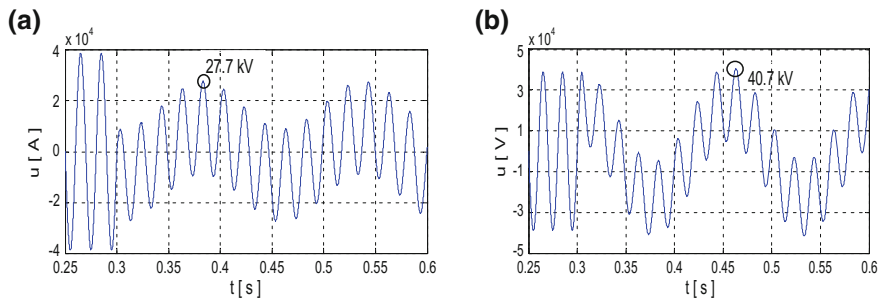


Fig. 9 Pantograph voltage in transient state 2. **a** Switching at 0.3 s. **b** Switching at 0.305 s

$$\mathbf{E}''_1 = \begin{bmatrix} C_{T1} & C_{T1R1} & \cdots & C_{T1F2} \\ C_{T1R1} & C_{R1} + C_t & \cdots & C_{R1F2} \\ \vdots & \vdots & \vdots & \vdots \\ C_{T1F2} & C_{F2R2} & \cdots & C_{F2} \end{bmatrix} + \mathbf{C}_1, \mathbf{E}''_2 = \begin{bmatrix} C_{N1} + C_t & C_{N1N2} \\ C_{N1N2} & C_{N2} \end{bmatrix}$$

$$\mathbf{E}''_3 = \begin{bmatrix} C_{N1T1} & C_{N1R1} - C_t & C_{N1F1} & C_{N1T2} & C_{N1R2} & C_{N1F2} \\ C_{N2T1} & C_{N2R1} & C_{N2F1} & C_{N2T2} & C_{N2R2} & C_{N2F2} \end{bmatrix}$$

As the solution of state 2, assume the time when the electric train arriving position b is t_2 s. Figure 9 represents the pantograph transient voltages with different switching times. It shows that the phase angles of source voltage when pantograph instantaneously disconnect with the contact wire T11 can also affect the amplitudes of oscillating voltage, which are lower than the amplitudes of voltages in state 1.

The same analysis method can be used for transient voltages in state 3 and state 4. As the train passes through the neutral section, the circuit topology changes. High-frequency oscillations lead to transient overvoltages.

4 Conclusion

This paper presents a state-space model of AT TPSS with the neutral section, which focuses on truly reflecting the structure of TPSS. The system is modularized by a neutral section module, parallel modules, series modules, the ESS, and ATs represented by differential equations. The proposed model is based on choosing and combining with those modularized differential equations in accordance with the specific structure and operation conditions.

Transient voltages as electric train passing through neutral section are analyzed using this state-space model. The whole process can be divided into four transient processes. As circuit topology changes in each process, high-frequency oscillation leads to transient overvoltage, in which amplitude is affected by the phase angle of source voltage at that moment.

References

1. Battistelli L, Pagano M, Proto D (2011) 2×25 -kV 50 Hz high-speed traction power system: short-circuit modeling. *IEEE Trans Power Del* 26(3):1459–1466
2. Lee H, Lee C, Jang G et al (2006) Harmonic analysis of the Korean high speed railway using the eight-port representation model. *IEEE Trans Power Del* 21(2):976–986
3. Wu ML, Roberts C, Hillmansen S (2010) Modeling of AC feeding system of electric railways based on a uniform multi-conductor chain circuit topology. In: *IET conference on railway traction systems (RTS 2010)*, pp 1–5

4. He ZY, Hu HT, Zhang YF et al (2014) Harmonic resonance assessment to traction power-supply system considering train model in China high-speed railway. *IEEE Trans Power Del* 29(4):1735–1743
5. Lv XQ, Wang XR (2017) State-space model of high-speed railway traction power-supply system. *Proc CSEE* 37(3):867–868 (in Chinese)
6. Jiang XF, Heng ZY, Hu HT (2013) Analysis on electromagnetic transient process of electric multiple unit passing neutral section devices. *J China Railway Soc* 35(12):30–35 (in Chinese)
7. Zhang XY (2009) Research on overvoltage mechanism and prevention in the network-locomotive coupling of the high speed and heavy haul railways. Southwest Jiaotong University 33–80 (in Chinese)
8. Wang Y, Liu ZG, Mu XQ et al (2015) Study on electromagnetic transient process in electric split-phase system considering electrical coupling of viaduct. *J China Railway Soc* 37(11): 37–43 (in Chinese)
9. Dommel HW (1992) EMTP theory book. In: Microtran power system analysis corporation. Vancouver, BC, Canada

Fault Diagnosis of Train Wheels Based on Empirical Mode Decomposition Generalized Energy

Yejian Chen, Xiaolong Wang, Yong Zhang and Zongyi Xing

Abstract In this paper, a fault diagnosis method of train wheel based on empirical mode decomposition (EMD) generalized energy is proposed. First, EMD is applied to rail vibration signal to obtain the intrinsic mode functions (IMFs) and weight coefficients of each IMF are calculated. Then, quantitative value of EMD generalized energy is determined by calculating weighted sum of IMFs' energy. Finally, the security domain threshold of EMD generalized energy is determined to distinguish faulted wheel. Experiment results by using simulation data showed that the proposed method can distinguish normal and faulted wheels and the accuracy reached above 90%.

Keywords IMF · EMD · Safety region boundary · Fault diagnosis

1 Introduction

The wheelset is one of the extremely important components of traveling system in the system of urban rail transit. The wheelset needs to bear the large static forces and dynamic forces, which is an important factor of trains' safe operation. Meanwhile, it is also the key test object in the security check of moving train system [1, 2]. Therefore, real-time monitoring of the train wheel has important practical significance for the safety of passengers and the rail vehicle operations [3, 4]. Vittorio Belotti et al. [5] installed the vibration acceleration transducers on one side of track, and found a fault diagnosis method to detect and estimate quantitatively the size of the flat scar of wheel by using wavelet transform. Stephen Lechowicz and et al. [6] designed and developed the wheel load distribution and fault defects monitoring system based on database application technology.

Y. Chen · X. Wang · Y. Zhang (✉) · Z. Xing
School of Automation, Nanjing University of Science and Technology,
No.200 Xiao Ling Wei, Nanjing, Xuan Wu District, China
e-mail: 34445721@qq.com

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_2

Based on EMD generalized energy, a diagnosis method of decomposing rail vibration signal EMD is proposed in this paper, which is to obtain effective IMF and to work out the threshold of EMD generalized energy security zone of a normal wheel for the partition between normal wheels and faulty wheels. The method has high veracity for the online identification of normal wheels and faulty wheels.

2 EMD Generalized Energy

Extraction of signal state characteristic value is the key to monitor and assess the wheels. The method of extraction for the characteristics of the signal state includes the method of time-domain parameters and the method of time-frequency parameters. The method of time-domain parameters cannot obtain the characteristics of the frequency domain signal, so the accuracy of the results cannot reach the application requirement. The time-frequency parameters are more comprehensive in characteristic extraction, so the accuracy of system fault diagnosis could be significantly improved. Therefore, this paper adopts the method of time-frequency parameters and decomposes the signals by using EMD, and then extracts the corresponding characteristic value.

Energy of vibration signal is the sum of squares of the absolute value of the signal amplitude, and the method of calculation is as follows:

$$E_n = \sum_{i=1}^N |x_i|^2 \quad (1)$$

E_n is the energy value of vibration signal, and x_i is the amplitude at time i .

Generalized energy value is the sum of energy value of IMF elements which could be obtained by decomposing the vibration signal by EMD.

$$E = \sum_{i=1}^N E_i \quad (2)$$

E is the generalized energy value of this vibration signal, E_i is the energy value of the i th IMF, and N is the IMF order after EMD decomposition.

3 The Method of Train Wheels Fault Diagnosis

According to the characteristics of the rail vibration signal, the single eigenvalue security domain estimation method based on EMD generalized energy is proposed. Estimating threshold value of security domain from generalized energy value helps to assess the state of the wheel failure.

EMD-generalized energy can reflect the change of contact force between wheels and rail, so the threshold value of security domain of the wheel fault state can be decided through EMD generalized energy, which helps to judge wheels fault state. The detailed algorithm flowchart is as follows (Fig. 1).

3.1 De-noising of Rail Vibration Signal

The wheel failure state assessment is based on the rail vibration signal, so it is necessary to collect the rail vibration signal. The existence of much noise in the signal can adversely affect essential characteristic of rail vibration signal, which is not good for analyzing signal characteristic. So it is necessary to reduce the noise, and to improve the accuracy of analyzing the characteristic.

In this paper, wavelet threshold de-noising method [7] is employed to filter signals. The principles of de-noising are as follows: after the original signal is decomposed by wavelet, the component of smaller wavelet coefficients is assigned to noise signal, and the component of larger wavelet coefficients is assigned to useful signal. So the wavelet coefficient that is less than this threshold value is set to 0, and noise signal is effectively filtered from original signal.

After signal de-noising, SNR is used to evaluate the performance of de-noising algorithm. The larger the value of SNR, the better is the noise reduction effect of this method. The formula of SNR is as follows, and the unit is dB:

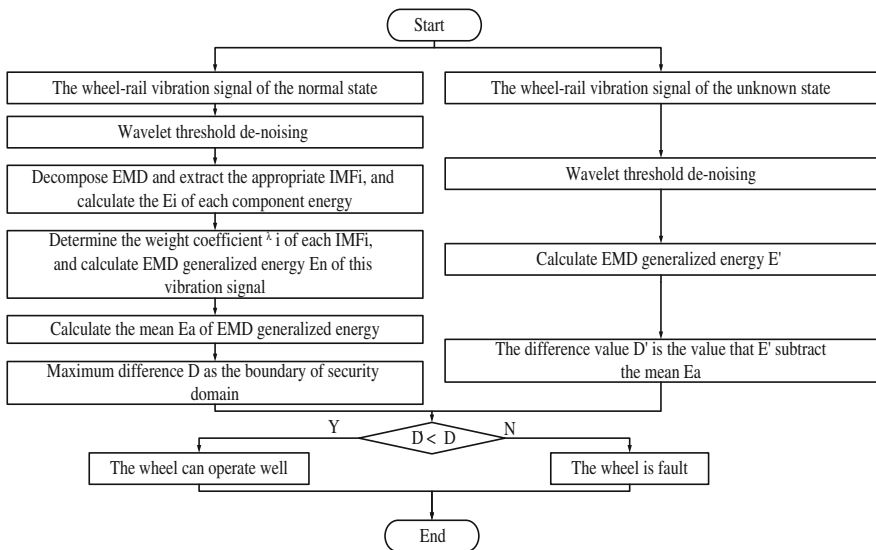


Fig. 1 The algorithm flowchart of wheel failure state assessment

$$\text{SNR} = 10 \log_{10} \left[\frac{\sum_n f^2(n)}{\sum [f(n) - \hat{f}(n)]^2} \right] \quad (3)$$

In Eq. (3), $f(n)$ is the original signal, $\hat{f}(n)$ is the signal after de-noising.

3.2 EMD Decomposition and IMF Decomposition of Rail Vibration Signal

The EMD algorithm proposes that a complex signal can be decomposed as the sum of the IMF component, and every IMF must conform to two basic conditions meanwhile: (1) In the signal of complete data segment, the difference between the amount of extreme point and zero-crossing point must be less than or equal to 1; (2) At any time, the mean value of the upper and lower envelopes formed by local extreme points is zero. The initial signal $x(t)$, concrete steps of EMD are as follows:

- (1) First, to recognize all the extreme points of the original signal, and connect these points by cubic spline curve to form the upper and lower envelopes of signal, and calculate the mean value $m_1(t)$ of the upper and lower envelopes.

$$m_1(t) = \frac{e_{\text{upp}} + e_{\text{low}}}{2} \quad (4)$$

- (2) $h_1(t)$ is got by subtracting $m_1(t)$ from x_i . If $h_1(t)$ is a basic IMF component, the decomposition is stopped. If $h_1(t)$ does not satisfy the condition of the IMF component, $h_1(t)$ is regarded as a new $x(t)$. Repeat the previous procedure to filter k times, until the basic IMF component $h_{1k}(t)$ is obtained.
- (3) The first basic IMF component $c_1(t)$ is obtained, $c_1(t) = h_{1k}(t)$. As its time characteristic scale is smallest, the frequency of $c_1(t)$ is highest. The rest part $r_1(t)$ is obtained, and the formula is as follows:

$$r_1(t) = x(t) - c_1(t) \quad (5)$$

- (4) $r_1(t)$ is regarded as the initial signal, and the processes of (1) (2) (3) are repeated for $r_1(t)$ to gradually separate each IMF component:

$$\begin{aligned} r_2(t) &= r_1(t) - c_2(t) \\ &\dots \\ r_n(t) &= r_{n-1}(t) - c_n(t) \end{aligned} \quad (6)$$

When $c_n(t)$ is less than a certain threshold, or $r_n(t)$ is monotone function, the cycle will be stopped.

(5) The original signal $x(t)$ will be obtained, and the formula is as follows:

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (7)$$

Among them, $c_1(t), c_2(t), \dots, c_n(t)$ are IMF components in different stages. $r_n(t)$ is the residual value, and represents the average trend of the signal.

In this paper, the IMF selection algorithm based on the combination of time-domain kurtosis and frequency dispersity is adopted. By choosing through time-domain kurtosis in low frequency and frequency dispersity in high frequency, the discrimination for effective IMF component will be improved.

Viewed from time domain, the IMF component with larger value of time-domain kurtosis includes more impact component of the signal. The calculation formula of time-domain kurtosis of the signal $x(t)$ is as follows:

$$\text{Kurtosis}(x) = \frac{(1/N) \cdot \sum (x_i - \bar{x})^4}{\text{Std}(x)^4} \quad (8)$$

Through the experimental analysis, choose 3 to be the general threshold, and the value higher than 3 is efficient component, and the value lower than 3 is inefficient component.

Viewed from frequency domain, frequency dispersity reflects the encircled energy of the signal in the frequency band. The calculation formula of frequency dispersity of the signal $x(t)$ is as follows:

$$B^2 = \frac{4\pi}{E_x} \int_{-\infty}^{+\infty} (f - f_m) |x(f)|^2 df \quad (9)$$

In Eq. (9), E_x is the energy of the signal. Through the experimental analysis, the general threshold is chosen to be 10, and the value higher than 10 is inefficient component and the value lower than 10 is efficient component.

When the value of time-domain kurtosis and the value of frequency dispersity are calculated from each IMF component through EMD decomposition, the IMF component that simultaneously satisfies the screening requirements of time-domain kurtosis and frequency dispersity is selected as efficient component to calculate its respective energy value E_i by Eq. (10).

$$E_i = \int_{-\infty}^{+\infty} |c_i(t)|^2 dt, (i = 1, 2, 3, \dots) \quad (10)$$

In Eq. (10), E_i is the energy value of each efficient IMF component.

3.3 The Determination of Weights of IMF and the Acquisition of EMD Generalized Energy

In order to better characterize the affect of each IMF component, the energy weight coefficient is introduced to characterize the affect of different modal components on total energy:

$$E = \sum_{i=1}^n \lambda_i \cdot E_i \quad (11)$$

In Eq. (11), E is the generalized energy of vibration signal, and E_i expresses the energy value of the i th IMF; λ_i is the weight value of coefficient of the corresponding IMF, and needs to meet the condition of Eq. (12):

$$\sum_{i=1}^n \lambda_i = 1 \quad (12)$$

In this paper, the correlation coefficient between the IMF component and the rail vibration signal after de-noising is regarded as the weight coefficient of each IMF component.

The correlation coefficient of the signal $x(n)$ and $y(n)$ is as follows:

$$\rho_{xy} = \frac{\sum_{n=0}^{+\infty} x(n) \cdot y(n)}{\sqrt{\sum_{n=0}^{+\infty} x^2(n) \cdot \sum_{n=0}^{+\infty} y^2(n)}} \quad (13)$$

The correlation coefficient λ'_i between the IMF component after screening and the original vibration signal is obtained, and the ratio between λ'_i and global correlation coefficient is obtained:

$$\lambda_i = \frac{\lambda'_i}{\sum \lambda'_i} \quad (i = 1, 2, 3, \dots) \quad (14)$$

In Eq. (14), λ'_i is the correlation coefficient between IMF $_i$ and the original signal. λ'_i is regarded as the weight coefficient of the IMF $_i$ component, and the generalized energy is obtained by combining formulas in (10), (11), and (12):

$$E = \sum \left(\frac{\lambda'_i}{\sum \lambda'_i} \cdot \int_{-\infty}^{+\infty} |c_i(t)|^2 dt \right) \quad (i = 1, 2, 3, \dots) \quad (15)$$

3.4 The Determination of the Threshold Value of Security Domain

In the aspect of rail transit, some scholars put forward that the threshold value of security domain is the safety margin that rail vehicles can operate stably [8]. The safety margin in this paper is the minimum distance of generalized energy of rail vibration signal and the threshold value of security domain and it should fluctuate in a certain range. In this study, the state characteristic that is extracted from the signal of rail vibration accelerometer is viewed as the security-related variable, so Euclidean distance is used to calculate the threshold value of security domain in this paper. The formula to calculate the Euclidean Distance between N-dimensional vector $U_i(U_{i1}, U_{i2}, \dots, U_{in})$ and N-dimensional vector $U_j(U_{j1}, U_{j2}, \dots, U_{jn})$ is as follows:

$$D_o(U_i, U_j) = \sqrt{\sum_{k=1}^n (U_{ik} - U_{jk})^2} \tag{16}$$

4 The Experiment and Analysis

4.1 The Fault Diagnosis of Wheel

The data of rail vibration signal of wheel failure state assessment based on EMD generalized energy is obtained from system simulation. The system selects 500 groups of signals, and the frequency is 20 kHz. Each signal has 2000 data points, and the duration is 0.1 s.

In this paper, the method of wavelet threshold de-noising is used to reduce the noise of rail vibration signal, and the SNR that is calculated from the data result after de-noising is evaluated comprehensively. The result shows that the noise reduction effect is optimal by using the selection of sym8 wavelet basis, three-level decomposition, heuristic rule, and soft threshold. Figure 2 shows the rail vibration signal after de-noising. Most noise is reduced and the shock response signal of wheel to rail is well reserved, which provides data for the EMD decomposition.

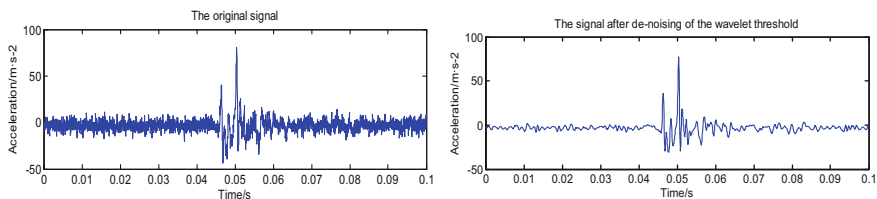


Fig. 2 Effect of rail vibration signal de-noising

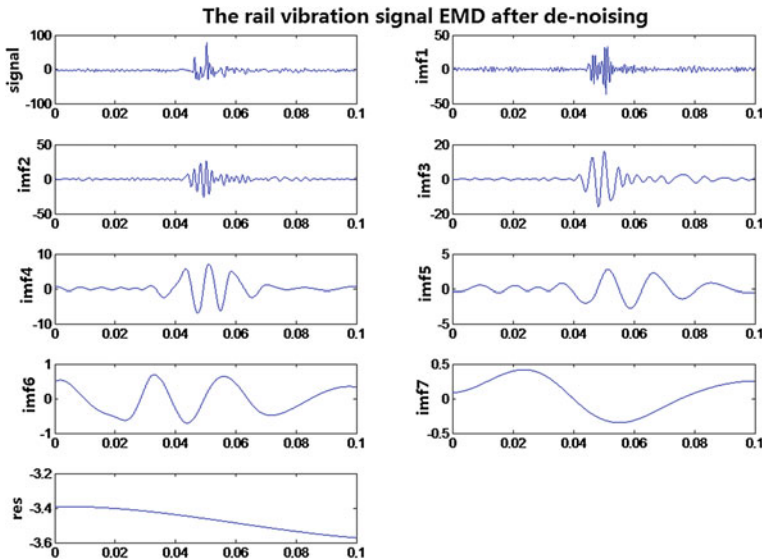


Fig. 3 Results of EMD decomposition of the de-noising rail signal

Figure 3 shows that the rail vibration signal after de-noising gets IMF1 ~ IMF7 and res through EMD. The signal is the rail vibration signal after original de-noising, IMF1 ~ IMF7 is 7 IMF components, and res is residual component.

After the EMD decomposition of the de-noising signal, the effective IMF component is extracted by the algorithm, and the value of energy is regarded as the cardinal number of the ultimate value of generalized energy. 7 IMF after the EMD decomposition is calculated in the time-domain kurtosis and the frequency divergence, and the results are shown in Table 1.

Viewed from Table 1, former five of IMF include most of the energy of the signal, reaching more than 90%, thus is regarded as the main IMF component of this rail vibration signal and the cardinal number to generalized energy calculation. According to the rule in Sect. 3.2, IMF1 ~ IMF5 are selected as the effective IMF components.

The selected five IMF are used to calculate the correlation coefficient, and the weight coefficient of each component is obtained according to Eq. (14). The results are in Table 2, and the value of generalized energy of this signal calculated by Eq. (15) is 2.8017×10^4 .

Table 1 Time-domain kurtosis and frequency divergence of each IMF component

Project	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
Time-domain kurtosis	25.03	18.01	12.18	6.124	3.612	1.681	1.881
Frequency divergence	9.264	6.458	3.210	2.354	1.868	0.523	0.482

Table 2 The weight coefficient of each component

Project	IMF1	IMF2	IMF3	IMF4	IMF5
The weight coefficient	0.2959	0.2752	0.2492	0.1658	0.0140

According to the above method of EMD generalized energy calculation, the generalized energy is obtained from the data of 500-segment simulated rail vibration. The results are in Fig. 4. The range of the generalized energy obtained from normal operation of wheel is between 2.1×10^4 and 3.0×10^4 . The average generalized energy value of 500 groups of data is 2.5246×10^4 . The difference between 500 sets of data and this average is calculated and viewed as the Euclidean Distance of safety margin. Viewed from Fig. 5, the maximum difference is the difference between the value of generalized energy of No. 277 and the average and is 0.5126×10^4 . So, the threshold value of security domain for the wheel fault state based on EMD generalized energy is 0.5126×10^4 .

4.2 The Verification of Simulation Data

50 groups of rail vibration signal of faulty wheels are obtained by the model simulation, and 50 groups of rail vibration signal of normal wheels are selected to verify the accuracy of the above threshold of security domain.

First, the generalized energy value of normal signal and fault signal are calculated. Then, the difference between each generalized energy value and the average value is obtained. The result of the normal signal is shown in Fig. 6, and the result of the fault signal is shown in Fig. 7.

When the state of the wheel is evaluated by the threshold of the safety domain constructed by the generalized energy value, a misjudgment may appear. As is seen from Fig. 6, there are three sets that are misjudged as fault wheel. The differences are respectively 5429, 6644, and 6748. But the boundary value of security domain is 5126, and the misjudgment rate is 6%; As is seen from Fig. 7, in the 50 sets of fault wheel data, there are five sets that are misjudged as normal wheel, and the differences are, respectively, 4931, 3315, 4028, 3219, 2880, and the misjudgment

Fig. 4 Energy value of vibration signals

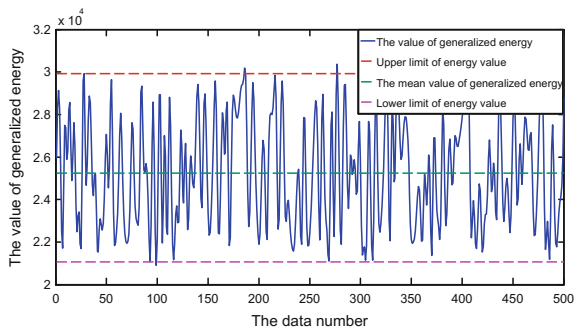


Fig. 5 Difference with the mean value

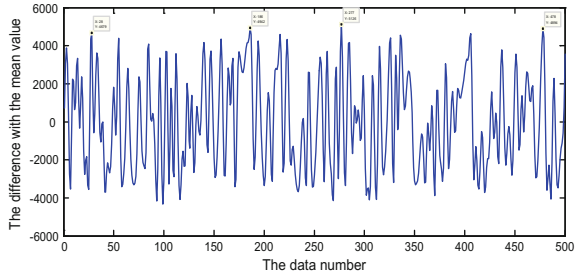


Fig. 6 Result of normal wheel

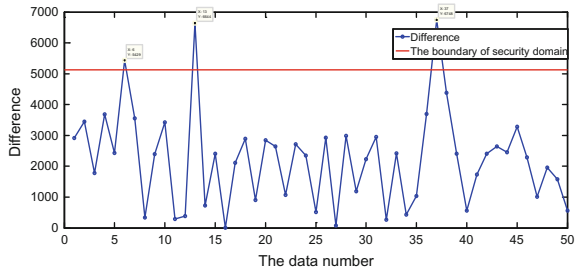
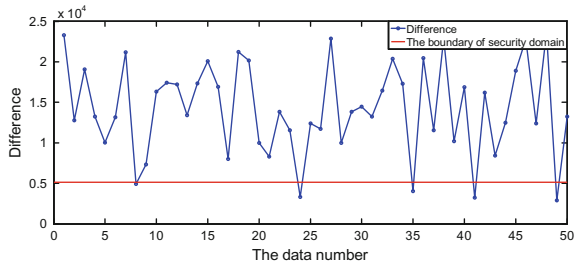


Fig. 7 Result of faulty wheel



rate is 10%. The reason of misjudgment might be the fact that wheel appears minor abrasions and leads to the obvious impact on the rail, so the generalized energy value does not exceed the normal range. When using the generalized energy single eigenvalue, the evaluation of wheel failure status is easy to achieve, but failure mode cannot be recognized, so this method can be applied to real-time monitoring of wheels, to issue warning when the generalized energy value of wheel is located at the edge of the security domain boundary or exceed the security domain boundary, and to provide maintenance decision-making reference for vehicle maintenance department.

5 Conclusion

In this paper, an evaluation method of the failure of train wheel based on EMD generalized energy is proposed. By processing the simulated vibration signal of rail, the energy weight sum of each component and the generalized energy value of this vibration signal are obtained. The boundary of security domain for wheel failure status is decided by the generalized energy value of a large number of normal signals, so it can be judged whether the generalized energy value of the rail vibration signal is within the safety domain and to distinguish faulty wheel from normal one. Finally, experimental verification is carried out through the rail vibration signal of simulated normal wheel and fault wheel, and the result shows that the method in this paper effectively distinguishes fault wheel and normal wheel. It verifies the accuracy and the effectiveness of the method in this paper and its great significance in the worksite application.

Acknowledgements This work is supported by National Key R&D Program of China under Grant (2017YFB1201201).

References

1. He G, Xing Z, Zuo C, Zhang Y (2015) Fault diagnosis method for rolling bearing of metro vehicle based on EMD and SVM. *Railway Comput Appl* 8:1–4 + 15 (in Chinese)
2. Belotti V, Crenna F, Michelini RC et al (2006) Wheel-flat diagnostic tool via wavelet transform. *Mech Syst Signal Process* 20(8):1953–1966
3. Wei Y, Wu Y, Chen E, Duan Z, Wang L (2015) Modal/spectrum dynamic response analyses of wheel/rail system under the condition of track irregularity based on ANSYS. *Railway Stand Design* 5:51–54 (in Chinese)
4. Asplund M, Famurewa S, Rantatalo M (2014) Condition monitoring and e-maintenance solution of railway wheels. *J Qual Maintenance Eng* 20(3):216–232
5. Belotti V, Crenna F, Michelini RC et al (2006) Wheel-flat diagnostic tool via wavelet transform. *Mech Syst Sig Process* 20(8):1953–1966
6. Lechowicz S, Hunt C (1999, May 3–5) Monitoring and managing wheel condition and loading. In: *Proceedings of the international symposium on transportation recorders*. Arlington, Virginia, pp 205–239
7. Zhou X, Wang X, Yang Y, Guo J, Wang P (2014) De-noising of high-speed turnout vibration signals based on wavelet threshold. *J Vib Shock* 33(23):200–206 (in Chinese)
8. Pu M, Liang S, Fang Z, Zhao W (2011) High-speed bogies nonlinearity and high-speed railway vehicle safety stability margin. *China Railway Sci* 32(3):86–92 (in Chinese)

Detection of Squats Length Based on Axle Box Vibration Analysis

Zhi Yang, Zongyi Xing and Jie Jiang

Abstract A method based on axle box vibration analysis is proposed to detect the length of squats. First, the dynamic model of wheel/rail is established for the research on relation between the frequency of track irregularity and that of axle box acceleration. Second, the basic principles of Adaptive Morphological Filtering (AMF) and Frequency Slice Wavelet Transform (FSWT) are introduced, and the slice function in FSWT is improved to advance the performance of FSWT to different signals. Finally, the measured axle box vibration signals from a metro are analyzed to estimate squats length. The results show that the method has a good engineering adaptability and feasibility.

Keywords Axle box · Vibration analysis · FSWT · AMF · Squats length

1 Introduction

During train operation, track irregularity is the primary cause of the abnormal vibration. Squats as one of the irregularities of short track waves can increase train vibration and decrease train service life, if it gets worse, rail abrasion and the safety of train operation would happen [1].

Experts of track traffic have already done lots of research in detecting track conditions. Molina et al. [2] have put forward a machine vision system to identify defects and symptomatic conditions. Lee et al. [3] have described a mixed filtering method including Kalman filtering, band-pass filtering, and adaptive filtering to estimate irregularities in railway tracks. Real [4] has used the methods of double

Z. Yang · Z. Xing (✉)

School of Automation, Nanjing University of Science and Technology,
Nanjing 210094, China
e-mail: xingzongyi@163.com

J. Jiang

School of Mechanical Engineering, Nanjing University of Science
and Technology, Nanjing 210094, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_3

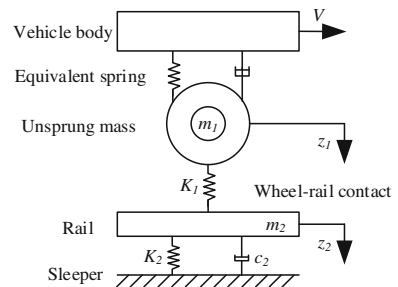
integral, high-pass filtering, and phase compensation filtering to eliminate the influence of interference signal, obtain vertical track geometry, and locate defects along the track. Yunpeng et al. [5] have established the finite element model of wheel/rail system based on ANSYS and made analysis on spectrum dynamic response of wheel/rail system under the condition of track irregularity. Jianming et al. [6] have used Frequency Slice Wavelet Transform (FSWT) to analyze the time-frequency characteristics of the leading and rear wheels caused by short track defects, and constructed the model of detecting short track defects to prove effectiveness of the method.

The above researches on the detection of track conditions remain at the qualitative level, and the characteristic parameters of the squats have not been studied. On the foundation of relevant research of track irregularity at home and abroad, a method based on Adaptive Morphological Filtering (AMF) and FSWT is proposed to detect the length of squats. Analyzing the vibration signal by filtering and frequency sliced wavelet transform, we can get fault characteristic frequency corresponding to maximum amplitude in the zoom spectrum graph. Then, according to train speed and fault characteristic frequency, the length of squats can be estimated.

2 The Dynamic Model of Wheel/Rail

The dynamic model of wheel/rail is a common tool to analyze mechanical characteristics of wheel pairs and track vibration under the influence of track irregularity. It includes two parts: vehicle model and wheel/rail model. In the practical analysis, the two parts are often simplified, only considering the high-frequency shock between wheel and rail and vibration caused by sprung [7]. The simplified parameter-lumped dynamical model of the train is shown in Fig. 1, where we can see that K_1 represents elastic contact stiffness between wheel and rail, and K_2 is vertical stiffness of the track. Also, m_1 is unsprung mass and m_2 is equivalent mass of the track. c_2 is vertical damping of the track. Besides, z_1 and z_2 denote, respectively, the displacement of the unsprung mass and the track relative to static equilibrium position.

Fig. 1 The simplified parameter-lumped model of the train



When the train running speed is v , the vertical dynamic equation of wheel/rail can be expressed as Eq. (1).

$$\begin{cases} m_1 \ddot{z}_1 + K_1(z_1 - z_2) = 0 \\ m_1 \ddot{z}_2 + c_2 \dot{z}_2 + K_2 z_2 - K_1(z_1 - z_2) = 0 \end{cases} \quad (1)$$

According to track irregularity η , the displacement of the rail relative to equilibrium position can be expressed as $z_0 + \eta$, so Eq. (1) can be transformed as Eq. (2).

$$\begin{cases} m_1 \ddot{z}_1 + K_1(z_1 - z_2 - \eta) = 0 \\ m_1 \ddot{z}_2 + c_2 \dot{z}_2 + K_2 z_2 - K_1(z_1 - z_2 - \eta) = 0 \end{cases} \quad (2)$$

Solving the differential equations with constant coefficients of Eq. (2), we can get

$$\begin{cases} \ddot{z}_1 = \sum_{i=1}^3 A_i e^{-p_i t} + B \cos(\omega t + \psi) \end{cases} \quad (3)$$

In Eq. (3), A_i , p_i and B are constants, and their values are influenced by parameters ω_1 , ω_2 , K_1 , and K_2 jointly. From Eq. (3), it can be seen that the first term on the right side of the equation is the frequency-independent attenuation and the second item is steady value of equation. That is to say, acceleration of the axle box and track irregularity have the same frequency. Thus, when the train runs at speed v on the track and wavelength of track irregularity is λ , the angular velocity of track irregularity can be expressed as $\omega = 2\pi v/\lambda$, and impact frequency of track irregularity on the train can be represented as $f = v/\lambda$. Besides, specific track defect corresponds to a specific particular characteristic frequency, and if the train speed and the irregularity wavelength are known, the characteristic frequency can be obtained. In other words, if train running speed and fault characteristic frequency are known, track irregularity wavelength can be obtained.

3 Detection Principle of the Squats Length

In this section, we mainly talk about detection principle of the squats length including two parts: AMF and FSWT.

3.1 Adaptive Morphological Filtering

Traditional morphological filters mostly adopt a single structuring element to process the signals, but filtering effect depends on the selection of structuring elements. In order to get a better filtering effect, the signal should be known well

before it is processed, but the actual situation is that engineering signals usually contain multiple types of noise. It is hard for us to select proper structuring elements based on the prior knowledge. Therefore, according to different types of noise, it is necessary to select different structuring elements and make corresponding combinations.

If multiscale morphological structuring elements are combined, the generalized morphological filter can be constructed as

$$y(n) = \frac{1}{2} \cdot [F_{Goc}(f(n)) + F_{Gco}(f(n))] \quad (4)$$

In Eq. (4),

$$F_{Goc}(f(n)) = (f \circ g_1 \bullet g_2) \quad (5)$$

$$F_{Gco}(f(n)) = (f \bullet g_1 \circ g_2) \quad (6)$$

In Eq. (5) and (6), \circ and \bullet , respectively, represent opening and closing operation.

In this section, intensity coefficient of the characteristic frequency is used to quantify the processing ability of the vertical vibration signal of the axle box under the influence of different structuring elements. The expression of the characteristic frequency strength coefficient is shown as in Eq. (7), which is defined as the ratio of the sum of octaves amplitudes to the sum of frequency amplitudes.

$$Cf = \sum_{i=1}^3 FC_i / \sum_{j=1}^{N-1} F_j \quad (7)$$

In Eq. (7), $FC_i (i = 1, 2, 3)$ represents octaves amplitude of characteristic frequency in the frequency spectrum, and $F_j (j = 1, 2, \dots, N - 1)$ represents frequency amplitude.

The characteristic frequency strength coefficient has different amplitude values under the influence of different structuring elements. According to different values, we can select optimal structuring element so that optimal filter can be constructed.

3.2 Adaptive Morphological Filtering

Based on the advantages of short-time Fourier transform and wavelet transform, a new time-frequency analysis method called FSWT is proposed by Yan [8]. For any signal $f(t) \in L^2(R)$, FSWT can be defined by Eq. (8).

$$W(t, \omega, \lambda, \sigma) = \frac{\lambda}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(u) \hat{p}^* \left(\frac{u - \omega}{\sigma} \right) e^{iut} du \tag{8}$$

In Eq. (8), σ is the scale factor and λ is energy coefficient. In addition, $\hat{f}(u)$ is the Fourier transform of the signal $f(x)$, and $\hat{p}(\omega)$ representing the frequency slice function is the Fourier transform of the signal $p(t)$.

The slice function shapes of FSWT usually depend on the signals to be analyzed, and these signals need to select different slice functions according to the characteristics of the signals, which limit the adaptability of FSWT. Biswal [9] has made improvement on the frequency slice function $\hat{p}(\omega) = e^{-0.5\omega^2}$ to make it in a greater adaptability, and applied it to detection of power signal disturbances. But in this paper, the improvement will be applied to the analysis of vertical vibration signals of the axle box.

Thus, the improvement of frequency slice function is shown in Eq. (9).

$$p(\omega) = e^{\frac{-2\lambda^2 \pi^2 \omega^2}{a + b\sqrt{\omega}}} \tag{9}$$

FSWT well solves the problem that traditional wavelet-reconstructed signal must rely on wavelet basis, in addition, its inverse transform has no direct connection to slice function, so fast Fourier transform (FFT) can be used to reconstruct signal separation.

In the time-frequency transform interval of the target signal, target interval signals can be separated and reconstructed by selecting time-frequency slice interval $(t_1, t_2, \omega_1, \omega_2)$.

$$f_s(t) = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} \int_{t_1}^{t_2} W_f(\tau, \omega, \lambda, \sigma) e^{i\omega(t-\tau)} d\tau d\omega \tag{10}$$

4 Experimental Analysis

In this paper, the measured vertical vibration signals of the axle box based on a car of tape A is taken as an example, and the signals are processed by AMF and FSWT in order to verify good engineering applicability of the method proposed. The field experiment is carried out on a re-profiled car so that wheel irregularity has little influence on the detection of squats length. The sampling frequency of accelerometer of A-Type car is 20 kHz, and the running speed of the train is 10 m/s.

Figure 2 shows the waveform graph of vertical vibration signals of the axle box in 1 s. It can be seen that the vertical vibration signals of the axle box at 0.190, 0.922 and 0.720 s are greatly affected by noise, and the signals at 0.6 and 0.7 s have

obvious vibration shocks, in addition, the vibration shocks at 0.7 s are relatively complex. In order to extract fault characteristic information, vibration signals should be filtered to reduce the influence of interference signal on the analysis results.

According to structure principle of adaptive morphological filters, structuring element parameters of measured vibration signals should be calculated firstly. The length scale of structuring elements L is $[6, 20]$, and the height scale of structuring elements H is $[0, 0.6826]$. On this basis, morphological opening operation, closing operation, open-closing, and close-opening operation are operated respectively on the analysis of vertical vibration signals of the axle box. The optimal structuring element under different morphological operation is shown in Table 1.

Figure 3 shows the frequency intensity coefficients under different operation and structuring elements. It can be seen that the frequency intensity coefficients of morphological open-closing operation and close-opening operation fluctuate between 0.05 and 0.12, while the frequency intensity coefficients of opening operation and closing operation are not more than 0.04, which indicates that open-closing and close-opening operation have a better filtering effect. Because individual open-closing operation increase negative impulsive noise, and individual close-opening cannot filter out positive impulsive noise completely, in this paper, upper triangular structuring element with length 16 and lower semicircular structuring element with length 11 are combined to construct a generalized morphological filter to process original signal.

Figure 4 shows the measured vibration signal of the axle box after AMF. It can be seen that the signals are much smoother, and abnormal vibration signals near 0.190 and 0.922 s can be eliminated effectively. Moreover, vibration impact becomes much clearer even near 0.70 s which is superior to the Butterworth filtering.

Fig. 2 The measured vibration signals of the axle box

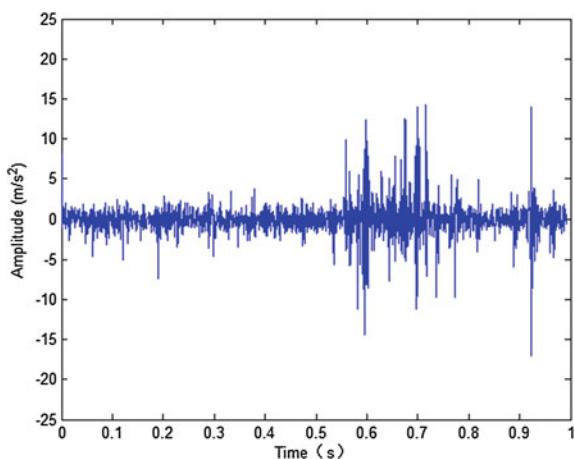


Table 1 Selection of optimal structuring elements

Morphological operators	Sensitive structuring elements
Opening operation	The length of line structuring element is 11
Closing operation	The length of upper triangular structuring element is 11
Open-closing operation	The length of upper triangular structuring element is 16
Close-opening operation	The length of lower semicircular structuring element is 11

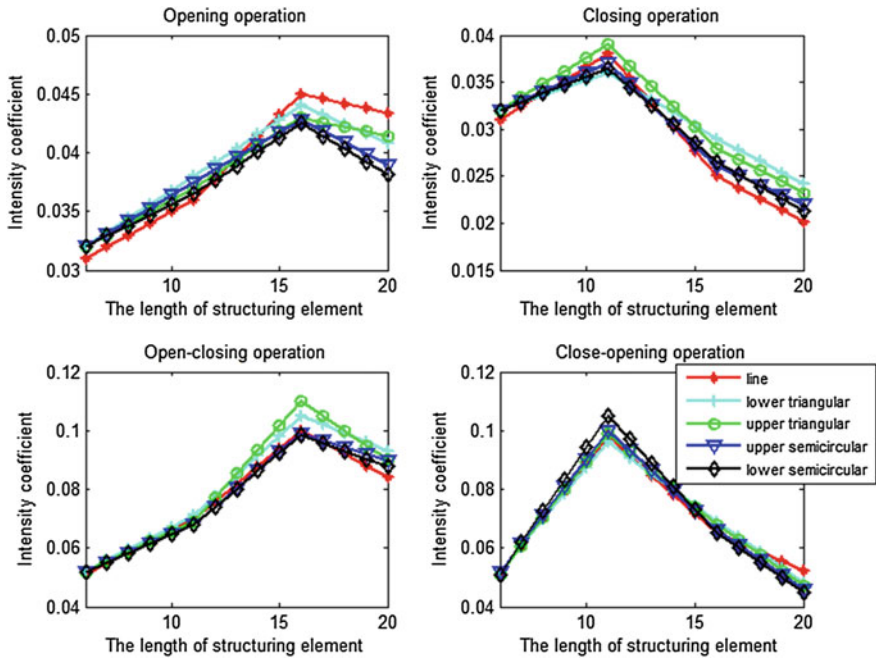


Fig. 3 The frequency intensity coefficients under different operations

Figure 5a shows analysis results of vibration signals of the axle box by using $\hat{p}(\omega) = e^{-\omega^2/2}$ as a slice function, and Fig. 5b shows analysis results by using $\hat{p}(\omega) = e^{-200\pi^2\omega^2/8+9\sqrt{|\omega|}}$ as a slice function. From Fig. 5a, it can be seen that there is an obvious target area at 0.6 s, while at 0.7 s, it is difficult to determine fault characteristic frequency, which has a contradictory with the obvious vibration impact at 0.7 s in the Fig. 4. From Fig. 5b, it can be seen that there are two maximum amplitude target areas at 0.6 and 0.7 s. Compared with the two time-frequency analysis results, it can be seen that the improved slice function has better time-frequency resolution and has a greater advantage in processing non-stationary vibration signals.

The two maximum amplitude target areas in Fig. 6 are [0.56 s, 0.63 s, 262 Hz, 332 Hz] and [0.65 s, 0.72 s, 313 Hz, 369 Hz]. According to Eq. (8), the two target

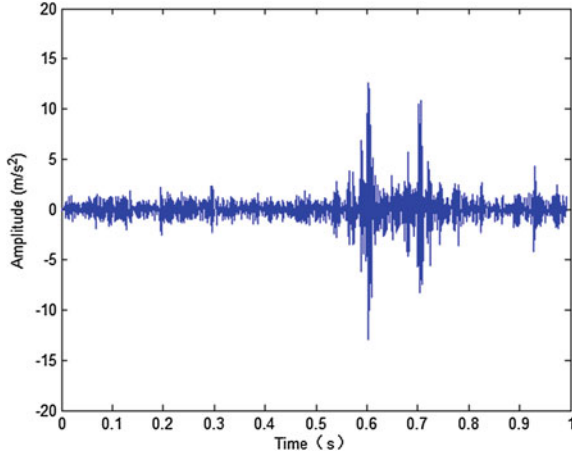


Fig. 4 The measured vibration signal of the axle box after AMF

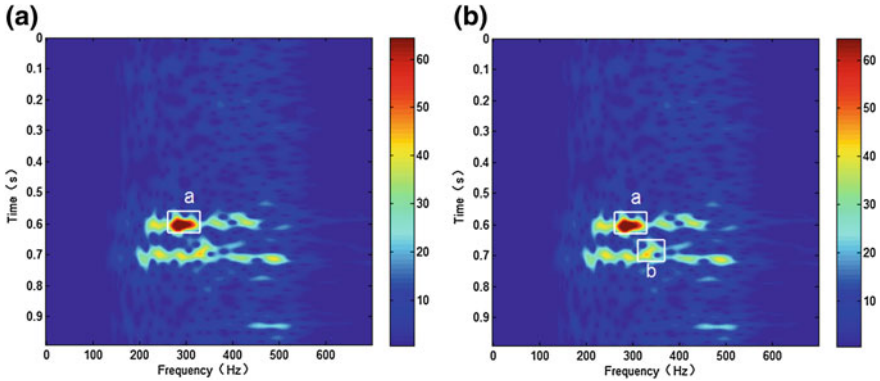


Fig. 5 Time-frequency graph of the vibration signals of the axle box

areas a and b can be separated. Figure 6 shows reconstructing signal of the target area. It can be seen that signal reconstruction has a better effect by FSWT.

The zoom frequency spectra of target areas a and b is shown in Fig. 7a, b. The peak amplitude of the target area a corresponds to a frequency of 281 Hz, and the peak amplitude of the target area b corresponds to a frequency of 338.5 Hz. So, we can conclude that there are train's vibration signals caused by track in 281 and 338.5 Hz. When train speed is 10 m/s, the squats length can be calculated as 35.4 and 29.5 mm respectively based on derived mathematical relationship. According to the squats length, we can get that the squats are medium, and this is consistent with squats on site. All these indicate that the method has the reliability and feasibility of the project.

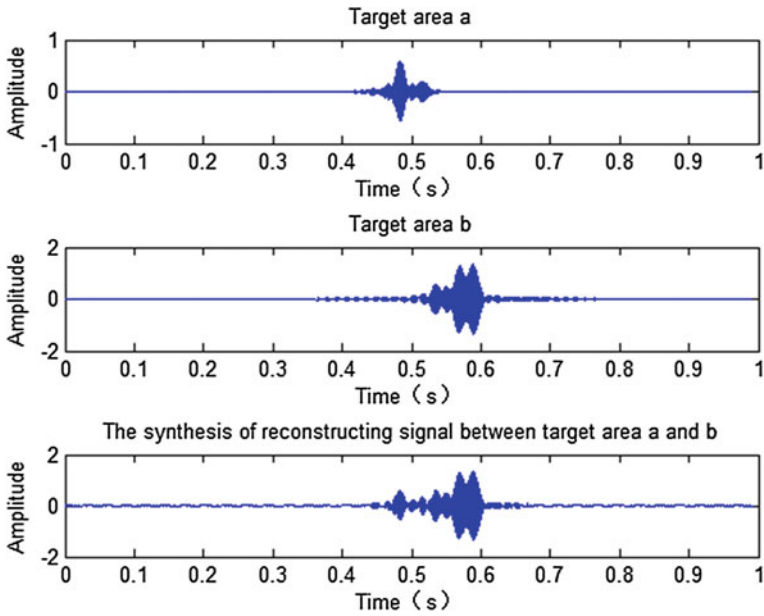


Fig. 6 Reconstructing signal of target areas

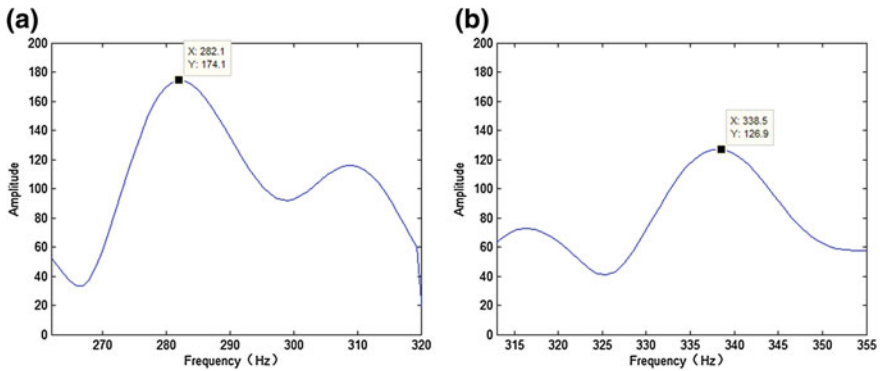


Fig. 7 The zoom frequency spectra of the target areas

5 Conclusion

In this paper, a method based on AMF and FSWT to detect the length of squats is proposed. First, based on the measured data of axle box, adaptive morphological combination filter can be constructed to eliminate the effects of noise and other interfering signals. Second, improved FSWT is applied to analyze vertical vibration

signals of the axle box. Target area is selected for zooming analysis to obtain more accurate fault characteristic information. Finally, the length of squats is calculated by using derived mathematical relationship. Comparing the actual site, the reliability and feasibility of the project can be verified.

Acknowledgements This work is supported by National Key R&D Program of China under Grant (2016YFB1200402).

References

1. Jianmin G, Wanming Z (2014) Dynamic effect and safety limits of rail weld irregularity on high-speed railways. *Sci Sin Tech* 44:697–706. <https://doi.org/10.1360/n092014-00081> (in Chinese)
2. Molina LF, Resendiz E, Edwards JR et al (2011) Condition monitoring of railway turnouts and other track components using machine vision. In: *Transportation research board 90th annual meeting* (11–1442)
3. Lee JS, Choi S, Kim SS et al (2012) A mixed filtering approach for track condition monitoring using accelerometers on the axle box and bogie. *Instrum Meas IEEE Trans Instrum Meas* 61(3):749–758
4. Real JI, Montalbán L, Real T et al (2012) Development of a system to obtain vertical track geometry measuring axle-box accelerations from in-service trains. *J Vibroeng* 14(2)
5. Yunpeng W, Yaping W, E C et al (2015) Modal/spectrum dynamic response analyses of wheel/rail system under the condition of track irregularity based on ANSYS. *Railway Standard Design* 5:51–54 (in Chinese)
6. Jianming D, Jianhui D, Han W et al (2014) Dynamic detection of short track defects comparing position of impact characteristics. *J Vib Shock* 33(6):113–117 (in Chinese)
7. Jian Z, Xuesong J et al (2011) Solution methods of rail model in vehicle-track coupling dynamics. *J Traffic Transp Eng* 30(10):104–108 (in Chinese)
8. Yan Z, Miyamoto A, Jiang Z et al (2010) An overall theoretical description of frequency slice wavelet transform. *Mech Syst Signal Process* 24(2):491–507
9. Biswal B, Mishra S (2014) Power signal disturbance identification and classification using a modified frequency slice wavelet transform. *Gener Transm Distrib IET* 8(2):353–362

Profile Calibration of Dynamically Measuring Rail Wear Using LIS

Chao Wang, Yanfu Li, Dawei Liu and Hongli Liu

Abstract Dynamically inspecting steel rail wear by the laser image system suffers from vibrations on the camera and the laser emitter, which leads to varying extrinsic camera parameters and distorted rail profile, respectively. In the paper, a novel profile calibrating method is presented to address these issues. This method involves a self-calibration of the extrinsic camera parameters, which is based on the coordinate sets projection constructed from the distorted profile to the standard one on quadruple projecting primitives including rail jaw, railhead, rail waist, and foot intersection. To remove the effect of random vibrations, we calibrate distorted profile by projecting it onto the plane $x - y$ of the world coordinate frame. We compared our procedure with other established methods. The results show its effectiveness and superiority for dynamically measuring rail wear.

Keywords Profile calibration · Coordinate sets projection · Laser image system
Rail wear

1 Introduction

To ensure transportation security, the track rail needs to be inspected periodically. In the process, the track wear is one of the key objects, which not only represents the levels of rail defect directly but also offers a dependable reference for rail maintenance [1].

C. Wang (✉) · Y. Li · D. Liu · H. Liu
College of Electrical and Information Engineering, Hunan University,
Changsha, China
e-mail: chao_wang@hnu.edu.cn

C. Wang · Y. Li · D. Liu · H. Liu
College of Electrical Engineering, University of South China, Hengyang, China

C. Wang · Y. Li · D. Liu · H. Liu
Department of Computer Science and Software Engineering,
Xi'an Jiaotong-Liverpool University, Suzhou, China

Laser image system (LIS) is the main technology of measuring the rail profile. As can be seen in Fig. 1, such systems are usually composed of a laser emitter and a camera. The laser emitter is utilized to highlight the rail profile, and camera captures interesting images. According to the intrinsic and extrinsic camera parameters, the measuring profile represented by the spatial coordinates of the laser stripe is calculated, and is calibrated with the standard one to compute the wear [2].

In order to dynamically measure rail wear, LIS is usually assembled under the rail inspection and maintenance trains. Since the movement of the train, the track un-flatness could lead to the multiple degrees of freedom vibration on the laser emitter and camera, i.e., swaying, stretching, bouncing, rolling, pitching, and heading [3]. Among all vibrations on line laser projector, the first four do not have any effect on the verticality between the laser plane and the rail longitudinal axis. The measured profile is correct and usable. However, the pitching and heading vibrations make the laser plane fastigiated. In Fig. 2, the profiles are stretched along vertical and horizontal axis due to the pitching and heading, respectively. The stretch transformation would cause the distorted profile.

On the other hand, all vibrations other than stretching on camera would lead to varying in relative attitude between the camera and rail. In camera model, the relative attitude between camera and object can be modeled as extrinsic parameters of camera [4]. The variations in relative attitude mean different extrinsic camera

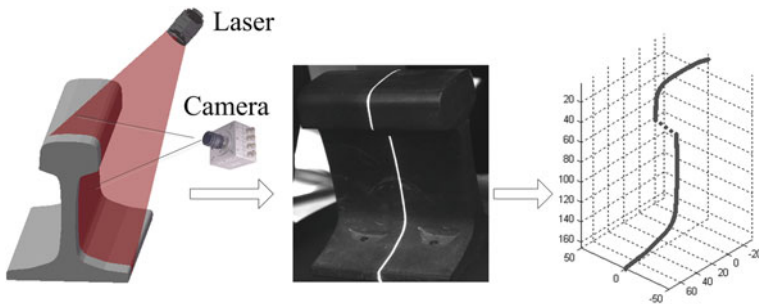


Fig. 1 The laser image system used for rail wear measurement

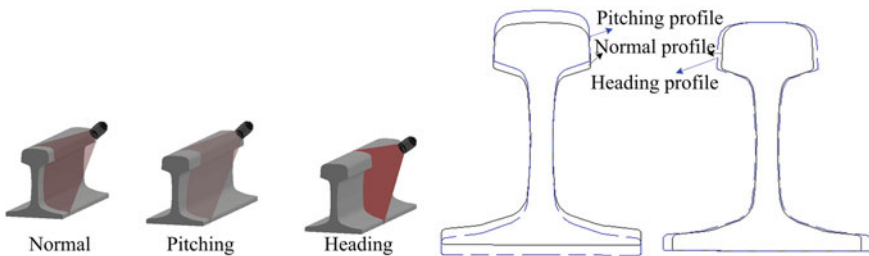


Fig. 2 The distorted profile caused by pitching and heading on the laser emitter

parameters, which are hardly calibrated during the dynamic measurement with the traditional calibration methods [5]. So, it is difficult to reconstruct 3D space coordinates of rail profiles from 2D pixel coordinates correctly.

To address the issues above, this study proposes a profile calibrating method for the dynamic measurement of rail wear.

2 Calibration of Distorted Profiles

First, three Euler angles are employed for the representation of the rotation matrix. Then, iterative extrinsic parameters self-calibration based on the coordinate sets projection are implemented. By projecting the distorted profile on the plane $x - y$ of the world coordinate system, we can calibrate it easily.

2.1 Rotation Matrix Denoted by Euler Angles

Assuming that α , β and θ are the rotation angles of camera in the axes x , y and z respectively, their rotation matrix of each axis can be denoted as

$$\begin{aligned} \mathbf{R}_x(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} & \mathbf{R}_y(\beta) &= \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \\ & & \mathbf{R}_z(\theta) &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (1)$$

The final rotation matrix can be denoted as

$$\mathbf{R} = \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\theta). \quad (2)$$

Given a profile point $p_{ki} = (u_{ki}, v_{ki})^T$ detected from the image i and its corresponding one $P_k = (x_{kw}, y_{kw}, z_{kw})^T$ from 3D space, we have

$$P_k = \mathbf{R}_i^{-1}\mathbf{A}^{-1}\hat{p}_{ki} - \mathbf{R}_i^{-1}\mathbf{t}_i, \quad (3)$$

where \hat{p}_{ki} is a homogeneous coordinate of p_{ki} , \mathbf{A} is intrinsic parameter matrix, \mathbf{R}_i and \mathbf{t}_i are rotation and translation matrix for image i , respectively. \mathbf{R}_i is determined by the Euler angles α_i , β_i and θ_i directly.

2.2 Coordinate Sets Projection

In order to estimate the extrinsic camera parameters, i.e., α_i , β_i , θ_i and t_i , we proposed a similarity quantization called coordinate sets projection. This method employs four rail regions, covering the rail jaw, railhead, rail waist, and foot intersection to act as the projecting primitives. Due to the invariable attitude of the standard profile which is vertical to the rail longitudinal axis, we project these primitives from the measured profile onto the standard one, as shown in Fig. 3.

In the implement, given a point on the railhead $A'B'$ of the measured profile, we first project it onto the standard one and determine its response point by the cubic spline interpolation. Then, we calculate the projecting error as follows

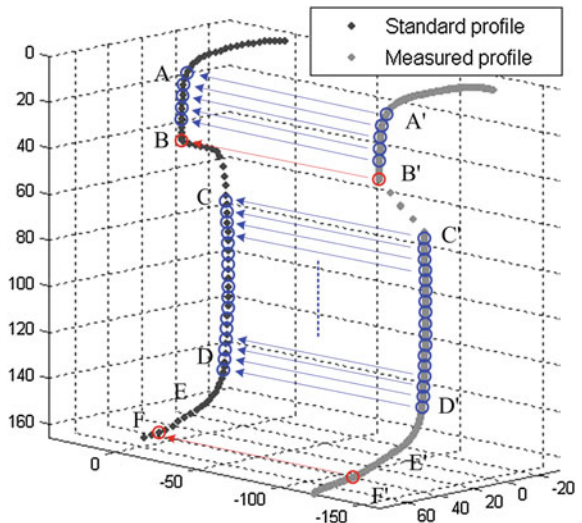
$$\text{fitness}_{AB} = \frac{1}{N_h} \sum_{j=1}^{N_h} \|x_{h_{pj}} - x_{h_{rj}}\|^2, \quad (4)$$

where N_h is the quantity of projecting points on $A'B'$, $x_{h_{pj}}$ acquired from (3) is the horizontal coordinate of the j th projecting point, and $x_{h_{rj}}$ is the coordinate of the response one. In the same way, the projecting error of the rail waist could be denoted as

$$\text{fitness}_{CD} = \frac{1}{N_w} \sum_{j=1}^{N_w} \|x_{w_{pj}} - x_{w_{rj}}\|^2, \quad (5)$$

where N_w is the quantity of projecting points on $C'D'$, $x_{w_{pj}}$ acquired from (3) is the horizontal coordinate of the j th projecting one, and $x_{w_{rj}}$ is the coordinate of the response one.

Fig. 3 The process of coordinate sets projection



The locations of the rail jaw B and rail foot intersection F are known and correct, the interpolation is therefore unnecessary. Their projecting errors could be represented as

$$\text{fitness}_B = \|B_p - B_s\|^2 \quad \text{fitness}_F = \|F_p - F_s\|^2, \quad (6)$$

where B_p and F_p are mapping point of the rail jaw B and the foot intersection F from the measured profile, respectively, B_s and F_s are from the standard profile.

If extrinsic parameters are inaccurate, all of the projecting errors are large. So, we can assign consistent weight coefficients for the quadruple primitives. The global projecting error will be represented as

$$\text{fitness} = \omega_{AB}\text{fitness}_{AB} + \omega_B\text{fitness}_B + \omega_F\text{fitness}_F + \omega_{CD}\text{fitness}_{CD}. \quad (7)$$

2.3 Minimization of the Global Projecting Error

In our procedure, the position vector $\mathbf{X}_n = [\alpha_n, \beta_n, \theta_n, t_{xn}, t_{yn}, t_{zn}]^T$ is established for each particle. The global projecting error described in (7) is served as the fitness. For each iteration, the position vector compares its fitness and records the best personal and swarm positions expressed as $\mathbf{P}_n = [\alpha_{pn}, \beta_{pn}, \theta_{pn}, t_{xpn}, t_{ypn}, t_{zpn}]^T$ and $\mathbf{G} = [\alpha_g, \beta_g, \theta_g, t_{xg}, t_{yg}, t_{zg}]^T$, respectively. On basis of the \mathbf{P}_n and \mathbf{G} , the velocity and position are updated iteratively for the optimization of the swarm. These procedures could be expressed as

$$\begin{aligned} \mathbf{V}_n^{k+1} &= \omega \mathbf{V}_n^k + c_1 r_1 (\mathbf{P}_n^k - \mathbf{X}_n^k) + c_2 r_2 (\mathbf{G}^k - \mathbf{X}_n^k) \\ \mathbf{X}_n^{k+1} &= \mathbf{X}_n^k + \mathbf{V}_n^{k+1}, \end{aligned} \quad (8)$$

where ω is the inertia weight. c_1 and c_2 are the accelerating factor. r_1 and r_2 are random digits between $[0, 1]$. k is the current number of iteration. When the iteration reaches the maximum number, the circulation stops and outputs the optimal position vector.

2.4 Rectifying of Distortions

The optimal position vector is utilized as the final extrinsic parameters to reconstruct 3D coordinates of the distorted profile. In addition, all the sections of the distorted profile are projected onto the plane $x - y$ of world coordinate system. Figure 4 exhibits a rectifying example. From the magnifying version of railheads, it is clear that the pitching and heading vibrations on the laser emitter cause the vertical and horizontal stretch of the distorted profile. The final rectifying profile is given in Fig. 4b, which presents obvious revivification of these stretches.

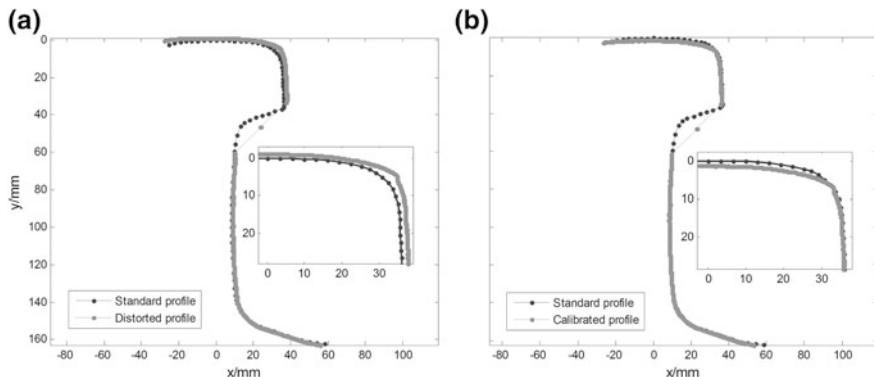


Fig. 4 An example of the rectifying distorted profile. **a** Comparison between the distorted profile and standard one. **b** Comparison between the recovered profile and standard one

3 Performance Evaluation

To evaluate the performance of the proposed calibration in Sect. 2, we used the classical calibration proposed by Zhang [6], self-calibration based on genetic algorithm (GA) [7, 8] and differential evolution particle swarm optimization (DEPSO) [9] for comparison.

We arbitrarily moved and rotated the camera to imitate swaying, bouncing, rolling, pitching, heading, and their composition, respectively. Under each camera perspective, we repeatedly took 10 images of the distorted profile. One of them is shown in Fig. 5. Each perspective image is corresponding to one group of extrinsic parameters. These parameters were estimated using the methods based on Zhang, GA, DEPSO, and ours, respectively, and their statistics were reported in Table 1.

As shown in Table 1, the results obtained by our method are also closest to those from Zhang. This could be explained as follows. Since the variation of reflection property of rail surfaces and surrounding light, the images of the same scene may be discrepant, especially in intensity, texture, and acutance. This would give rise to the unreliable correspondences which are key primitives of the methods proposed in literatures [7–9]. They therefore suffer from the lower precision, although these

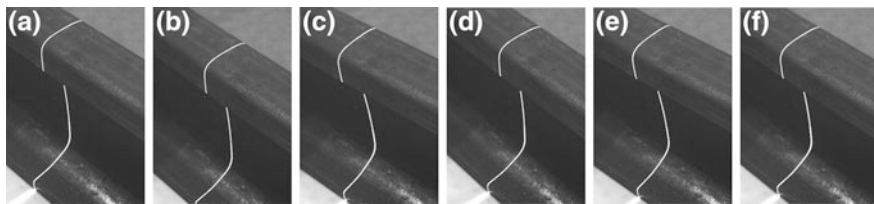


Fig. 5 Images of the normal profile under different camera perspectives. **a** Swaying. **b** Bouncing. **c** Rolling. **d** Pitching. **e** Heading. **f** Their composition

Table 1 Extrinsic parameters for distorted profiles obtained by different methods

	Method	α (rad)	β (rad)	θ (rad)	t_x (mm)	t_y (mm)	t_z (mm)
a	Zhang	0.771	0.775	0.694	-352.44	351.14	281.78
	GA	0.740	0.788	0.678	-352.11	351.08	281.63
	DEPSO	0.795	0.813	0.674	-352.28	351.29	281.56
	Ours	0.770	0.776	0.686	-352.47	351.10	281.75
b	Zhang	0.765	0.749	0.659	-351.86	356.01	286.97
	GA	0.756	0.740	0.630	-352.22	356.24	286.80
	DEPSO	0.741	0.743	0.684	-351.65	356.31	286.58
	Ours	0.770	0.741	0.668	-351.85	356.02	286.94
c	Zhang	0.712	0.815	0.673	-351.54	351.37	287.51
	GA	0.713	0.842	0.635	-351.76	351.49	287.38
	DEPSO	0.695	0.835	0.685	-351.40	351.10	287.24
	Ours	0.721	0.811	0.682	-351.49	351.35	287.49
d	Zhang	0.769	0.805	0.644	-352.04	352.54	287.51
	GA	0.803	0.813	0.676	-352.04	352.24	287.70
	DEPSO	0.744	0.780	0.626	-351.86	352.84	287.16
	Ours	0.776	0.802	0.647	-352.04	352.54	287.48
e	Zhang	0.807	0.785	0.647	-351.59	351.08	287.54
	GA	0.769	0.773	0.623	-351.73	350.96	286.70
	DEPSO	0.821	0.799	0.621	-351.77	350.06	287.15
	Ours	0.815	0.792	0.654	-351.53	351.16	287.62
f	Zhang	0.760	0.804	0.668	-352.01	355.34	283.05
	GA	0.769	0.777	0.644	-351.66	355.52	283.19
	DEPSO	0.795	0.812	0.666	-351.90	355.13	282.90
	Ours	0.762	0.801	0.665	-352.06	355.25	282.96

methods employ different search algorithms to enhance the accuracy of convergence. On the other hand, coordinate sets projecting, one of the main ideas of our work, utilizes the interpolation operation to ensure the true match, or rather the interpolation operation can reconstruct the counterparts which may be inexistent in the standard rail profile.

To further assess the validity of our procedure, we calculated the wears of 10 distorted profiles, and compared them with the truth manually acquired by the rail wear gauge and those calculated by Sun [10]. Table 2 lists the results.

It is clear that the errors between the wear values calculated by Sun and the truth are large relatively. One interpretation may be that Sun ignores the vibrations on camera which would cause the unreliable extrinsic parameters. Although he projects the profiles onto an auxiliary plane which is vertical to the rail longitudinal axis, the rectifying is still dubious. However, these influences, caused not only by the vibrations on the line laser projector but also on the camera, are taken into account fully by our method. So, the wear acquired by our proposed method nearly

Table 2 Results of wear measurement test

True wears (mm)			Results obtained by Sun (mm)			Results obtained by us (mm)			
	Vertical wear	Horizontal wear	Vertical wear	Error	Horizontal wear	Vertical wear	Error	Horizontal wear	Error
1	1.25	-0.40	1.10	0.15	-0.57	1.21	0.04	-0.45	0.05
2	1.23	-0.35	1.01	0.22	-0.52	1.12	0.11	-0.44	0.09
3	1.29	-0.43	1.13	0.16	-0.48	1.19	0.1	-0.47	0.04
4	1.35	-0.32	1.16	0.19	-0.51	1.20	0.15	-0.42	0.1
5	1.21	-0.34	1.08	0.13	-0.55	1.13	0.08	-0.41	0.07
6	1.18	-0.36	1.05	0.13	-0.54	1.07	0.11	-0.46	0.1
7	1.24	-0.38	1.11	0.13	-0.59	1.13	0.11	-0.46	0.08
8	1.34	-0.41	1.22	0.12	-0.56	1.20	0.14	-0.51	0.1
9	1.16	-0.43	1.04	0.12	-0.51	1.08	0.08	-0.49	0.06
10	1.28	-0.30	1.11	0.17	-0.44	1.14	0.14	-0.38	0.08
RMSE (mm)			0.1446	-	0.1691	0.1012	-	0.0791	-
MAPE (%)			11.12	-	44.53	7.802	-	20.76	-

Bold entries denote optimal results

approach to the truth. Moreover, The RMSEs decrease to 0.1012 and 0.0791 mm, respectively. The MAPEs decrease to 7.802 and 20.76%, respectively. These results exhibit the validity and superiority of our method.

4 Conclusion

To improve the accuracy of the rail wear measurement, we focused on the profile calibration under the vibrations both on the laser emitter and on the camera in this paper. The distorted profile is calibrated via the coordinate sets projection which is the basis for the self-calibrations of the extrinsic camera parameters. We performed a set of experiments, including comparison of profile calibration against the state of the art and validity test of the wear measurement, to assess our method. Experimental results exhibit that the proposed method is superior to the related classic ones, owning fine precision of profile calibration and wear measurement, even under the random vibrations on the laser emitter and camera.

References

1. Attivissimo F, Danese A, Giaquinto N, Sforza P (2007) A railway measurement system to evaluate the wheel–rail interaction quality. *IEEE Trans Instrum Meas* 56(5):1583–1589
2. Molleda J, Usamentiaga R, Millara ÁF, García DF, Manso P, Suárez CM, García I (2016) A profile measurement system for rail quality assessment during manufacturing. *IEEE Trans Ind Appl* 52(3):2684–2692
3. Liu H, Li Y, Ma Z, Wang C (2017) Recognition and calibration of rail profile under affine-distortion-based point set mapping. *IEEE Trans Instrum Meas* 66(1):131–140
4. Ramalingam S, Sturm P (2016) A unifying theory for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 1–11
5. Long C, Zhu J, Yi W (2016) Portable visual metrology without traditional self-calibration measurement model. *Measurement* 90:424–437
6. Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 22(11):1330–1334
7. Merras M, El Akkad N, Saaidi A, Nazih AG, Satori K (2015) Camera self calibration with varying parameters by an unknown three dimensional scene using the improved genetic algorithm. *3D Res* 6(1):1–14
8. El Akkad N, El Hazzat S, Saaidi A, Satori K (2016) Reconstruction of 3D scenes by camera self-calibration and using genetic algorithms. *3D Res* 7(1):1–17
9. Deng L, Lu G, Shao Y, Fei M, Hu H (2016) A novel camera calibration technique based on differential evolution particle swarm optimization algorithm. *Neurocomputing* 174:456–465
10. Sun J, Liu Z, Zhao Y, Liu Q, Zhang G (2013) Motion deviation rectifying method of dynamically measuring rail wear based on multi-line structured-light vision. *Opt Laser Technol* 50:25–32

Design and Implementation of Online Monitoring System of Contact System Tensioner

Wanjiao Han, Mingguang Liu and Hui Yu

Abstract Tensioner is an important equipment of contact system. The real-time monitoring is carried out to running state of contact system, so as to ensure the railway system operating safely and stably. First, this paper analyzes the necessity of real-time online monitoring and summarizes the effect of environment on contact system. Second, the selection of monitoring information is introduced and monitoring terminal is designed based on the requirement of transmission and monitoring information. Third, the main function and implementation method of computer monitoring software are introduced. Finally, through installation and debugging, monitoring system can operate safely and stably and achieve alarm function so that overhaul can be guided.

Keywords Contact system · Tensioner · Environment · Real-time online monitoring

1 Introduction

1.1 Research Background

Contact system is a special form of transmission line, which sets up over the rail lines and supplies electricity to train [1]. Contact system is exposed to outdoor environment throughout the year and there is no spare. Once the malfunction of contact system occurs, it will seriously affect the passenger safety, railway transportation and production, etc. Therefore, it is very necessary to know the condition of contact system timely.

Tensioner is an important equipment of contact system and is installed on both ends of wire of anchoring section. Under the action of gravity of tensioning weight, tensioner can automatically adjust tension of wire and ensure wire sag to meet technical requirements. Thus, stability and flexibility of catenary suspension can be

W. Han (✉) · M. Liu · H. Yu
Beijing Jiaotong University, Beijing, China
e-mail: 15121412@bjtu.edu.cn

improved and operating quality of contact system can also be improved [2]. Real-time monitoring tensioner and meteorological environment can evaluate the working state of contact system. When malfunction occurs, overhaul measures can be taken in time, making contact system be able to restore run quickly.

1.2 Effect of Meteorological Environment on Contact System

1.2.1 Strong Wind

By the end of 2007, the number of railway accidents reached 30 on Lanxin and Nanjiang railway line [3]. On Dahuai railway line, the number of faults of pantograph and OCS caused by strong wind accounts for 10% of the total number of catenary malfunction [4].

Because of long time working in strong wind environment, contact system in gale area is prone to failure. The wind will cause lateral displacement and vibration for wire, which will make the different potential wires close to each other, then trigger arc discharge or short trouble, leading to wire burn, or even break line.

1.2.2 Icing

Snow disaster occurred in southern China in early 2008. Freezing rain occurred in Liaoning Province in October 2009 and in February 2010. The snow and ice disaster caused that trains can not operate normally, which affected the transport order seriously [5]. On December 15, 2008, Japanese JR Ome Line and Hachiko Line caused 17 trains to withdraw from schedule because of icing [6].

Ice coating of contact system may affect train collection. Due to the poor electrical conductivity of ice shell, arc discharge is easy to be triggered, which can burn pantograph and contact wire. When icing seriously, pantograph cannot get supply, which makes the train lost power and unable to run normally. After icing, wire will be more sensitive to wind vibration because of the irregular wind area, so as to make contact system prone to dancing [7].

2 Design of the Monitoring Program

2.1 The Selection of Monitoring Content

2.1.1 Environmental Parameters

Environmental parameters monitored include temperature, wind speed, etc. Contact system is a complex temperature response system. Wires will expand with heat and

contract with cold when the temperature change, which makes the tension of wires and b value of tensioner change. Monitoring the environment temperature is to predict tensioner fault combined with and b value. Monitoring the wind speed is to sound a warning when over the critical wind speed that will endanger the traffic safety.

2.1.2 Temperature of Clip

There are joints on contact wire and catenary wire, and they all use clip to connect. When loose connection occurs, traction current will make joint fever. Clip temperature monitoring can judge whether the clip is loose, which can not only solve the problem that it is not easy for railway to repair joint, but also prevent abnormal power supply or break line timely, so as to improve the reliability and security of contact system.

2.1.3 Values of a and b

The b value refers to the distance between the bottom of tensioning weight and the ground. The b value is an important parameter which reflects the catenary work state, and is used for fault detection. Monitoring the b value can alarm accident in time, such as tensioning weight stagnation, break line, etc.

2.2 Hardware Design of Monitoring Terminal

Monitoring principle is as shown in Fig. 1. The PC and the monitoring terminal are equipped with the matching communicators. The terminal sends collected data to receiver by wireless transmission. Receiver sends data to PC by serial communication to store and show to users.

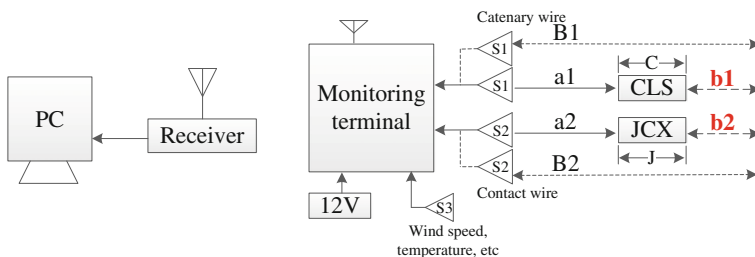


Fig. 1 Monitoring principle

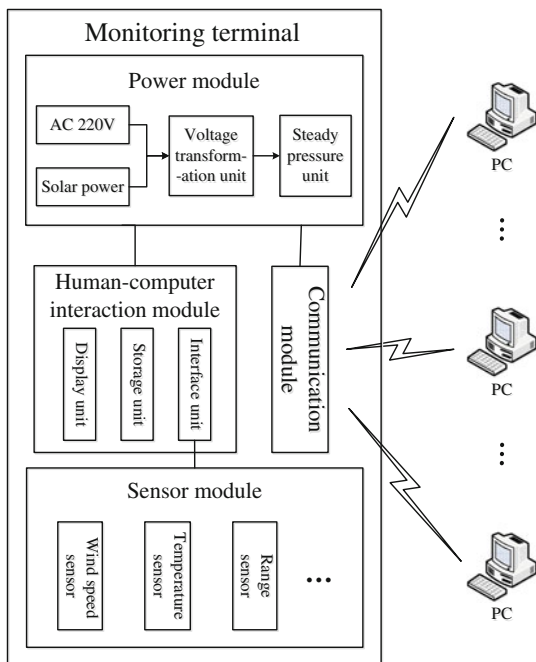
Data such as temperature, wind speed adopt the direct measurement method. While, in order to fix sensors conveniently, b value uses the indirect measurement method. The range sensor measures the distance between itself and tensioning weight as $a1$ and $a2$. The administrator inputs the distance between the sensor and the ground as $B1$ and $B2$, and the distance between the sensor and the top of tensioning weight as C and J . Then, the PC will calculate $b1$ and $b2$.

$$b1 = B1 - a1 - C \tag{1}$$

$$b2 = B2 - a2 - J \tag{2}$$

As shown in Fig. 2, monitoring terminal hardware system mainly includes power module, human-computer interaction module, communication module, and sensor module. Power module provides stable DC 12 V for terminal. The power can not only be obtained from the current which is from traction power supply system but also obtained from the solar. Two power supply modes can ensure long-term stable operation of the terminal. Human-computer interaction module adopts Kinco MT4210T for receiving, handling and storing real-time data collected by sensors, and providing the query function to view data at the scene. Communication module adopts Hongdian H7210 GPRS module, connecting to PC by wireless transmission. High measurement precision and slight error of the sensors are selected for sensor module.

Fig. 2 The structure of monitoring terminal



3 Design of Software

In Visual Basic 6.0, the monitoring system writes PC software program by VB language, and uses SQL Server 2008 for data storage and management.

3.1 Communications Protocol

Monitoring terminal connects to GPRS module by RS232 serial port. GPRS module can transform serial data to TCP/IP data packet. So, the original communication data does not need to change. So, this paper chooses Modbus to realize data transmission between PC and monitoring terminal.

Modbus is a request/response protocol. The data frame includes address code, function code, data area, and error-checking code [8]. The following part takes function code 03 which is mostly applied in this system as an example to introduce the message frame structure.

Table 1 is the message frame format of the command from PC. The data field includes the start address of register and the amount of words which need to be read. The terminal only accepts and deals with the message which is send to its own address. After receiving the message with wrong address, the terminal will discard the message directly. Only when the frame is complete and correct the terminal will response to PC. The response message frame structure is as shown in Table 2.

Table 1 The message frame format of PC command

Terminal address	Function node	Start address which will be read		The amount of words which will be read		CRC	
		High byte	Low byte	High byte	Low byte	High byte	Low byte
01	03	1B	53	00	12	33	32

Table 2 The message frame format of terminal response

Terminal address	Function node	The amount of bytes	Monitoring data	CRC	
				High byte	Low byte
01	03				

3.2 *Function and Implementation Technology of PC Software*

The overall structure of PC monitoring software is as shown in Fig. 3. Software can be divided into six surfaces: landing, real-time monitoring, setting, data management, historical data view and analysis, and print options.

On landing surface, users need to input name and password same as in the database to login. The user can also modify password. After the password was changed, it will be synchronized in the database.

Real-time monitoring surface provides users with real-time monitoring data, alarm threshold, and alert content. The software read data from the terminal by using MSComm control calls the command function to complete sending and receiving data, and then displays and stores data. The real-time data will be compared with the threshold value set before. If monitoring data is beyond the threshold, the software will show the corresponding overrun alert in the alarm box and sing warning tone. The software will also evaluate catenary security status based on real-time data.

The software connects to the database by using Adodc control. The administrator has access to change basic settings and can also import and export data by DataGrid control and EXCEL type library.

Historical data view and analysis surface connects to database and calls line drawing function to draw and analyzes the curve of three groups of monitoring parameter. a , b value are the most important parameters of monitoring software. There are tensioning weight installation curve and actual data curve in one picture, which are more intuitive for user to observe difference between the actual measured value and theoretical value, and provide the basis for further analysis.

After user selects content that are needed to be printed, the surface can print information selected by user as a report by calling print function. The report contains alarm content and monitoring parameter curve. It is easy for user to circulate and analyze.

Fig. 3 The overall structure of PC monitoring software

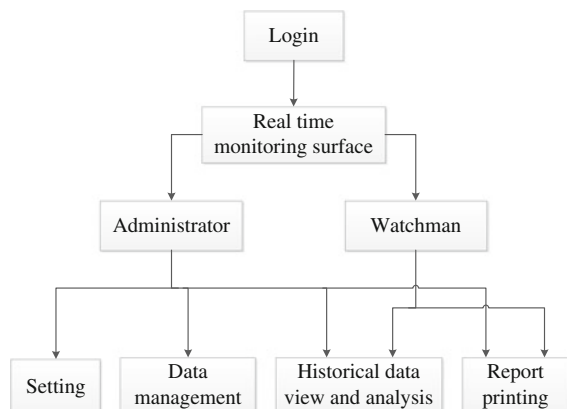
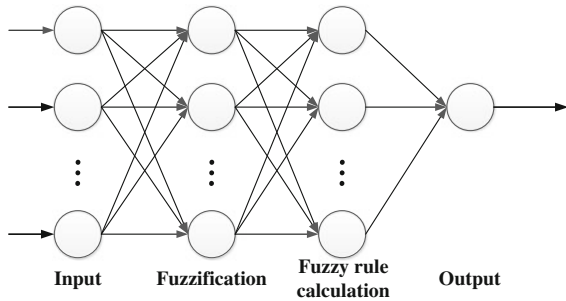


Fig. 4 The structure of fuzzy neural network



3.3 Data Analysis

The system adopts four-layer neural networks to evaluate catenary security status. The structure of fuzzy neural network is as shown in Fig. 4. At first, real-time data will be normalized and network parameters are set. Then data will be fuzzed. Finally, the software will figure out security level. This paper adopts T-S model for fuzzy treatment. This model can automatically update model according to real circumstance and correct membership function. In the process of study, gradient descent method is used to optimize the system parameter, reduce error between output and expected requirements, so as to complete network building, so that the network can analyze real-time data and evaluation catenary security status.

4 Conclusion

The monitoring terminal was installed on contact system of a power feed section in Gansu province for installing and debugging. Field installation result is as shown in Fig. 5. The human-computer interaction module can receive information measured

Fig. 5 Field installation result



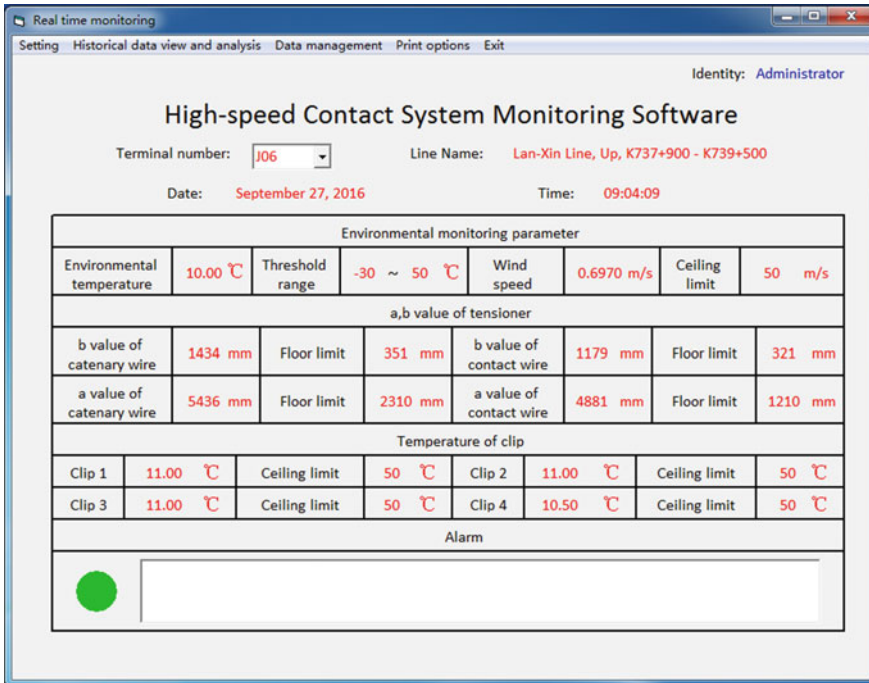


Fig. 6 Monitoring software surface

by sensors. The monitoring software was debugged on PC. As shown in Fig. 6, the software can run stably, read monitoring data in real-time, and give alarm information timely.

References

1. Dong S (2010) Contact system. China railway publishing house, Beijing (in Chinese)
2. Liu S (2013) Contact system. Southwest Jiaotong University Press, Chengdu (in Chinese)
3. Cao S, Qin J, Ke J (2010) Status and prospect of reliability of contact system under wind loads. World Sci-tech R & D 4:491–493 (in Chinese)
4. He Y (2004) Analysis on receive electric accident under strong wind touching the electric net. Inner Mongolia Sciencetech Econ 18:57–58 (in Chinese)
5. Zhen Lei (2011) Analysis of hazard of ice coating on overhead contact system and its counter-measures. Electr Railway 3:30–32 (in Chinese)
6. Kantsuu (2008) Icing of JR Ome Line contact system, power outage and stop running [EB/OL]. http://news.kantsuu.com/200812/20081216083555_134108.shtml
7. Zhang A (2008) Discussion of ice coating of overhead contact system in snowstorm disaster. Electr Railway 2008(3):32–33, 37 (in Chinese)
8. Naskar S, Basuli K, Sarma SS (2008) Serial port data communication using MODBUS protocol. ACM, America

A Method of Detecting Pantograph Slide's Abrasion Based on Image Processing

Saiyan Yu, Shuang Chen and Zongyi Xing

Abstract This paper presents a design of a detection method of pantograph slide's abrasion. First, cascade filter with dynamic intensity is adopted to filter original images. The local contrast enhancement algorithm based on local standard deviation is used to enhance the target area. Next, the adaptive threshold Canny edge detection algorithm is then performed on the preprocessed images to extract the edge of slides. The slide edges are located by Hough transform, and the actual slide residual thickness can be calculated by the camera calibration method.

Keywords Pantograph slide's abrasion · Cascade filter · Adaptive threshold Canny edge detection · Hough transform

1 Introduction

Pantograph slide plate is the only contact part between the train and the contact network [1], which is the most important power supply device of the train power supply system [2]. Serious slide abrasion may lead to serious pantograph–catenary accident. The accurate detection of the pantograph slide abrasion can not only prevent the occurrence of safety accidents but also provide the basis for the maintenance of urban rail vehicles.

The main method of pantograph detection is manual detection. Manual detection requires the train back to the library and leads to the interruption of the contact network, which makes low detection efficiency and poor detection accuracy [3]. The method of online detection is still in the initial stage and the basic principles of the system include ultrasonic ranging, CCD (Charge-coupled Device) imaging, image processing, and image recognition. Xie [4] used three sets of cameras and a

S. Yu (✉) · S. Chen · Z. Xing
School of Automation, Nanjing University of Science and Technology,
No. 200 Xiaolingwei, Xuanwu, Nanjing, Jiangsu, People's Republic of China
e-mail: yusaiyan11@163.com

group of video camera to the multi-angle shoot after the analysis of the actual installation environment. Sun [5, 6] used the principle of ultrasonic to design a set of electric bow online thickness detection device. Photoelectric sensor is used to measure the current train speed, and time-controlled ultrasonic sensor with a time difference calculates the slide abrasion value. Yue [7] proposed a pantograph slide abrasion online detection method based on active shape model (ASM).

This paper presents a design of a detection method of pantograph slide's abrasion based on image processing. The whole process does not require the train back to the library, not affecting the normal operation of the train. When the train pulled into the detection area, the automatic detection of skateboard wear is achieved and the pantograph fault can be alarmed. Filtering, image enhancement, and edge extraction algorithm have been improved accordingly and the measuring accuracy of skateboard wear is ± 0.5 mm, which meets the requirements of actual field test.

2 Design of Wear Detection Algorithm

In the process of online detection of pantograph slider wear, the original image is preprocessed to eliminate the influence of the external environment on the image quality. Next, the edge feature information is enhanced to extract the pantograph slide upper and lower edge curves. Finally, the actual slide residual thickness can be calculated by the camera calibration method, as shown in Fig. 1.

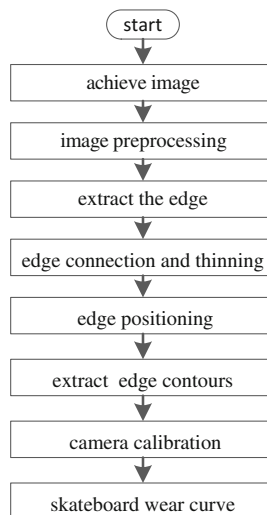


Fig. 1 Flow chart: the measurement of pantograph slide's abrasion

2.1 Image Preprocessing

In this paper, cascade filters with dynamic intensity are adopted to reduce the influence of noise and improve edge feature as much as possible. The first filter is a median filter, and the second filter is a bilateral filter. According to the denoising ability of first-class filter, the denoising ability of second-class filter can be dynamically calibrated, as shown in Fig. 2.

First, the denoising ability of the previous filter is estimated by the SAD (Sum of absolute differences) method. Assuming the partial block data of the original image to be filtered is I_1 , the partial block data of the output image after denoising is I_2 , and then the denoising ability of the previous filter can be expressed as

$$L = \frac{\|I_1 - I_2\|}{N}, \tag{1}$$

where N is the total number of pixels of the data block I_1 and I_2 .

Supposing the denoising ability of the previous filter is K , the difference between calibrated denoising ability K and the estimated noise intensity of the previous filter L is

$$\Delta K = \|K - L\|, \tag{2}$$

where ΔK is the calibrated denoising ability of the second-class filter.

The two-stage filters are cascaded to obtain a cascade filter and the dynamic denoising intensity of bilateral filter can be calculated by Eq. (2).

2.2 Contrast Enhancement

Local Adaptive Contrast Enhancement (LACE), which centered on each of the pixels in processing image, calculates the average of the pixels in the local region. Supposing $g(i,j)$ is the gray value of pixels (i,j) in processing image and the above local region can be defined as centered on (i,j) , the area of window size is $(2n + 1) * (2n + 1)$, where n is positive integer. The low-frequency part of the image can be replaced by the average pixel value in the local region, the concrete formula as follows:

$$m_g(i,j) = \frac{1}{(2n + 1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} g(k,l) \tag{3}$$

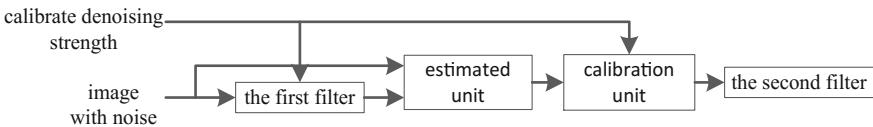


Fig. 2 Flow chart: the dynamic calibrated denoising intensity

the local variance is

$$\sigma_g^2(i,j) = \frac{1}{(2n+1)^2} \sum_{k=i-n}^{i+n} \sum_{l=j-n}^{j+n} [g(k,l) - m_g(i,j)]^2 \quad (4)$$

Defined $f(i,j)$ as the pixel gray value after the contrast enhancement of pixel points (i,j) , the LACE algorithm can be described as

$$f(i,j) = m_g(i,j) + G(i,j)[g(i,j) - m_g(i,j)], \quad (5)$$

where the function $G(i,j)$ represents the size of the contrast gain CG.

In general, the contrast gain CG is always greater than 1, so that the high-frequency part of the image $[g(i,j) - m_g(i,j)]$ can be enhanced. In this paper, the contrast gain function $G(i,j)$ is calculated as follows when amplifying the high-frequency part image:

$$G(i,j) = \frac{D}{\sigma_g(i,j)}, \quad (6)$$

where D is the constant, which represents the global mean of the image.

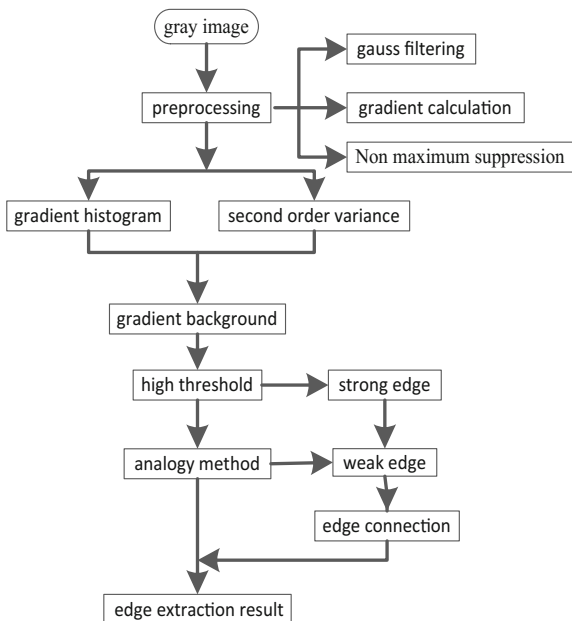
2.3 Adaptive Threshold Canny Edge Detection

Canny algorithm is the most widely used edge detection method [8]. However, the Canny operator needs to be fixed at high or low threshold in advance, which makes it easy to detect spurious edges caused by noise. This paper improves the Canny operator and proposes an adaptive threshold edge detection algorithm based on Canny, as shown in Fig. 3.

(1) Adaptive threshold calculation

According to the histogram information of processed images, the gradient value G_p corresponding to the peak value of the histogram is used as the first value of the gradient background and the global three-order variance of the mean G_p is calculated. The sum of the G_p and three-order variance is used as the background representation of the gradient. The gradient higher than the gradient background is represented as high threshold T_h . After removing the edge points higher than the high threshold, the remaining edge points are then calculated as the center of G_p to obtain the three-order variance, so that the background representation of the residual points is obtained as low threshold T_l .

Fig. 3 Flow chart: an adaptive threshold edge detection algorithm based on Canny operator



The calculation formula of high threshold is as follows:

$$T_h = G_p + \delta_1, \tag{7}$$

where G_p is the gradient value corresponding to the histogram peak value;

$$\delta_1 = \left[\frac{\sum_{i=0}^{N_1} (G_i - G_p)^n}{N_1} \right]^{\frac{1}{n}},$$

N takes 2 or 3, which is the degree of edge extraction in the whole image. The larger the value n is, the smaller the proportion of the extracted edge points to the whole image is and N_1 is the total number of pixels in the corresponding range.

The formula of low threshold is as follows:

$$T_l = G_p + \delta_2, \tag{8}$$

where G_p is the gradient value corresponding to the histogram peak value;

$$\delta_2 = \left[\frac{\sum_{i=0}^{N_2} (G_i - G_p)^n}{N_2} \right]^{\frac{1}{n}},$$

N takes 2 or 3, which is the degree of edge extraction in the whole image. The larger the value n is, the smaller the proportion of the extracted edge points to the whole image is and N_2 is the total number of pixels in the corresponding range.

(2) Strong and weak edge connection

The gray value of the pixel whose gradient less than the threshold is set to 0 and the edge above T_h becomes a strong edge H ; the edge between T_h and T_l becomes weak edge L . Centered on the strong edge H , search the weak edge in its neighborhood. If there is a weak edge in the strong edge neighborhood, the weak edge and the strong edge can be connected to extend the strong edge until the final edge map is obtained, as shown in Fig. 4.

As a matter of fact, the edge contour has been extracted to include more than one-pixel width and further refinement of the image is required. The pantograph edge after morphological thinning is shown in Fig. 5.

2.4 Contact Line and Skateboard Edge Positioning

Hough transform can be used to locate the upper edge of the pantograph slide, the lower edge of the slide support frame, and the catenary wire.

Contact line positioning: Hough transform can be used to detect and curve fitting the image. Calculate the slope of the line and extract a line that slopes up to the range of $[80^\circ, 135^\circ]$ and $[45^\circ, 110^\circ]$. Remove a false contact wire edge of a catenary or catenary formed on a slide plate according to the detected length of the line. The lead edge is positioned to the desired contact wire portion, as shown in Fig. 6.

Upper edge positioning: If the upper edge does not have an intersection with the located wire, select the detected straight lines. By determining whether the slope of the straight line in the $[-20^\circ, 20^\circ]$ range, remove some pseudo-skateboard edge curve. Then, determine whether the straight line with the contact line edge of the

Fig. 4 Image: pantograph edge extraction

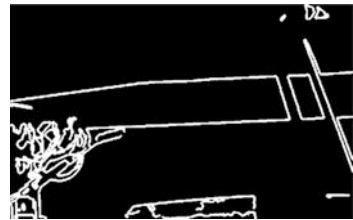


Fig. 5 Image: the edge extraction after refinement



Fig. 6 Image: the catenary wire in left half bow image

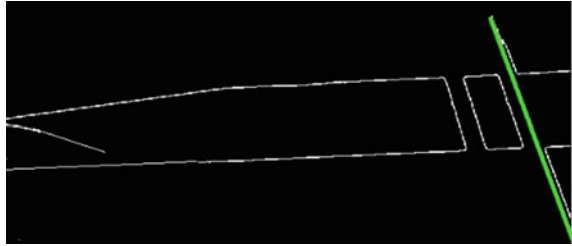
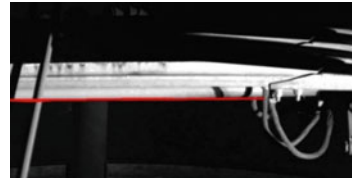


Fig. 7 Image: the lower edge of pantograph support



existence of the intersection or the distance between the two less than 70 pixels. Finally, select the image at the top of the straight line as the upper edge of the skateboard.

Under the edge location: Make sure the gradient value of the lower edge of the support frame is larger than the set threshold value. According to the information of the upper edge of the positioned skateboard, search the upper edge of the support frame and put the upper edge near the end of the contact line as the starting point. Choose the lower edge of the support frame as the candidate point and connect the edge line segment whose horizontal direction is discontinuous but the angle between the straight segments less than 20° . Then, use Bézier to connect the horizontal to the discontinuous edge point, so the lower edges of the support frame can be detected (Fig. 7).

2.5 Skateboard Wear Curve

Supposing the number of pixels in the vertical direction of the upper edge of the pantograph and the reference line is n and the correction factor after the camera calibration is S_c , the vertical distance between the upper edge of the slide plate and the reference line is $l = n * S_c$. Since the fixed thickness of the pantograph slide support frame is 20 mm and the distance between the reference line and the upper edge is l , the actual residual thickness of the pantograph slide plate is h :

$$h = 20 - l \tag{9}$$

3 Experimental Results and Comparative Analysis

The pantograph skate images collected in a variety of cases include the normal skateboard image, the rainy skateboard image, the insufficient fill light, and the excessive exposure. Result is shown in Table 1.

Dealing with images captured on rainy days, the accuracy of the edge positioning is close to that of the normal skateboard image, which indicates that the algorithm proposed in this paper can effectively eliminate the influence of the raindrop on the lower edge and the upper edge of the slide support frame. In the case of insufficient illumination or too strong, the accuracy of the skateboard positioning is 81.25% and 87.5%, respectively, which indicates the accuracy of the fill light directly affecting the accuracy of the skateboard positioning.

Poor fill light effect or shadows produced by light projection results in skateboard edge characteristics not obvious, which may mistake the skateboard edge curve and inaccurate skateboard positioning.

In order to further analyze the accuracy of the measurement results and the stability of the measurement results, the repeatability of the pantograph slider's wear test is carried out to verify the robustness of the algorithm. Table 2 shows the minimum remaining thickness of system value under different circumstances and speeds. Table 3 shows the minimum manual measured thickness of the skateboard.

From the above two tables, the number of error more than 0.5 mm of the slide abrasion value is 7, accounting for 7% of the total number of the measuring slide, and the number of error more than 0.3 mm of the slide abrasion value is 26, accounting for 26% of the total number of the measuring slide. The system of repeated measurement accuracy is 0.49 mm, meeting the requirements of field inspection accuracy.

Table 1 System resulting data

The type of image	Normal	Rainy	Insufficient light	Excessive exposure
The sum number of image	396	140	48	16
The number of correct positioning	382	132	39	14
The proportion of correct positioning	96.46%	94.28%	81.25%	87.5%

Table 2 Minimum remaining thickness of system value

Slide number	The measurement times				
	1	2	3	4	5
1	16.56	16.34	16.78	16.15	16.17
2	16.4	16.43	16.48	16.12	16.47
3	15.95	15.48	16.44	15.97	16.52
4	16.43	16.51	16.14	16.15	16.61
5	17.89	18.1	18.06	17.69	17.99
6	15.6	15.32	15.13	15.9	16.12
7	13.58	13.45	13.32	13.79	13.8
8	11.33	11.4	11.81	11.82	11.61
9	14.61	14.68	14.45	14.34	14.16
10	9.77	9.52	9.61	9.49	9.28
11	10.57	10.32	10.61	10.29	10.33
12	17.89	18.1	18.06	17.79	17.69
13	12.41	11.99	12.39	12.48	12.49
14	12.97	13.01	12.79	12.58	12.94
15	11.89	12.3	11.99	11.81	12.13
16	8.14	8.53	8.16	8.26	8.74
17	11.38	11.33	11.36	11.44	11.57
18	12.05	11.65	11.76	11.78	12.06
19	15.1	15.3	15.52	15.34	15.2
20	6.22	6.46	6.48	6.36	6.36

Table 3 Minimum residual thickness measurement of manual value

Slide number	1	2	3	4	5	6	7	8	9	10
Manual value	16.9	16.7	15.7	16.6	17.8	15.6	13.5	11.5	14.4	9.3
Slide number	11	12	13	14	15	16	17	18	19	20
Manual value	10.3	17.8	12.3	12.7	12.1	8.2	11.4	11.8	15.4	6.4

4 Conclusion

This paper presents a design of a detection method of pantograph slide’s abrasion. First, the original image is denoised by the cascade filter with dynamic intensity, and the target region is enhanced by the local contrast enhancement algorithm based on local standard deviation; Then, an adaptive threshold Canny edge detection algorithm is proposed, and the slide edges is located by Hough transform. Finally, the distance between the upper and lower edges of the slider is converted to the actual thickness of the actual slide to obtain the remaining thickness curve of the skateboard.

Acknowledgements This work was supported by the National Key Research and Development Plan (No.2017YFB1201201). The author gratefully acknowledges the anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this paper substantially.

References

1. Hao X, Wang S (2013) Optimization modeling and simulation of power supply network for urban rail transit. *Comput Simul* 30(10):194–198 (in Chinese)
2. Shi S (2016) Study on measures to reduce wear of overhead rigid suspension contact net. *Eng Engl* (10):00166 (in Chinese)
3. Cui Z (2005) Pantograph skateboard external state image detection. Southwest Jiaotong University (in Chinese)
4. Xie L (2009) Research on pantograph state detection technology based on image processing. Southwest Jiaotong University (in Chinese)
5. Sun F, Wang B (2011) Ultrasonic testing method for wear of double slide plate. *Develop Innov Electromechanical Prod* 24(3):129–131 (in Chinese)
6. Sun F (2011) Design and debugging of wear test system for pantograph slide. Southwest Jiaotong University (in Chinese)
7. Yue A, Zhao Z, Wang C (2003) Pontoon automatic detection method based on active shape model. *Chinese Soc Vis Image Anal* (in Chinese)
8. Canny JF (1987) A computational approach to edge detection. In: *Readings in computer vision: issues, problems, principles, and paradigms*. Morgan Kaufmann Publishers Inc

Subsystem Characteristics-Based Modeling Method for the Simulation of Electromagnetic Compatibility of Rail Transit Vehicles

Dafa Jiang and Zhongcheng Jiang

Abstract Rail transit vehicles consist of subsystems and components with diverse electrical and electromagnetic characteristics, which brings difficulties to the modeling and simulation of their electromagnetic compatibility (EMC). In this paper, we proposed a modeling method for the EMC simulation of rail transit vehicles based on the subsystem characteristics. Vehicle body model, cable model, and equivalent circuit model are built and integrated for multilevel simulation of the EMC of rail transit vehicles, including body shielding effect, interior magnetic field distribution, and electromagnetic interference emission. This modeling method extracts the electrical and electromagnetic characteristics of subsystems with high accuracy and efficiency, and contributes to the practicability of EMC simulation of rail transit vehicles.

Keywords Rail transit vehicles Electromagnetic compatibility Subsystem characteristics Modeling Multilevel simulation

1 Introduction

Rail transit vehicles, including locomotives, multiple-unit trains, metro vehicles, streetcars, etc., consist of a series of subsystems and components with different electrical and electromagnetic characteristics, including current collection device, traction drive system, train network system, and so on. Currently, the power electronics devices are evolving toward higher power and frequency, causing severe high-order harmonics and electromagnetic radiation. On the other hand, electronic devices are getting increasingly sensitive to electromagnetic interference (EMI)

D. Jiang (✉) · Z. Jiang
State Key Laboratory of Heavy Duty AC Drive Electric
Locomotive Systems Integration, CRRC Zhuzhou Locomotive Co. Ltd,
No. 1 Tianxin Road, Zhuzhou, Shifeng District, Hunan Province, China
e-mail: jiangdafa@crrezelc.cc

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_7

[1, 2]. Therefore, the EMC of rail transit vehicles is attracting more and more attention all over the world.

EMC simulation at design stage is an important approach for ensuring the EMC of products. Currently, EMC simulation is widely used for the development of consumer electronics products. Simulation methods such as full-wave 3D simulation and mixed-mode simulation are developed [3, 4]. However, rail transit vehicles are much more complex than consumer electronics devices. The abovementioned methods individually are not qualified for the simulation of EMC of rail transit vehicles [5, 6]. Therefore, it is essential to develop modeling and simulation methods for investigating the EMC of rail transit vehicles.

In this paper, we proposed a modeling method for the simulation of EMC of rail transit vehicles based on subsystem characteristics. This paper consists of three parts. First, the subsystem characteristics-based modeling method of rail transit vehicles is introduced. Second, multilevel EMC simulation is carried out to evaluate the body shielding effect, interior magnetic field distribution, and cable coupling of a metro vehicle. Finally, a case is introduced to demonstrate the application of this method.

2 Subsystem Characteristics-Based Modeling Method of Rail Transit Vehicles

Accurate modeling of the electrical and electromagnetic characteristics of rail transit vehicles is essential for EMC simulation. As abovementioned, the currently available modeling methods individually are not qualified for EMC simulation of rail transit vehicles. For the application of EMC simulation in the development of rail transit vehicles, it is a rational approach to employ the well-developed modeling methods according to the specific characteristics of the subsystems and components. The models are then integrated for multilevel EMC simulation [7, 8], as shown in Fig. 1. In this section, the methods employed for the modeling of the vehicle body, electronic devices, and cables are discussed in detail.

2.1 Vehicle Body

Generally, the vehicle body consumes the largest amount of calculation time during the EMC simulation. The vehicle body consists of body structure and interior decoration components. The former is made of stainless steel or aluminum alloy, while the latter is made of nonmetal materials. The nonmetal materials have so little impact on electromagnetic fields as to be neglected. Therefore, the modeling of vehicle body takes consideration of only the metallic body structure.

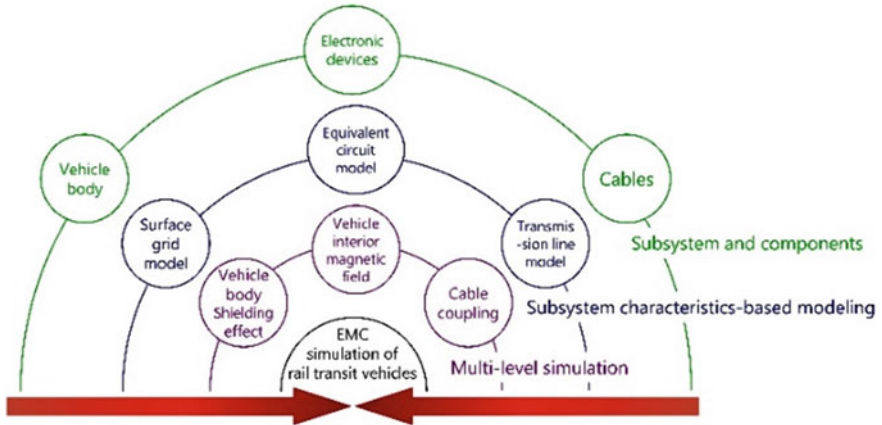


Fig. 1 Schematic illustration of subsystem characteristics-based modeling and simulation of EMC of rail transit vehicles

Currently, rail transit vehicles are designed using computer-aided design (CAD) techniques. The CAD file (Fig. 2a) is employed to improve the accuracy and efficiency of the modeling procedure. Nonmetal material components are removed, left behind the metallic body structures, as shown in Fig. 2b. The following two steps are taken for further simplification of the model:

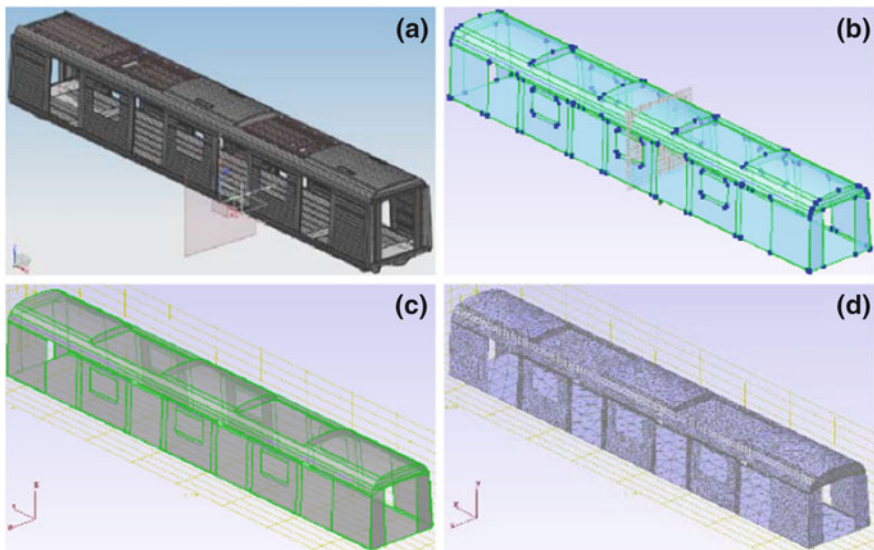


Fig. 2 The modeling procedure of vehicle body. **a** CAD file, **b** metallic body structure, **c** simplified as surface model, **d** surface grid model

1. The metallic vehicle body structures have specific thickness. Considering the skin effect, the electric current mainly distributes within a very thin surface layer of the vehicle body. Therefore, the vehicle body can be simplified to a shell with zero thickness, i.e., surface model, as shown in Fig. 2c, so as to reduce the grid amount and calculation time significantly.
2. The vehicle body contains a lot of small-size structures having negligible impact on its EMC behavior, for example, the equipment installation sites. These small-size structures are removed after careful evaluation.

The surface model is then meshed into grids to obtain the surface grid model, as shown in Fig. 2d. The size of the grids depends on the concerned frequency (f) of the simulation, which generally should be smaller than the one-tenth of the wavelength (λ) corresponding to the concerned frequency, i.e., $\lambda/10$.

2.2 Electronic Devices

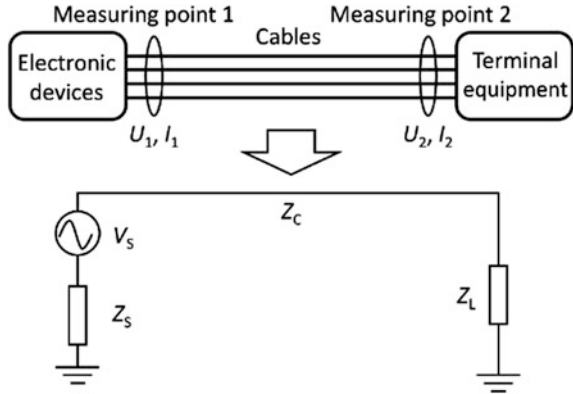
Electronic devices are the radiation source and victim equipment of EMI, thus becoming the focus of EMC simulation of rail transit vehicles. The electronic devices, such as the transformer and inverter, are carefully protected from EMI via smoothing and shielding carried out by the suppliers. Therefore, compared with the interference coupled via cables, the radiation and conduction interference emission by electronic devices, as well as the coupling to external interference, are negligible. On the other hand, electronic devices are usually supplied by the suppliers, rail transit vehicles manufacturers are not aware of their internal circuit structures. Therefore, electronic devices are usually regarded as black boxes connected by cables, and frequency-domain equivalent circuit models are built.

Some tests are needed to acquire the characteristic parameters of equivalent circuit models. Figure 3 shows an interference emission model, which contains an interference voltage source, a source-end impedance, and a terminal impedance. This model demonstrates the interference associated with the electronic devices and makes an important part of the integrated model for EMC simulation of rail transit vehicles. The voltage source V_S , source-end impedance Z_S , and terminal impedance Z_L are calculated based on the measurement results of the voltage V and current I in the cable at both the electronic devices side and the terminal equipment side, i.e., measuring point 1 and 2 in Fig. 3. V , I , and Z are all complex function of frequency f .

$$\frac{U_1(f)}{I_1(f)} = Z_c(f) + Z_L(f)$$

$$\frac{U_2(f)}{I_2(f)} = Z_c(f) + Z_S(f),$$

Fig. 3 Schematic illustration of the equivalent circuit model of electronic devices



where $U_1(f)$ and $U_2(f)$ are the source-end and terminal voltage, $I_1(f)$ and $I_2(f)$ are the source-end and terminal current, $Z_S(f)$ is the impedance of the equivalent voltage source, $Z_c(f)$ is the impedance of the cables, and $Z_L(f)$ is the impedance of the terminal equipment.

$$Z_C(f) = \frac{\frac{j}{2\pi fC} (R + j2\pi fL)}{\frac{j}{2\pi fC} + R + j2\pi fL},$$

where R is the resistance of the cable, L is the inductance of the cable, and C is the capacitance of the cable.

2.3 Cables

The cables play an important role in the EMC of rail transit vehicles, because they are the main source of EMI emission, spreading way of conductive interference and coupling way of radiative interference. The model of cables is usually built using transmission line model (TLM) [9] to calculate the voltage and current distribution inside the cables.

Cable model describes the structural and electrical information of the cables. The structural information contains the length and path of the cables, which can be exported from the cable design files. The electrical information includes the amount, type, diameter, and materials of the cables. The model is meshed into grids, of which the size should also be smaller than $\lambda/10$, as shown in Fig. 4.

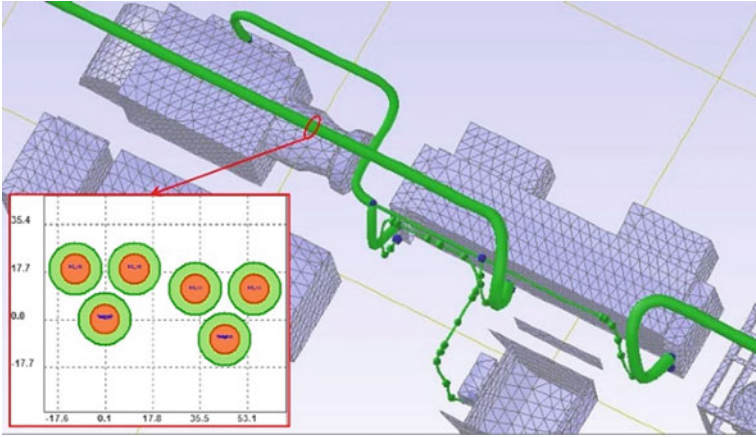


Fig. 4 TLM of the traction system cables

3 Multilevel Simulation of the EMC of Rail Transit Vehicles

Having finished the modeling procedure, the subsystem characteristics-based models are integrated for EMC simulation. Appropriate simulation methods are selected based on the electromagnetic characteristics of subsystems. For example, employing circuit simulation method to calculate the interference source noise, transmission line simulation method to calculate the current distribution in cables, and full-wave simulation method to calculate the cable coupling and current distribution in vehicle body. As soon as the calculations are finished, the field intensity of electromagnetic radiation both inside and outside the vehicle, and induced current distribution in the cables and components are revealed, which are inspiring for the modification of the EMC of rail transit vehicles.

In this section, a metro vehicle is taken as example to demonstrate the application of the above subsystem characteristics-based modeling method in the EMC simulation of rail transit vehicles. The integrated model is shown in Fig. 5; it consists of the surface grid model of vehicle body, the equivalent circuit model of electronic devices (including high-voltage power supply network, driving module, braking module, etc.), and transmission line model of cables. The body shielding effect, interior magnetic field distribution, and cable coupling are calculated based on the integrated model.

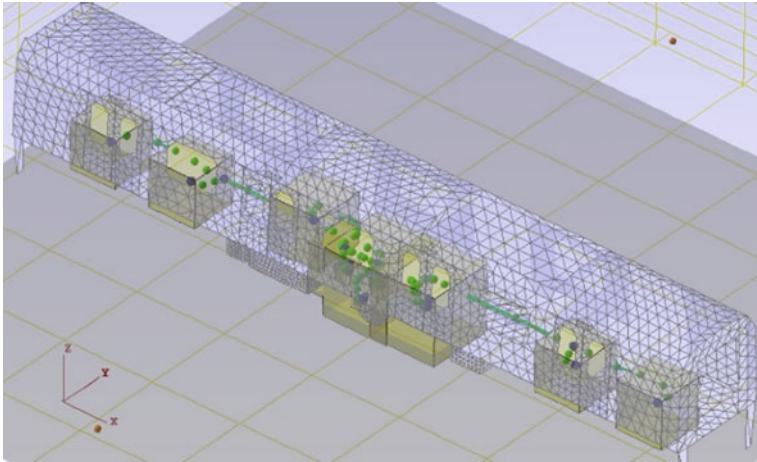


Fig. 5 The integrated model for EMC simulation of a metro vehicle

3.1 Vehicle Body Shielding Effect

The simulation result of vehicle body shielding effect is shown in Figs. 6 and 7. It reveals the dependence of vehicle body shielding effect on vehicle body structure and electromagnetic radiation frequency. As shown in Fig. 6, the continuity of the vehicle body is key to its shielding effect. At the position of windows and doors, where vehicle body is discontinuous, the vehicle body shielding effect is significantly lower than at the other parts of the vehicle body. With the increase in electromagnetic radiation frequency, the vehicle body shielding effect deteriorates. Therefore, within low-frequency range, the vehicle body is capable of protecting the passengers from low-frequency electromagnetic energy which is usually

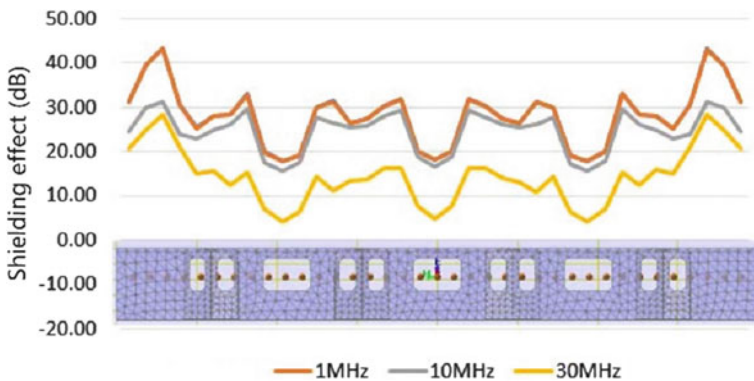


Fig. 6 The simulation result of the vehicle body shielding effect

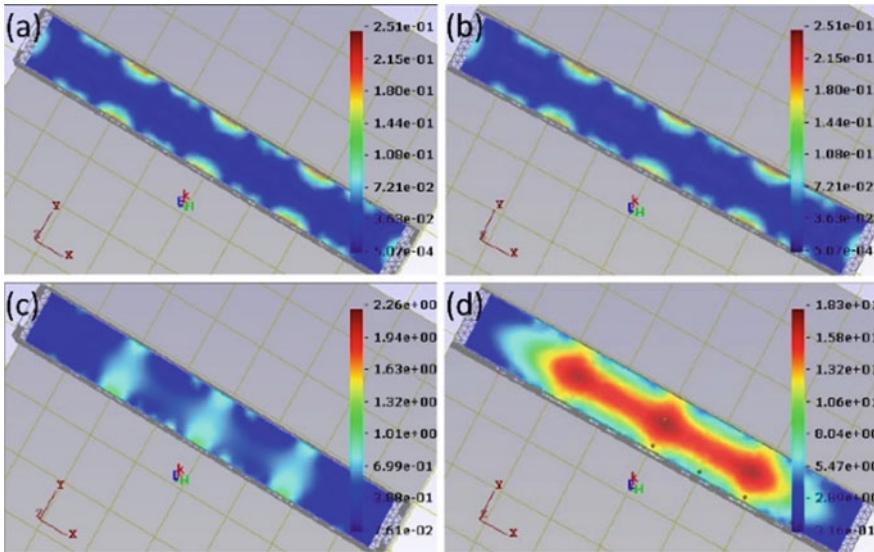


Fig. 7 The simulation results of the electromagnetic field distribution within the vehicle body at different frequencies. **a** 1 kHz, **b** 1 MHz, **c** 40 MHz, **d** 53 MHz

harmful to human body, while within high-frequency range, the vehicle body causes little influence to the onboard high-frequency communication equipment. 53 MHz is a resonant frequency of the system. At this frequency, the shielding effect of vehicle body is destroyed, as shown in Fig. 7d.

The simulation results of the vehicle body shielding effect offer guidance for the design of rail transit vehicles. On one hand, the simulation method can be employed to predict the shielding effect of vehicle body and cabinets, thus evaluating the systematic EMC risk. On the other hand, the simulation results are helpful for identifying the resonant frequency and position, thus avoiding the use of devices of which the working frequency is close to the resonant frequency.

3.2 Vehicle Interior Magnetic Field Distribution

The interior low-frequency magnetic field distribution of a Mp vehicle is revealed by simulation, as shown in Fig. 8. The result indicates that the electronics devices installed at the bottom of the Mp vehicle, such as the inductor, traction inverter, and braking resistor, have decisive impact on the interior magnetic field distribution of Mp vehicle. The position right above the braking resistor demonstrates the highest magnetic flux. The simulation result of the interior magnetic field distribution offers guidance for the design of shielding measures aimed at avoiding excessive electromagnetic field intensity in the vehicle.

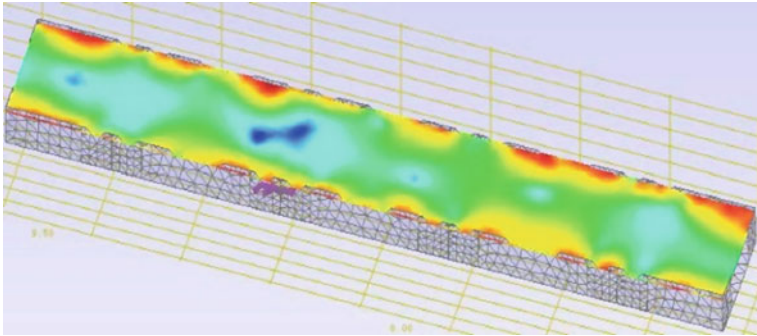


Fig. 8 The simulation result of the low-frequency magnetic field distribution in Mp vehicle

3.3 Cable Coupling

By simulation using the integrated model, the influence of EMI noises from the auxiliary inverter to the cables 1 m away is taken into consideration. The resistance to EMI of three types of cables, i.e., single wire, twisted pair wire, and shielded twisted pair wire, is compared. A pulse-width modulated (PWM) pulse source is employed to simulate the EMI noises from the auxiliary inverter. Figure 9 demonstrates the comparison between the time-domain measurement results and simulation results, which confirms quite good consistency between the two results, i.e., the main peaks of cable-coupled voltage are accurately identified by simulation. Frequency-domain simulation of cable-coupled voltage is conducted covering frequency range from 9 kHz to 30 MHz. As shown in Fig. 10, among the three types of cables, the coupled voltage is higher than 60 mV in single wire, while in twisted pair wire it is about 20 mV and in shielded twisted pair wire about 0.01 mV. It is clearly revealed that the resistance of these three types of cables to cable coupling: shielded twisted pair wire better than twisted pair wire and better than single wire.

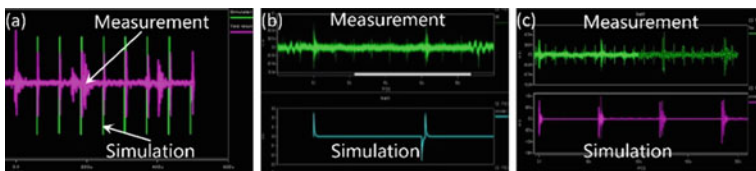


Fig. 9 Comparison between the time-domain measurement results and simulation results of the cable-coupled voltage

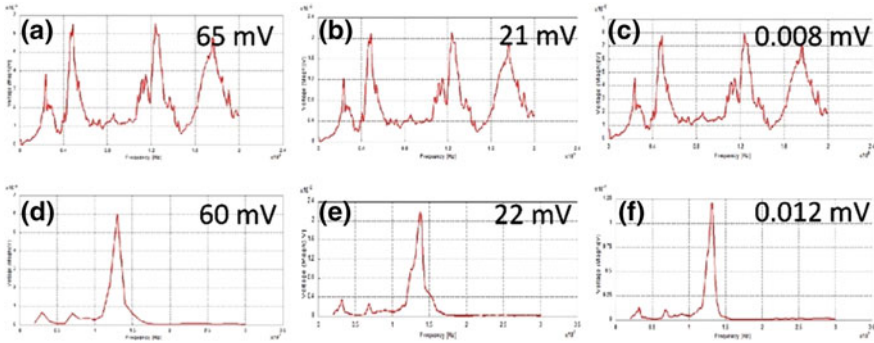


Fig. 10 Comparison between the frequency-domain measurement results and simulation results of cable-coupled voltage, frequency range: 9 kHz–2 MHz (top) and 2–30 MHz (bottom) **a, d** single wire, **b, e** twisted pair wire, **c, f** shielded twisted pair wire

4 Application Case of EMC Simulation of Rail Transit Vehicles

The above EMC modeling and simulation method make a powerful tool for the design and management of the EMC of rail transit vehicles. The case discussed in this section demonstrates the application of this method in the EMC modification of a metro vehicle.

A metro vehicle was applied with low-frequency magnetic field test. The result reads the DC magnetic field intensity to be 1.32 mT 30 cm right above the inductor, higher than the limit defined in EN 45502-2-1:2003 standard, which is 1 mT. Further analysis deduced that this excessive DC magnetic field intensity might originate from the DC current in the input/output cables of the inductor. Based on the analysis, two modification approaches were proposed. One was to exchange the position of the input/output cables of the inductor, as shown in Fig. 11a. The other was to change the direction of the input/output cables of the inductor, as shown in Fig. 11b. By EMC simulation and calculation, we figured out that both methods were capable of reducing the DC magnetic field intensity at the testing spot, and the first one was better. Therefore, we exchanged the position of the input/output cables of the inductor. Subsequent tests indicated that the DC magnetic field intensity at the testing spot decreased to 0.4–0.7 mT, in compliance with the requirement of EN 45502-2-1:2003. In this application case, EMC simulation result offers guidance for the modification of the EMC of a metro train, and thus help to improve the efficiency of our job.

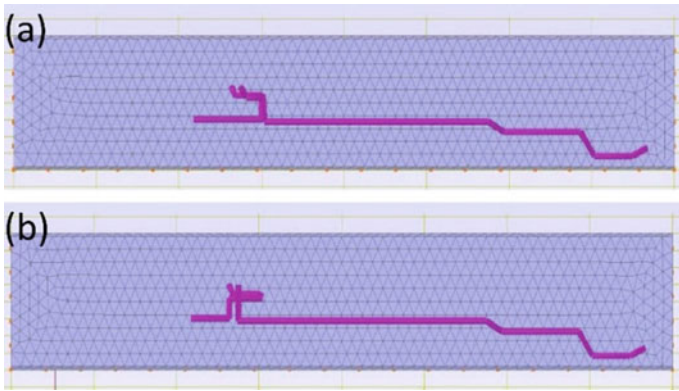


Fig. 11 Simulation model of the proposed two cable layouts toward the modification of the EMC of the metro vehicle

5 Summary

Currently, rail transit vehicles are getting increasingly intelligent and using more and more electronic devices, which increases the risk of EMI. EMC simulation is an important approach to ensure the EMC of products. To meet the needs of EMC simulation of rail transit vehicles, we proposed a subsystem characteristics-based modeling method for the EMC simulation of rail transit vehicles in this paper. Surface grid model, equivalent circuit model, and transmission line model are built based on the electrical and electromagnetic characteristics of the vehicle body, electronic devices, and cables. The models are integrated for multilevel simulation to calculate the vehicle body shielding effect, vehicle interior magnetic field distribution, and cable coupling. The simulation results show high consistency with the measurement results. This method achieves high-accuracy EMC simulation of rail transit vehicles and offers guidance for the EMC design and modification of rail transit vehicles.

References

1. Midya S, Thottappillil R (2008) An overview of electromagnetic compatibility challenges in European rail traffic management system. *Transport Res Part C Emerg Technol* 16(5):515–534
2. Jiao J, Shi D, Zhang Y (2012) Research of the electromagnetic compatibility problems caused as putting heavy and weak current facilities together in the subway system. In: 2012 6th Asia-pacific conference on environmental electromagnetics (CEEM). IEEE
3. Zhao H, Li Y, Zheng S, Yu L (2011) Status of electromagnetic compatibility research and technology application in some areas. *Appl Mech Mater* 50–51:688–692

4. Qian Z, Xin W, Lu Z, Pong M (2000) Status of electromagnetic compatibility research in power electronics. In: Power electronics and motion control conference, 2000. IP EMC 2000. The third international IEEE, vol 1, pp 46–57
5. Yuan Y, Chu Z, Tong H, Li B, Ding H (2014) Simulation methods and engineering practices of EMI/EMC in railway vehicles. *Transportation electrification Asia-Pacific IEEE*, pp 1–4
6. Luan X, Zhu H, Qiu B, Han B (2016) EMC in rail transportation. *Energy Procedia* 14:526–531
7. Li X, Xiong R, Wang L, He J, Tang K (2014) Electromagnetic compatibility multi-software co-modeling technology of vehicle level complex structures. *J Microwaves* 30(1):15–19 (in Chinese)
8. Lv J, Jiang Z, Xie L, Wang X, Zhang J, Wang Y (2014) Multi-level simulation of metro EMC. *Res Urban Rail Transit Netw* 6:121–125 (in Chinese)
9. Xie L, Lv J, Jiang Z, Yang Y, Wang X, Zhang J, Wang Y (2016) Simulation analysis and test verification of metro cable crosstalk. *Res Urban Rail Transit Netw* 5:78–82

Failure Analysis and Discussion of Bogie Temperature Sensor for Electric Multiple Unit

Yantong Liu, Wenjun Wang, Jintian Wang, Wanxiu Teng and Zhaoyu Ma

Abstract In order to solve the problem of delay and safety of EMU, caused by insulation decrease, temperature jump, open circuit, and structural damages of bogie temperature sensor, this paper analyzes the causes of failure, expounding the defects of traditional structure and installation method of the sensor on the bogie. By discussing two kinds of temperature monitoring methods (infrared and wireless technologies), it is shown that wireless sensor technology is usable in temperature monitoring redundancy design. The proposal of installing the emergency wireless temperature monitoring device at the point of abnormal temperature is put forward, so as to ensure that the bogie temperature is always controlled and that the punctuality and safety of the EMU are improved.

Keywords Temperature sensor · Failure analysis · Scheme discussion

1 Introduction

Nowadays, China's high-speed rail mileage has reached 22,000 km, and more than 2600 standard groups of train are operating online. As the EMU brings convenience to people, community has paid more attention to the security and punctuality of high-speed EMU. In order to ensure the safety of EMU operating and to avoid the phenomenon of hot axle and axle sheared, all of the bogies of EMU are equipped with temperature monitoring system to monitor the axlebox and gearbox temperature. However, it is not uncommon that a bogie temperature sensor failure occurs during train operation, and this failure makes the train lose the temperature monitoring capability, which will influence the punctuality and safety of the train. This paper mainly introduces the typical failures of the bogie temperature sensor in revenue service, analyzes the causes of the failure, discusses the defects of the

Y. Liu (✉) · W. Wang · J. Wang · W. Teng · Z. Ma
Railway Vehicle Laboratory, R&D Center, CRRC Changchun Railway
Vehicles Co., Ltd., No. 2001 Changke Road, Changchun, Jilin, China
e-mail: liuyantong@cccr.com.cn

structure and installation method of the sensor on the bogie, and then proposes a reasonable and effective solution to protect the safety of the EMU punctual operation [1].

2 Research Background

According to the statistics, A-type EMU had 144 temperature sensor failures in 2016, including false alarm, circuit disconnection, and other issues. 28 faults out of 144 were caused by the temperature sensor, 116 by hardware connectivity, and other problems. Figure 1 presents the time record of the gearbox temperature of EC01 carriage on a A-type EMU. The temperature exceeded the limit between 23:00 PM and 23:03 PM. It was obviously a temperature jump problem, the big gear of gearbox 4 was in test error, the train was required to limit the speed to level 2, and the drive shaft was required to be monitored. The train drivers found no temperature exception after 23:03. According to IEC61373 standard, the sensor was taken to the lab and tested in the longitudinal vibration environment (frequency range 10–500 Hz, longitudinal acceleration 6.43 g). The results are shown in Fig. 2: the vertical axis is temperature, the horizontal axis is time, channel 1 of gearbox temperature sensor generates a temperature jump, and it was the cause of temperature failure. Therefore, it is urgent to eliminate the safety hazard and to avoid the unnecessary speed limit caused by the failure of the temperature sensor. This paper analyzes the failure of service temperature sensor through the sensor structure and the assembly [2, 3].

The probe of temperature sensors use Pt100 or Pt1000 platinum thermal resistance as a sensitive element. The working principle is that the platinum thermal resistance changes with temperature, and the temperature of the measured environment is assessed by measuring its resistance.

3 The Failures of Temperature Sensor

Most of the bogie temperature sensors are exposed after installation, and some of them are fixed on the axlebox which has large vibration levels. So the IP level and anti-vibration performance requirements of the sensors are rather severe [4]. In

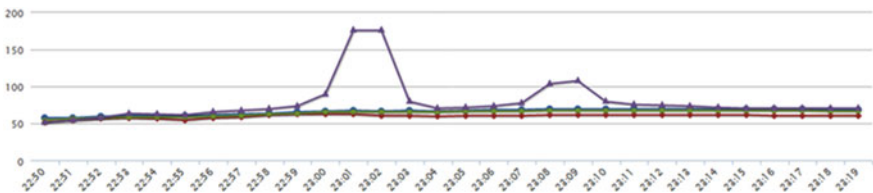


Fig. 1 Time record of gearbox temperature (remote data, time range 22:50–23:19)

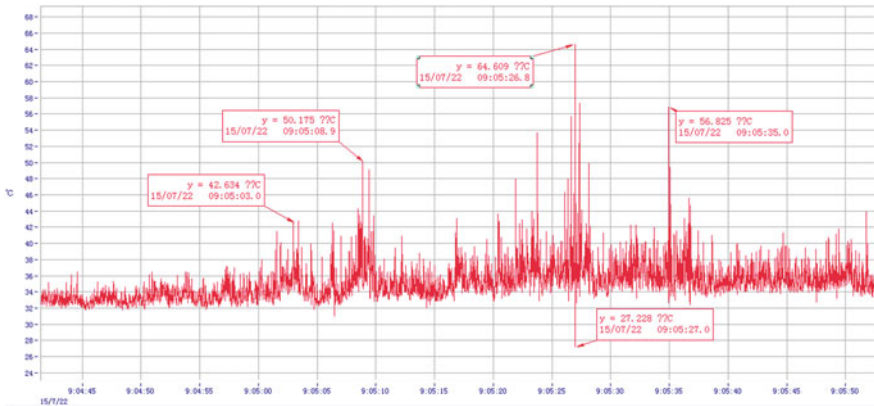


Fig. 2 Longitudinal vibration test results of gearbox temperature sensor

addition, the temperature sensor is assembled manually, which may lead to the defects of sensor and increase the possibility of sensor failure; besides, the connections between temperature sensor and other electrical parts and working voltage stability can also cause the sensor failure; and the paper presented the unreliability reasons of the structure and installation of the temperature sensor.1

3.1 Deterioration of Insulation and Voltage Resistance

The decrease of the insulation and of the voltage resistance performance of the temperature sensor can lead to false alarm. Indeed, first, because of the influence of the internal structure, of the potting process, and of the cable material performance [5], the screened wire is likely to stab the insulation layer inside and to get connected with the mechanical shell outside. In fact, the insulation performance of point to shield and point to the shell decreases (see the shield of sensor in Fig. 3). Second, because of the multiple parts connected to each other on the temperature

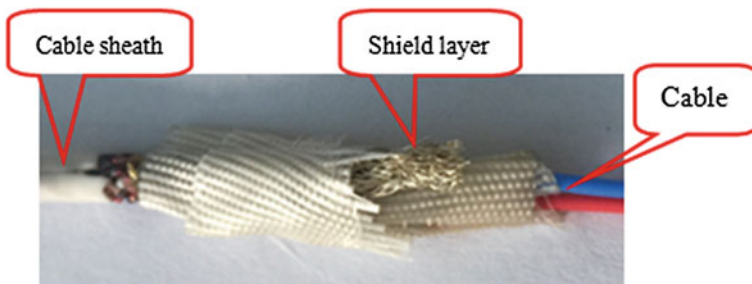


Fig. 3 Shield layer of axlebox temperature sensor

sensor, the reliability of the sealing ring and sealing plug faces great challenge, the moisture can easily permeate the interior of the components through these connection points under the harsh working environment. Figure 4 shows a temperature sensor that does not meet the requirements of IP68 during the laboratory waterproofing test.

Most of the manufacturers enhance insulation and voltage resistance performance by potting process and cable protection. Potting process is the pretreatment process of bonding surface [6]: the surface area and the bonding strength will be enhanced by adding surface roughness and cleanliness of the metal bonding. Another aspect is the choice of glue type. Conventional potting glues include silicone rubber, epoxy resin, and polyurethane. Silicone rubber is soft rubber; the disadvantage is that the bonding performance between the solidified elastomer of silicone rubber and metal is poor; there is a certain internal stress after solidifying; so silicone rubber is generally not suitable for the mechanical seal of the outermost layers. The bonding performance of polyurethane is good but the temperature resistance is poor, often less than 120 °C.

Moreover, the probability of insulation layer piercing and of connection between mechanical shell and insulation layer increases for the parts which are often bent, such as the extremity of the sensor probe and the root of cable shield. They are often subjected to external forces induced by the bending. This situation increases the probability of piercing, and it also damages the cable itself. At present, there is no sufficient experimental feedback to confirm that the reliability of shield and cable is good or not under the condition of long-term bending state. As shown in Fig. 5, a cable fracture occurred during the service of a temperature sensor.

The mechanical connections between the different parts of a temperature sensor can be improperly handled: this can affect the insulation and voltage resistance performance insulation. The following are shown in Fig. 6: (A) crimping position between inside, outside sleeve, and the rubber sheath at the end of sensor; (B) crimping position between inside, outside sleeve, and the rubber sheath; (C) the connection between inner sleeve and tail nut; (D) the connection between end nut

Fig. 4 The water permeated into sensor after waterproofing test

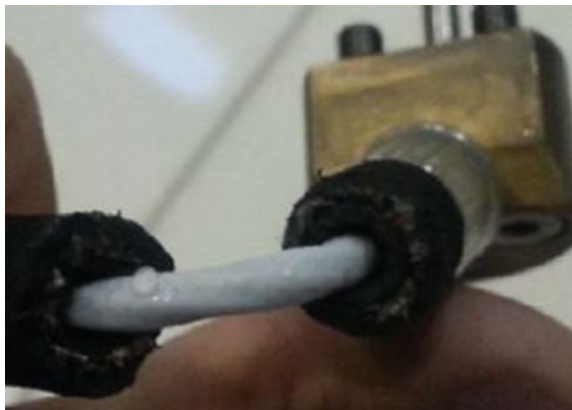
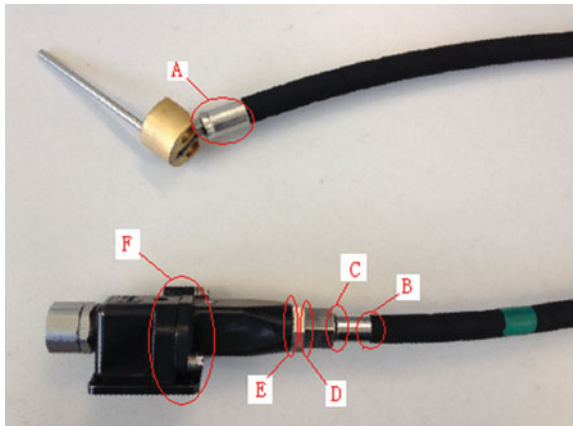


Fig. 5 Damaged part of the extremity of an axlebox temperature sensor



Fig. 6 Mechanical connections of an axlebox temperature sensor



and middle joint; (E) the connection between middle joint and connector; and (F) the connection between connector and docking device. So the sealing ring and plug structure designs have an important influence on the waterproof capacity of the sensor.

3.2 Temperature Jump and Open Circuit

The main cause of the temperature jump and open circuit of electrical circuit is the structure of the sensor. The entire electrical circuit of the sensor consists of the

platinum resistance, the welding of platinum resistance and cable, the cable, the crimp of the cable, and the pinhole.

In order to avoid the impact of the vibration shock, the strength of the wire and the cable length margin retained in operating conditions are fully considered before assembled; special positioning tools should be used for the crimp of the cable and the pinhole, besides the crimp position is fixed on the frame, where the vibration is smaller than the vibration level of the wheel, so the possibility of failure caused by the crimp is very small. The welding unreliability between the platinum resistance and the wire is the main reason for temperature jump; the welding structure between the platinum resistance and the wire is shown in Fig. 7.

At present, there are two main welding methods. The first one is the tin welding: by melting the tin at high temperature, the tin is wrapped around the platinum resistance wire and the core of wire. The second method is the resistance welding: the core of the wire and of the platinum resistance wire is melted and then melted into a whole. The final result of these two welding methods is spherical balls. As the diameter of platinum wire is 0.2 mm, in ideal conditions, the platinum wire should be in the center of the ball. Because the operation wire is too small, the platinum resistance cannot be fixed by the welding fixture, and using manual welding operation, the welding between the platinum resistance and the wire cannot be performed in the best condition. As shown in Fig. 8a, the platinum resistance wire is welded to the surface of the ball, and Fig. 8b shows the portion of the wire after the platinum resistance wire is disconnected. So far, no manufacturers have been able to provide certification of reliability for platinum resistance welding.

3.3 *The Damage Caused by External Reason*

In addition, installation methods and operating environment factors can also lead to the failures of sensor. For example, the cable of the sensor can be mechanically damaged by abnormal objects, or the sensors are likely to be broken down by a sudden change of the electrical potential between the car body and the track [7], or

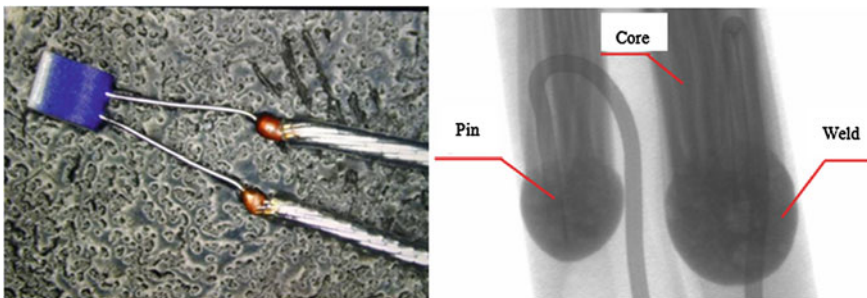


Fig. 7 X-ray photography of platinum resistance and wire welding

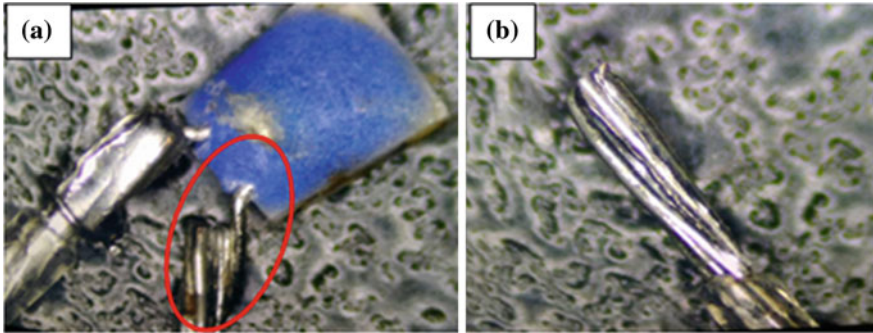
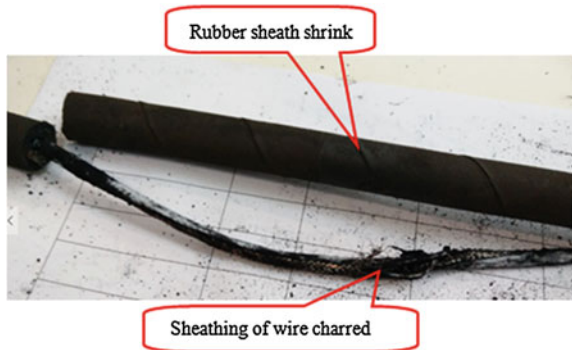


Fig. 8 Structure of the weld between platinum resistance and wire

Fig. 9 Burnt cable for an axlebox temperature sensor



by the protective cover of sensor seal aging and so on. In October 2016, four temperature sensors of a train were identified with failure at the same time.

By testing the axles where the failure sensor is fixed on, the radial wear of the wheels corresponding to the fixation point was high: the wheel has reached the standard reprofiling limit. This means that the vibration exceeds the standard, so the relative movement between the cable and the rubber sheath was exacerbated. A long-time impact caused the breaking of the shield wire; the burr from the shield would hit the wire core inside the insulation layer random. Thus, this may lead to lower dielectric strength, and ultimately the phenomenon of puncture happened between the shield and the wire core. The faulted sensor is shown in Fig. 9.

4 Discussion

Based on the analysis, a conclusion can get that safety risks exist in temperature sensor whatever the structure or installation method. By simply enhance methods, they were not possible to avoid false alarm, circuit breaks, and other failures. Some

institutions have already done research on redundant design of the monitoring of axlebox temperature [8]. There two monitoring methods; the infrared axlebox temperature monitoring equipment is installed inside the vehicle [9]; it is a kind of point monitoring for the low-speed vehicle axlebox temperature; it cannot achieve the continuous uninterrupted monitoring for the temperature of axlebox in the EMU. At present, most of infrared axlebox temperature monitor equipments are applied to the trucks. If we want to install the infrared axlebox temperature monitoring equipment on the EMU line, it is essential not only to do scientific research but also time and huge amounts of money investment are needed. So the infrared axlebox temperature monitoring can only be a subject for scientific research and research.

The second monitoring method is to install a set of emergency wireless axlebox temperature monitoring device on the bogie [10]. When the vehicle axlebox temperature sensor fails, the wireless axlebox temperature monitoring device should be quickly installed to replace the traditional sensor for temperature monitoring. At present, some institutions have already developed a wireless device for axlebox temperature monitoring; the device has the ability to perform real-time monitoring for the axlebox temperature, solving the problem of temperature out of control when the traditional sensor has a failure during the train operating.

5 Conclusion

Based on the analysis of the failure causes of the bogie temperature sensor, this paper presents the unreliability of the traditional structure and the installation method of the sensor on the bogie. The traditional method cannot meet the need of safe operation of EMU, so that some scientific and advanced methods of monitoring such as the redundant design should be used, and preventive measures should be taken to ensure foolproof running and punctuality of the EMU. In order to enhance the bogie temperature monitoring system, it is recommended to install an emergency wireless temperature monitoring device immediately when the traditional temperature sensor has a failure. In that way, the bogie temperature system of the EMU can be always controlled; this program can keep the train running at constant speed, improve the positive rate of EMU, and ensure the absolute safety of the operation.

References

1. Lian N, Jiemin Z (2007) Research on condition monitoring technique for internal temperature of rolling bearings. *Bearing* 2:25–27 (in Chinese)
2. Weigang M, Siyu T (2017) A prediction method based on stepwise regression analysis for train axle temperature. *IEEE Trans Image Process* 26(8):386–390

3. Chuliang W, Zemin C, Hwa-yaw T (2012) Real-time fault diagnosis of train bogie using FBG sensors on intelligent system design and engineering application. In: IEEE (ISDEA) 2012 second international conference, pp 1087–1090
4. Chenyang B, Huabo S (2016) Design of CRH axle temperature alarm based on digital potentiometer. *IEEE Trans Image Process* 25(8):8842–8851
5. Roderich Z, Suleman A, Michael B (2014) Failure mode analysis and optimization of assembled high temperature pressure sensors. In: IEEE thermal, mechanical and multi-physics simulation and experiments in microelectronics and microsystems, international conference on thermal, pp 1–6
6. Hua G, Zhao Haixia (2003) Application of pouring technology to electronic production. *Electr Process Technol* 24(6):257–259 (in Chinese)
7. Qiumin Z, Xinghua W (2014) Sensors burned cause analysis of high-speed EMUs. *Electr Drive Locomotives* 6:57–59 (in Chinese)
8. Jianqin Q, Jia C, Chun Z (2010) Design of a 0.8 V low power CMOS temperature sensor for RFID-based train axle temperature measurement. In: IEEE (ICSICT), 2010 10th IEEE international conference, pp 1404–1406
9. Changbo Z, Lei C (2009) Development and consideration of infrared axle temperature detecting. *Syst Rolling Stock* 47(1):28–30 (in Chinese)
10. Jaehoon K, Lee KS (2011) A study on the wireless onboard monitoring system for railroad vehicle axle bearings using the SAW sensor. In: International conference on sensor systems and software, LNICST, vol 57, pp 52–58

Study on Fault Diagnosis for Bearing Based on VMD-SVD and Extreme Learning Machine

Qiang Zhou, Yong Qin, Zhipeng Wang and Limin Jia

Abstract Bearings are key components in many mechanical facilities, and the research on fault diagnosis for bearing is of great importance to the safe operation of those facilities. Thus, a method for fault diagnosis based on VMD-SVD and extreme learning machine is proposed in this paper. First, the bearing vibration signal is decomposed into a number of stationary intrinsic mode functions (IMF) by VMD method. Second, the initial feature matrix of each IMF component is decomposed by SVD, and the obtained singular value is used as the eigenvector of the signal. Finally, extreme learning machine is used as the classifier for fault diagnosis. This method's feasibility and effectiveness have also been verified by experiment.

Keywords Fault diagnosis · Variation mode decomposition · Singular value decomposition · Extreme learning machine

1 Introduction

The bearing faults might bring about the breakdown of machine, causing enormous economic loss and even casualties [1]. Thus, the study of efficient fault diagnosis method for bearing has great practical significance to the safe operation of rotary machine. Generally, the main parts of bearing fault diagnosis are as follows: the acquisition of vibration signals, the extraction of the fault features, and the recognition and diagnosis of fault types [2].

In terms of fault feature extraction of bearings vibration signals, variation mode decomposition (VMD) proposed by Dragomiretskiy et al. [3] is introduced in this paper. VMD can decompose a signal into a number of IMFs adaptively, and time–frequency divisions of VMD is more refined compared to EMD and LMD [4].

Q. Zhou · Y. Qin · Z. Wang (✉) · L. Jia
State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University,
Beijing, China
e-mail: zpwang@bjtu.edu.cn

In order to extract fault feature further, singular value decomposition (SVD) is also used in this paper.

In terms of condition identification and fault diagnosis, this paper applies extreme learning machine (ELM) to fault diagnosis for bearing. Compared to other learning algorithms, ELM algorithm which is developed based on single-hidden layer feedforward neural network not only avoids falling into local optima but also accelerates the running speed of the identification model [5].

In 2015, Tian et al. proposed a method of rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine [6]. In this study, a new method based on VMD-SVD and ELM is put forward for fault diagnosis. The process of this method is shown in Fig. 1.

2 Methodology

2.1 Variation Mode Decomposition (VMD)

VMD can estimate the modes and determine the correlative bands of fault feature at the same time, which can decompose the signal into several IMFs [7]. VMD is a constrained variational problem represented by the following equation:

$$\min_{\mu_k, \omega_k} \left\{ \sum_k \left\| \partial_t \left[\left(\sigma_t + \frac{j}{\pi t} \right) \mu_k(t) \right] e^{-j t \omega_k} \right\|_2^2 \right\} \text{ Subject to } \sum_k \mu_k = f, \quad (1)$$

where μ_k and ω_k are IMF components and their center frequencies.

To obtain the optimal solution of the problem, introduce the augmented Lagrange function:

$$L(\mu_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\sigma_t + \frac{j}{\pi t} \right) \mu_k(t) \right] e^{-j t \omega_k} \right\|_2^2 + \left\| f - \sum \mu_k \right\|_2^2 + \langle \lambda, f - \sum \mu_k \rangle \quad (2)$$

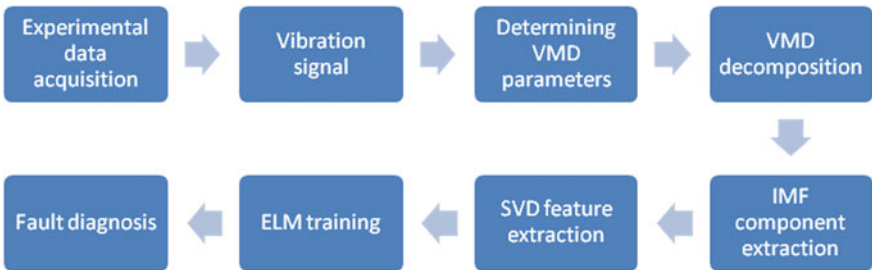


Fig. 1 Flow chart of fault diagnosis based on VMD-SVD and ELM

The mode number k and quadratic penalty α are set in advance, while the sub-mode function μ_k^1 , the center frequency ω_k^1 , and the Lagrangian multiplier λ^1 are initialized [8]. Then, modes μ_k and the center frequency ω_k are renewed, respectively, by Eqs. (3) and (4):

$$\mu_k^{n+1} \leftarrow \frac{\hat{f} - \sum_{i < k} \hat{\mu}_i^{n+1} - \sum_{i > k} \hat{\mu}_i^n + \frac{\hat{\lambda}^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2}, k \in \{1, K\} \quad (3)$$

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega \left| \hat{\mu}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{\mu}_k^{n+1}(\omega) \right|^2 d\omega}, k \in \{1, K\} \quad (4)$$

After the modes and center frequencies are updated, the Lagrangian multiplier λ is also updated by Eq. (5):

$$\lambda^{n+1} = \lambda^n + \tau \left(x - \sum_k \mu_k^{n+1} \right) \quad (5)$$

μ_k , ω_k and λ are updated iteratively until Eq. (6) is satisfied.

$$\sum_k \left\| \mu_k^{n+1} - \mu_k^n \right\|_2^2 / \left\| \mu_k^n \right\|_2^2 < \varepsilon \quad (6)$$

2.2 Singular Value Decomposition (SVD)

SVD is a matrix decomposition and transformation technique, which has been widely used in signal processing and data compression. Assuming that the measured data constructs a $m \times n$ matrix H , H can be decomposed as follows:

$$H = UAV^T = \sum_{i=1} \sigma_i \mu_i \nu_i^T = \sum_{i=1} \sigma_i B_i, \quad (7)$$

where $U \in R^{m \times m}$, $V \in R^{n \times n}$ are orthogonal matrix, $A = \begin{bmatrix} A_r & 0 \\ 0 & 0 \end{bmatrix} \in R^{m \times n}$, $A_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $\sigma_i (i = 1, 2, \dots, r)$ are singular values of H .

2.3 Extreme Learning Machine (ELM)

Extreme learning machine (ELM) is an emerging learning algorithm, which is presented based on single-hidden-layer feedforward neural networks (SLFNs) [9].

The structure model of ELM network is shown in Fig. 2.

Consider a training data set (x_i, y_i) , where $x_i (i = 1, 2, \dots, N)$ is input vector, $y_i (i = 1, 2, \dots, M)$ is output vector. If the classification of ELM with L hidden neurons, the output expression of ELM network is as follows:

$$y_i = \sum_{i=1}^L \beta_i G(\alpha_i, b_i, x_i), \tag{8}$$

where α_i is the weight vector connecting the i th hidden node and the input nodes. b_i is the threshold of the i th hidden node. β_i is the weight vector connecting the i th hidden node and the output nodes.

Equation (8) can be written simply as

$$H\beta = Y \tag{9}$$

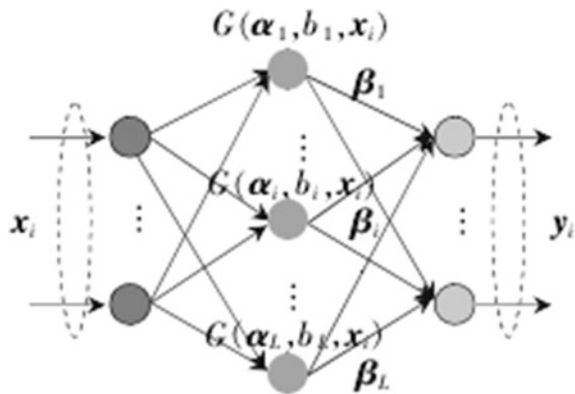
where H is the hidden layer output matrix of the network [10]:

$$H = \begin{bmatrix} G(\alpha_1, b_1, x_1) & \cdots & G(\alpha_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(\alpha_1, b_1, x_N) & \cdots & G(\alpha_L, b_L, x_N) \end{bmatrix}_{N \times L} \tag{10}$$

and, $\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times M}$, $Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times M}$. The least square solution of β is

$$\hat{\beta} = H^+ Y, \tag{11}$$

Fig. 2 The structure model of ELM



where H^+ is the Moore–Penrose pseudo-inverse of matrix H .

Then, the ELM algorithm can be summarized as follows:

1. Set the number of hidden layer neurons L and the activation function G .
2. Randomly assign the input connection weight α_i and the threshold b_i ;
3. Calculate output matrix H of the hidden layer; and
4. Calculate the output connection weight β .

3 Experimental Results

In this paper, the data from Case Western Reserve University Bearing Data Center were used to test and verify this fault diagnosis method.

In the experiment, there were four kinds of states of bearing: normal condition, inner race fault, outer race fault, and rolling element fault. The vibration signals were, respectively, collected in above states by a motor with a constant speed at 1797 r/min, and the sampling frequency of the signals is 12 kHz.

And the number of samples under four kinds of states of bearing is shown in Table 1.

3.1 Signal Decomposition Using VMD

VMD needs to give the preset IMF component number K and penalty parameter α . The value of α takes the default value 2000. In order to obtain the best result of decomposition, the vibration signals were decomposed corresponding to different K values under four kinds of states of bearing, and the mode center frequencies corresponding to different K values are shown in Table 2. From Table 2, under four kinds of states of bearing, it can be clearly seen that the mode components with similar central frequencies begin to appear when $K = 5$: this phenomenon is called over-decomposition. Therefore, the value of K was 4 in this experiment.

Therefore, $K = 4$ and $\alpha = 2000$ were substituted into VMD program to decompose vibration signals. The results of decomposition under four kinds of states of bearing are, respectively, shown in Figs. 3, 4, 5, and 6.

Table 1 Four kinds of datasets

	Normal condition	Inner race fault	Outer race fault	Rolling element fault
Number of samples	24	12	12	12

Table 2 Central frequencies of IMF components

Mode number	Normal condition					
	Center frequency(Hz)					
2	13	169				
3	11	87	175			
4	11	87	175	407		
5	11	56	87	175	408	
6	11	58	87	175	252	406
Mode number	Inner race fault					
	Center frequency(Hz)					
2	57	232				
3	56	229	299			
4	51	110	230	299		
5	51	109	228	280	302	
6	51	109	212	237	282	302
Mode number	Rolling element fault					
	Center frequency(Hz)					
2	8	276				
3	7	242	278			
4	6	117	242	278		
5	6	117	239	272	283	
6	2	49	238	271	280	286
Mode number	Outer race fault					
	Center frequency(Hz)					
2	3	284				
3	3	237	286			
4	2	236	279	295		
5	2	229	245	286	433	
6	2	123	236	279	295	433

3.2 Feature Extraction Using SVD

In order to extract fault feature further, the matrices composed of each IMF component under four kinds of states were decomposed using SVD method.

3.3 Fault Detection and Classification Using ELM

In order to identify the fault type effectively in training and testing, the fault types were given the corresponding class labels in this experiment, as shown in Table 3. And the samples of singular values were divided into two parts: the training samples and testing samples.

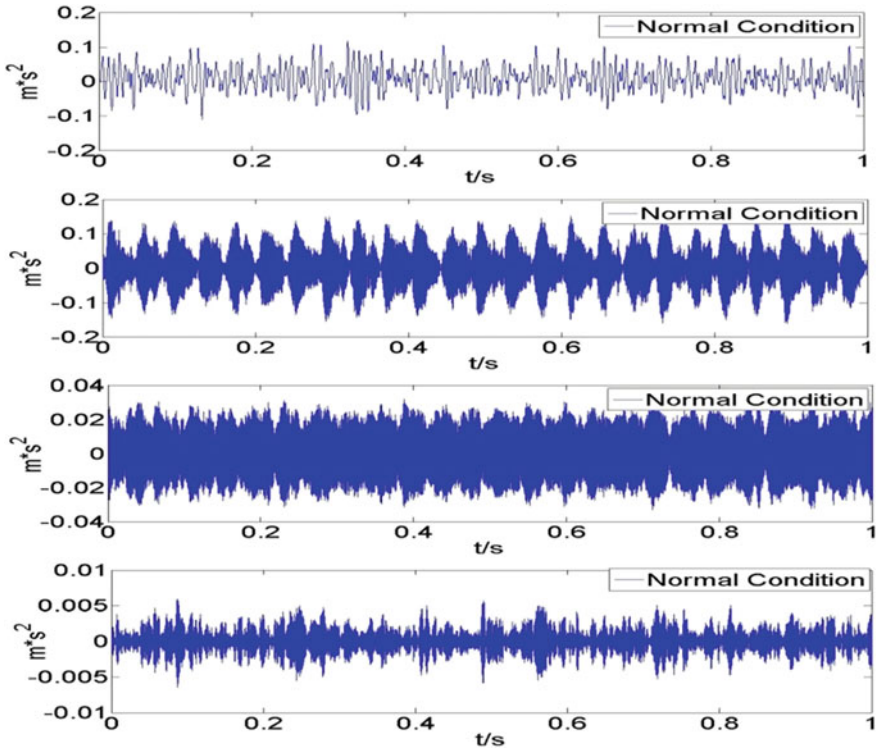


Fig. 3 Decomposition signal under normal condition

The number of hidden neurons and activation function need to be set while training the network of ELM. In order to improve the accuracy of fault identification, the number of hidden neurons was determined to be 10, and the Sigmoidal function was chosen as activation function. The experimental results are shown in Fig. 7. As can be clearly seen from Fig. 7, the curves of the real class and identification class overlap completely. So the classification accuracy is 100%.

For comparative analysis, the models of fault diagnosis based on the support vector machine (SVM) and the BP neural network (BPNN) were built in this paper. The experimental results are shown in Table 4.

According to Table 4, the result of fault diagnosis based on the ELM is significantly superior to the SVM and BPNN. In addition, ELM has notable superiority not only in running time but also in classification accuracy.

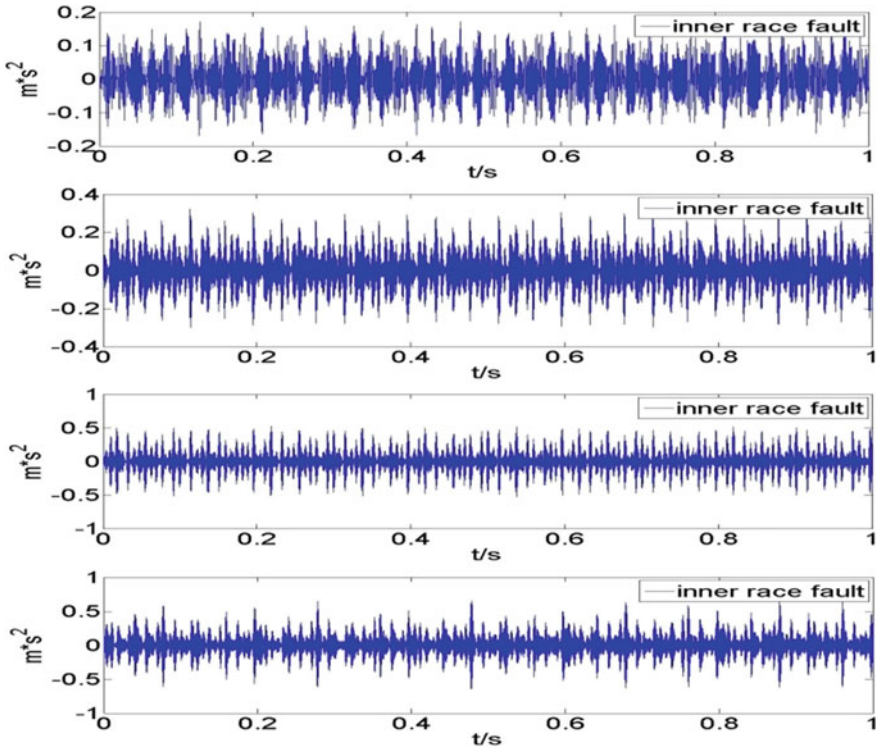


Fig. 4 Decomposition signal under inner race fault

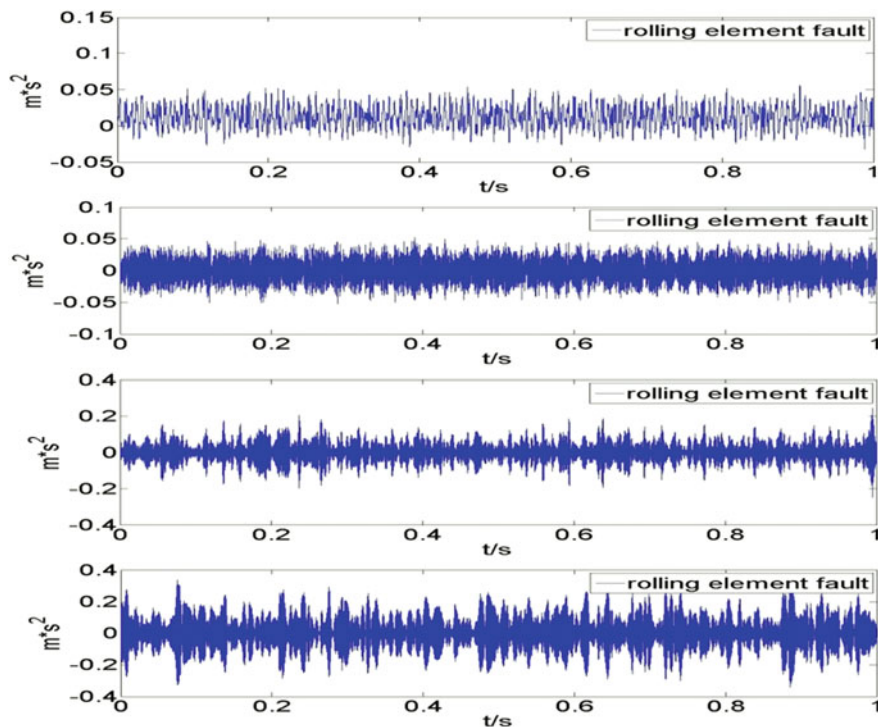


Fig. 5 Decomposition signal under rolling element fault

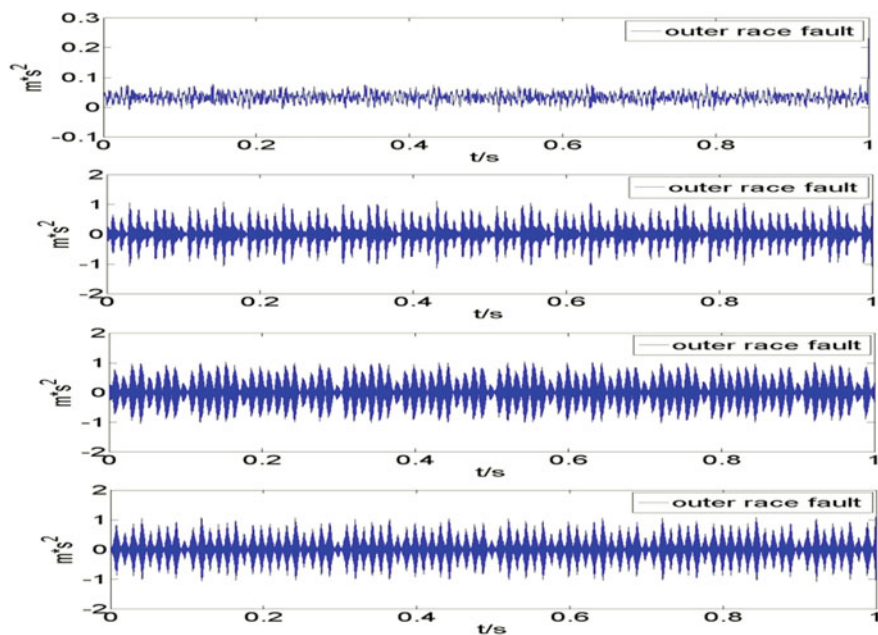


Fig. 6 Decomposition signal under outer race fault

Table 3 Label of fault types

Fault type	Normal condition	Inner race fault	Rolling element fault	Outer race fault
Class labels	1	2	3	4

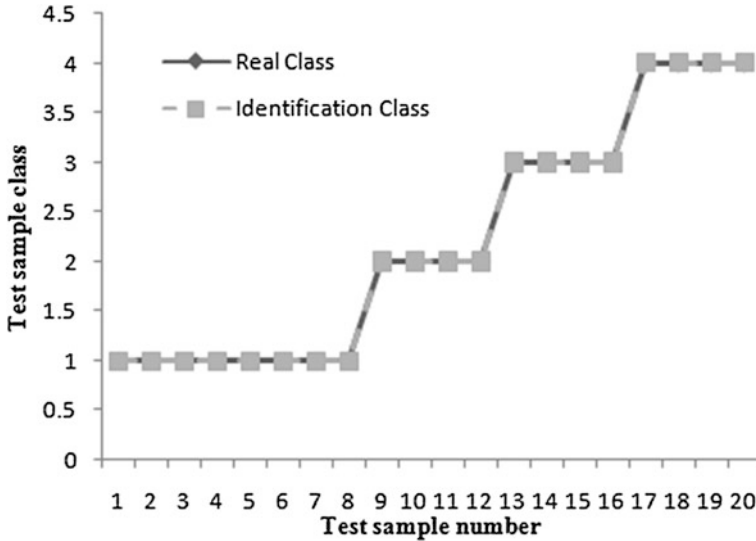


Fig. 7 ELM recognition result

Table 4 Comparison of experimental results

Type of network	The number of training samples	The number of testing samples	Running time	Classification accuracy (%)
ELM	40	20	0.0312	100
SVM	40	20	0.658	90
BPNN	40	20	0.9414	80

4 Conclusion

This paper introduces a fault diagnosis method for bearing based on VMD-SVD and ELM. VMD-SVD is able to extract the fault feature effectively. Then, the fault type of bearing is identified by ELM. Furthermore, ELM is better suited for fault identification compared to SVM and BPNN. It can be seen from the experimental results that VMD-SVD and ELM method presented in this paper can identify the state of bearing quickly and accurately.

Acknowledgements This research is supported by the National Key Research and Development Programs of China (Nos. 2016YFB1200203 and 2016YFB1200402), as well as the State Key

Laboratory of Rail Traffic Control and Safety (Contract Nos. RCS2016ZQ003 and RCS2016ZT018), Beijing Jiaotong University.

References

1. Liu X, Bo L, He X, Veidt M (2012) Application of correlation matching for automatic bearing fault diagnosis. *J Sound Vib* 331:5838–5852
2. Liu W, Han J, Jiang J (2013) A novel ball bearing fault diagnosis approach based on auto term window method. *Measurement* 46(10):4032–4037
3. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Trans Signal Process* 62(3):531–544
4. Wang Y, Markert R, Xiang J et al (2015) Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mech Syst Signal Process* 60:243–251
5. Yuan Q, Zhou W, Liu Y, Wang J (2012) Epileptic seizure detection with linear and nonlinear features. *Epilepsy Behav* 24(4):415–421
6. Tian Y, Ma J, Lu C, Wang Z (2015) Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine. *Mech Mach Theory* 90:175–186
7. Wang YX, Markert R, Xiang JW, Zheng WG (2015) Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mech Syst Signal Process* 60–61:243–251
8. Li Z, Chen J, Zi Y et al (2017) Independence-oriented VMD to identify fault feature for wheel set bearing fault diagnosis of high speed locomotive. *Mech Syst Signal Process* 85:512–529
9. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. *Neurocomputing* 70:489–501
10. Huang GB (2003) Learning capability and storage capacity of two hidden-layer feedforward networks. *IEEE Trans Neural Netw* 14(2):274–281

Optimization of Wheel Re-profiling Strategy Based on a Statistical Wear Model

Bingkui Li, Zhi Yang, Zongyi Xing and Xudong Gao

Abstract Aiming at improving the current unreasonable wheelset re-profiling schedule caused by class-lathing, this paper proposes a re-profiling strategy based on historical wear data. First, the statistical product and service solutions (SPSS) software is used to analyze the interdependency between wheel diameter wear and wheel flange thickness based on the wheel wear data of Guangzhou Metro Line 8. Second, flange thickness wear model based on state transition and wheel diameter wear model based on mathematical statistics are built, and single-stage and multistage planned turns strategies of an individual wheel are built on the two models. Finally, Monte Carlo simulation is used to compare two strategies, and the results show that multistage planned turns strategies can prolong expected life of wheels and effectively save operating cost.

Keywords Wheel wear · Optimization · Single-stage planned turns strategy
Multistage planned turns strategy

1 Introduction

With the continuous development of rail transit, wheel wear has always been the concern of domestic and foreign scholars. Ansari et al. [1] established a wheel wear model based on the theory of multibody dynamics and verified the feasibility of model by analyzing the wheel wear data of Tehran subway vehicle. Arizon et al. [2] analyzed wheel wear by means of analyzing the model to achieve wheel wear failure modeling and life prediction. Through analyzing Wuhan-Guangzhou high-speed EMU tread wear data, Huang [3] and Diao [4] found that tread wear increases with the increase in wheel thickness wear. The above researchers have

B. Li · Z. Yang · Z. Xing (✉) · X. Gao
Nanjing University of Science and Technology, No. 200 Xiao Lin Wei,
Xuanwu District, Nanjing 210094, China
e-mail: xingzongyi@163.com

done lots of research from the dynamic model and actual wear data and so on, laying the foundation for the optimization of wheel repair.

Tao et al. [5] established wear model based on mathematical statistics after the analysis of D20E locomotive wheel wear data and achieved different types of locomotive wheel cutting repair cycle prediction. Sun [6] analyzed the feasibility of class-lathing; he believed that different lathing class was chosen based on wheel flange width. The above methods are based on the statistical analysis of wheel wear and give some recommendations, but they do not touch the essence of optimization of wheel re-profiling strategy.

Due to the difference of train model and operation line, the wheel re-profiling scheme cannot meet the actual application based on the dynamic performance analysis. However, it is effective and feasible to formulate or adjust the maintenance cycle of wheel re-profiling by analyzing the train wheel wear data of actual operation line. Therefore, this paper proposes a multistage wheel re-profiling control strategy model to meet the actual needs of wheel maintenance based on the wear data of Guangzhou Metro Line 8.

2 Correlation of Wheel Wear and Wheel Flange Thickness

Xu [7] and Wang [8] shown that the wheel wear rate of rail transit is related to wheel flange thickness, but the trend and correlation of wear rate and wheel flange thickness are different. Therefore, it is necessary to analyze the correlation of wheel wear and wheel flange wear before establishing the wheel wear model. In this paper, the wheelset size data of Guangzhou Metro Line 8 are taken as sample data. Wheel wear rate and wheel flange wear rate of per month are calculated by formulas (1) and (2).

$$v_{d,i} = \frac{D_{i+1} - D_i}{t_{i+1} - t_i} \cdot 30 \quad (1)$$

$$v_{s_d,i} = \frac{S_{d,i+1} - S_{d,i}}{t_{i+1} - t_i} \cdot 30 \quad (2)$$

where D_{i+1} and D_i , respectively, are the wheel diameter measurement data of t_{i+1} and t_i moments without wheel Re-profiling. $S_{d,i+1}$ and $S_{d,i}$, respectively, are the wheel flange thickness data of t_{i+1} and t_i moments without wheel Re-profiling. $v_{d,i}$ and $v_{s_d,i}$, respectively, denote wheel wear rate and wheel flange wear rate for each time period.

The SPSS is used to analyze the interdependency between wheel diameter wear rate, wheel flange wear rate, and wheel flange thickness. By analyzing these data, the Pearson correlation of wheel wear rate and wheel flange thickness is -0.05 . But the Pearson correlation of wheel flange wear rate and wheel flange thickness was

-0.286. Therefore, it can be considered that wheel flange wear rate is related to wheel flange thickness, and wheel diameter wear rate is independent of wheel flange thickness.

3 Wheel Flange and Wheel Diameter Wear Model

3.1 Wheel Flange Wear Model Based on State Transition Process

According to the safety limit of wheelset size: $26 \text{ mm} \leq S_d \leq 32 \text{ mm}$ [9]. The safety field is divided into 20 state intervals. They are, respectively (31.7,32], (31.4,31.7], (30.1,31.4], ..., (26.3,26.6], [26,26.3], denoted as S_1, S_2, \dots, S_{20} . Due to the size difference of wheel wear, the wheel flange thickness value may remain in the original state interval, denoted as $S_{(i,i)}$, and it may be transferred to another state interval, denoted as $S_{(i,i+m)}$. The state transition diagram is shown in Fig. 1.

The sample data are calculated according to the state transition mode; thus, the state transition probability of each initial state is obtained. The state transition situation is mainly concentrated in state stay, state a transition, and state secondary transition. However, state secondary transfer is relatively stable. So, this paper mainly analyzes the state stay and state a transfer, as shown in Fig. 2.

According to Fig. 2, the probability of state stay has a tendency to increase at first and then decrease with the decrease in wheel flange thickness. In the initial state S_6-S_9 , wear rate decreases as the state stay probability increases. The probability of secondary transition is basically stable, so the change rate of state a transfer is opposite to state stay. By calculating state transition probability of each state in the initial state, the possibility that each initial state corresponds to the three kinds of transition is obtained in the wheel flange wear process. The probability of each transition is determined by the statistical probability and satisfies the uniform distribution.

Fig. 1 Schematic diagram of state transition

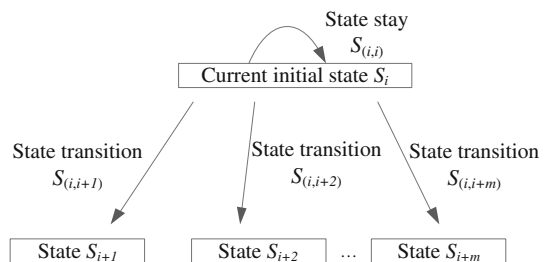
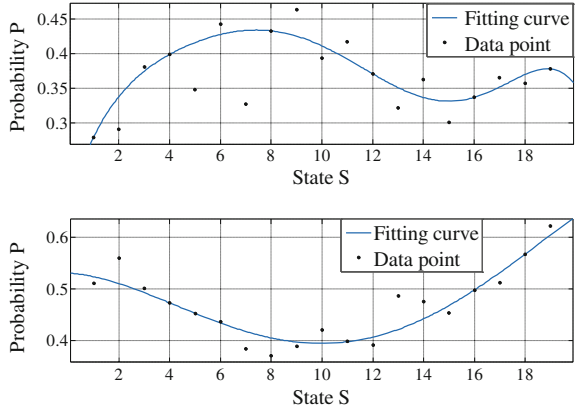


Fig. 2 The change trend of wheel flange wear and wheel thickness: (top) state stay and (bottom) state a transition

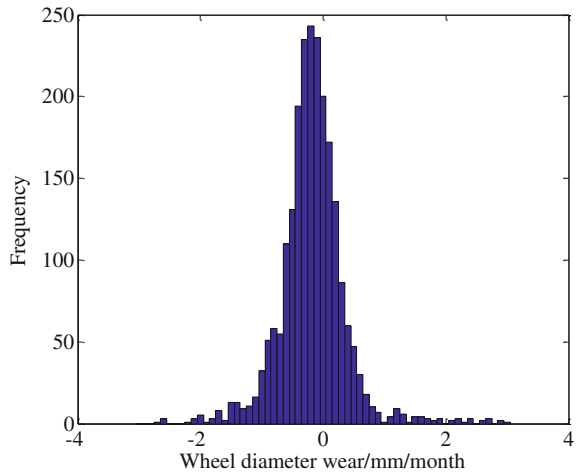


3.2 Wheel Wear Model Based on Mathematical Statistics Theory

According to the wheel wear data of Guangzhou Metro Line 8, wheel wear rate is basically -3 to 3 mm/month. For the convenience of statistical analysis, the interval $[-3, 3]$ is divided into 60 small intervals. The frequency of wheel diameter wear rate is counted in each interval, and the statistical analysis of wheel diameter wear rate is shown in Fig. 3. Among them, the frequency of wheel diameter wear rate mainly focuses on the -0.3 to 0.1 mm/month.

In order to obtain the rule of wheel diameter wear, wheel diameter wear is modeled by normal distribution, logarithmic normal distribution, and Γ distribution. And the specific parameters of three distributive probability density functions can be fitted and calculated by Matlab.

Fig. 3 Frequency histogram of wheel diameter wear rate



Normal distribution probability density function:

$$f_1(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

where $\mu = -0.1733$ and $\sigma = 0.5438$.

Logarithmic normal distribution probability density function:

$$f_2(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma(x+3)}} e^{-\frac{[\ln(x+3)-\mu]^2}{2\sigma^2}}, & x + 3 \geq 0 \\ 0, & x + 3 < 0 \end{cases} \tag{4}$$

where $\mu = 1.0188$ and $\sigma = 0.2134$.

Γ distribution probability density function:

$$f_3(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} (x+3)^{\alpha-1} e^{-\beta(x+3)}, & x + 3 \geq 0 \\ 0, & x + 3 < 0 \end{cases} \tag{5}$$

where $\alpha = 24.8086$ and $\beta = 0.1139$.

This paper uses χ^2 to test the goodness of three distribution models, the number of fitting goodness test k is 60, the significance level α is 0.05, and unknown parameter r is 0. And the critical value can be obtained by referring to the χ^2 distribution, $\chi^2_\alpha(k - r - 1) = \chi^2_{0.05}(59) = 77.931$. Table 1 shows test results.

According to Table 1 and Fig. 3, it can be seen that the normal distribution is more consistent with the wheel diameter wear distribution.

4 Class Planned Turns Strategy Model of an Individual Wheel

4.1 Single-Stage Planned Turns Strategy Model

The concept of state transfer is proposed and the safety threshold of wheel flange thickness is divided when wheel flange wear model is established. The following description is made: When wheel flange thickness is worn to a certain state S_j , then restored it to a state of S_i . S_i corresponds to the control upper limit state, denoted as Sh . S_j corresponds to the control lower limit state, denoted as Sl . The control upper and lower limit states correspond to the respective interval ranges, and the interval

Table 1 Test results of fitting goodness

Distribution type	χ^2 statistical value	Result comparison	Decision
Normal distribution	31.314	$\chi^2 < \chi^2_{0.05}(59)$	Accept
Logarithmic normal distribution	233.793	$\chi^2 > \chi^2_{0.05}(59)$	Refuse
Γ distribution	64.788	$\chi^2 < \chi^2_{0.05}(59)$	Accept

is expressed as $[Sh_r, Sh_l]$ and $[Sl_r, Sl_l]$, respectively. The schematic diagram of single-stage planned turns strategy is shown in Fig. 4. The wheel flange follows the state transfer probability table in the wear process.

According to the principle of wheel single-stage re-profiling control limit, there are 210 kinds of strategy schemes. In this paper, the principle of making the control strategy is to ensure that wheel flange thickness is maintained within a safety threshold of 26–32 mm, and the optimal control limit strategy is obtained to prolong service life of the wheel and ensure the smaller wheel re-profiling times. Therefore, wheel life and re-profiling times are served as the optimization target, after Monte Carlo simulation, life expectancy and expected re-profiling times of each strategy are obtained.

The steps of single-stage re-profiling Monte Carlo simulation are as follows:

Step 1. Parameter initialization: the initial value of wheel flange thickness and wheel wear size are, respectively, $D(0) = 840$ mm, $S_d(0) = 32$ mm. The cumulative life cycle $T = 0$, re-profiling times $N = 0$, re-profiling interval time $t = 0$, and the control upper and lower limit states are Sh and Sl , respectively. The corresponding re-profiling recovery upper and lower limits are Sh_l and Sl_l , respectively. The 210 policies corresponding values will be imported here.

Step 2. Calculating the wear cycle $t = t + 1$.

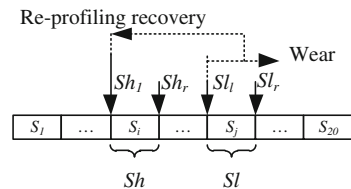
Step 3. According to the wheel diameter value $D(T + t - 1)$ and the wheel flange thickness value $S_d(T + t - 1)$ at $T + t - 1$ moment, the normal distribution of wheel diameter wear randomly generates wheel diameter wear probability p_D , and state transition uniform distribution randomly generates the state transition probability p_{S_d} . According to the probability of wear, wheel flange wear rate v_d and wheel diameter wear rate v_{S_d} were obtained, $D(T + t) = D(T + t - 1) - v_d$ and $S_d(T + t) = S_d(T + t - 1) - v_{S_d}$ is used to calculate the current wheel diameter and wheel flange thickness value.

Step 4. If the current wheel diameter value $D(T + t)$ is greater than or equal to 770 mm, go to step 5, otherwise the end of the simulation, get the final life cycle T_{end} and re-profiling number N_{end} .

Step 5. If the current wheel flange thickness value $S_d(T + t)$ is less than Sl_l , go to step 6, otherwise go to step 2.

Step 6. Re-profiling wheel, recovery flange thickness to Sh_l , and initializing the parameters: $T = T + t$, $N = N + 1$, $t = 0$, $D(T) = D(T) - k * (Sh_l - S_d(T))$, $S_d(T) = Sh_l$, then enter into step 2. Among them, k is the re-profiling proportional coefficient. According to the study of Pang [10] and Zhao [11], the re-profiling proportion coefficient is 4.2.

Fig. 4 Schematic diagram of single-stage planned turns strategy



4.2 Multistage Planned Turns Strategy Model

In the re-profiling process, single-stage control limit strategy reduces the wheel life at the expense of the larger wheel diameter, and single-stage control limit strategy is only for a single wheel. Therefore, this paper puts forward multistage planned turns strategy model in order to prolong service life of the train and meet the actual needs of re-profiling.

The multistage re-profiling control limit has a number of upper and lower limits, and they appear in pairs. The multilevel control limit strategy is shown in Fig. 5.

In the multistage control limit model, second-stage control strategies are 1140 and third-stage control strategies are 3876. In order to simplify work and the effectiveness of test method, Monte Carlo simulation of second-stage control limit strategies is taken as an example. Compared with the simulation of the single-level control strategy, the second-stage re-profiling limit strategy initialization need set the upper and lower limits. The wheel flange thickness threshold range of second-stage control strategy is judged, that is, wheels are repaired in first stage if the current wheel flange thickness satisfies primary re-profiling condition, wheels are repaired in second-stage if the current wheel flange thickness satisfies second-stage re-profiling condition.

5 Simulation

Monte Carlo simulation is used to analyze single-stage and second-stage re-profiling control limits, each scheme sets 1000 times. After averaging, single-stage control strategy simulation results are shown in Fig. 6. When the lower limit of the scheme is same, life expectancy of wheel increases first and then decreases with the increase in the upper limit, which is in accordance with the phenomenon that the probability of state stay increases first and then decreases with the thickness of flange. And when the lower limit of the scheme is same, re-profiling time expectancy of wheel decreases with the increase in re-profiling upper limit.

According to Fig. 2, control limit scheme numbers 749–832 are analyzed. The left value of first-stage control lower limit is 29.9, the left value of secondary-stage

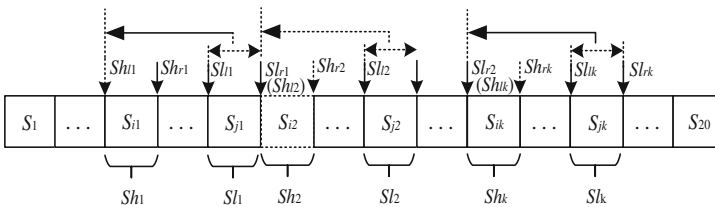


Fig. 5 Multistage wheel re-profiling control limit diagram

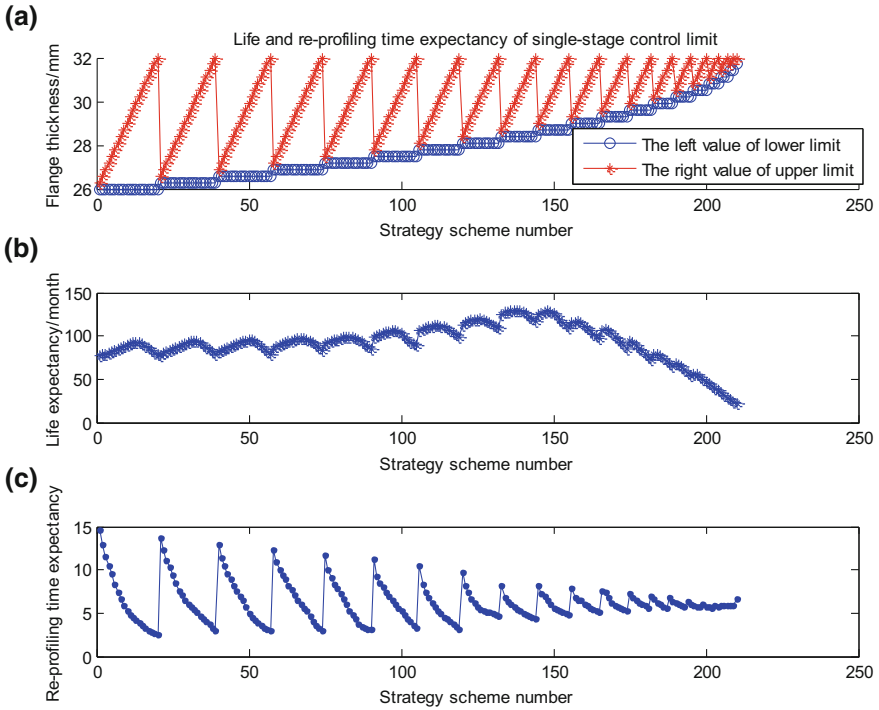


Fig. 6 The simulation result diagram of single-stage re-profiling control limit: **a** the 210 strategic scheme number, **b** life expectancy, and **c** re-profiling time expectancy

control upper limit is 29.6. The secondary-stage control strategy simulation results are shown in Fig. 7. When the upper and lower limits of secondary control are same, life expectancy of wheel increases first and then decreases, and re-profiling time expectancy of wheel decreases with the increase in first control upper limit. And with the increase in secondary control lower limit, life expectancy of wheel increases first and then decreases, and re-profiling times increase.

Compared with single-stage control strategy, the advantages of multistage control strategy are mainly reflected in the following two aspects.

From the perspective of operating cost, the wheel life expectancy of multistage control strategy is longer than single-level control strategy, so it can more effectively save the operating cost of the train.

From the view of practical application, although it is not obvious to improve re-profiling times for multistage control strategy, and re-profiling time expectancy also increased with the increase in the secondary control lower limit, it increases the flexibility of re-profiling times, and it is easy to apply to actual operation.

Four better schemes are selected for secondary-stage control strategy, as shown in Table 2.

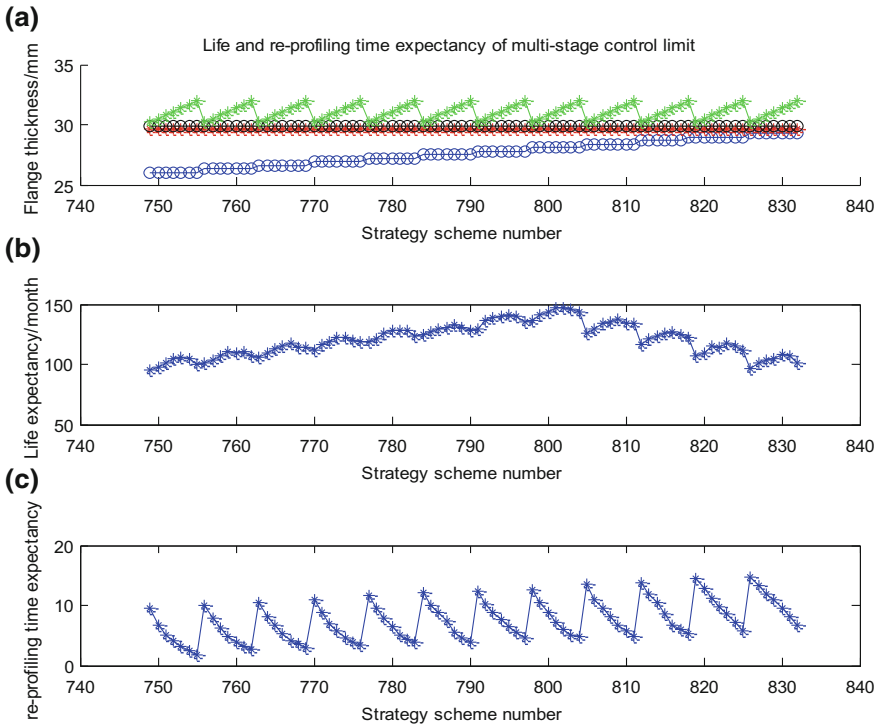


Fig. 7 Simulation result diagram of secondary-stage control strategy: **a** number 749–832 strategic schemes, **b** life expectancy, and **c** re-profiling time expectancy

Table 2 The optimal scheme of secondary-stage control strategy

Single-stage control limit	Secondary-stage control limit	Life expectancy	Re-profiling time expectancy
[29.9,31.4]	[28.1,29.6]	147	6.5
[29.9,31.7]	[28.1,29.6]	142	5.5
[29.9,31.7]	[27.8,29.6]	136	5.2
[29.9,31.1]	[27.5,29.6]	130	4.6

6 Conclusion

This paper proposes a re-profiling strategy based on the historical wear data, and wheel flange wear model based on state transfer process and wheel diameter wear model based on mathematical statistics are established. On this basis, single-stage and multistage planned turns strategy model are established. Monte Carlo simulation is used to get multistage planned turns strategy model, which can effectively improve wheel life and save train operating cost. This paper only studies

secondary-stage re-profiling strategy. It does not carry on the simulation analysis of third-stage re-profiling strategy. And it does not check whether there is the superior multistage re-profiling control limit strategy.

Acknowledgements This work is supported by national key R&D program of China (No. 2017YFB 1201102).

References

1. Ansari M, Hazrati IA, Esmailzadeh E et al (2008) Wear rate estimation of train wheels using dynamic simulations and field measurements. *Veh Syst Dyn* 46(8):739–759
2. De Arizon J, Verlinden O, Dehombreux P (2007) Prediction of wheel wear in urban railway transport: comparison of existing models. *Veh Syst Dyn* 45(9):849–866
3. Huang Y (2012) Wear law analysis of Wuhan-Guangzhou train line. *Gansu Sci Technol* 28(18):48–49 (in Chinese)
4. Diao X, Zhu S, Dong X (2013) WuGuang lines EMU's wheel wear and vibration performance tracking and research. *Railway Locomotive Car* 02:1–6 (in Chinese)
5. Chien TV, Li F, Qi Z, Ding J (2015) Processing method of locomotive wheel wear statistical data and prediction model of turning period. *J China Railway Soc Railway Eng* 12:14–19 (in Chinese)
6. Sun X (2016) Economic analysis of wheel lathing in metro vehicles. *Urban Mass Transit* 19(5):91–92 (in Chinese)
7. Xu H, Yuan H, Wang L et al (2010) Modeling of metro wheel wear and optimization of the wheel re-profiling strategy based on gaussian processes. *J Mech Eng* 46(24):88–95 (in Chinese)
8. Wang L, Yuan H, Na W et al (2011) Optimization of the re-profiling strategy and remaining useful life prediction of wheels based on a data-driven wear model. *Syst Eng Theory Practice* 31(6):1143–1152 (in Chinese)
9. Wang Z (2013) Study on the model of the re-profiling strategy optimization about the whole wheelset wearing electric multiple unit. Chengdu: Southwest Jiaotong University (in Chinese)
10. Pang S (2016) Research on wheel maintenance strategy of CRH2 EMU. Beijing Jiaotong University, Beijing (in Chinese)
11. Zhao W (2014) Prediction of metro wheel wear and optimization of the wheel re-profiling strategy. Hangzhou: China Jiliang University (in Chinese)

Research on Bearing Fault Diagnosis Based on Cyclic Statistics

Dong Yan and Xiukun Wei

Abstract Bearing is an important component of equipment. Different types of failures lead to different changes of the entire equipment and operating parameters. In the urban rail train operation, the quality and status of bearings have a significant impact on traffic safety. In rotary machinery fault diagnosis, second-order cycle statistics has become an important signal analysis tool. In this paper, the early fault point is realized by CUSUM combined with spectral correlation density function, and the characteristics of different fault are analyzed using second-order cyclic statistics. The research of the theory of cyclic stationary provides the direction and power for the development of bearing fault diagnosis with a case.

Keywords Bearing · Fault point recognition · Fault diagnosis
Cyclic statistics

1 Introduction

Bearings are the most widely used components of rotating machinery and railway vehicle, and are the most prone to failure. When the bearing damage occurs, the damage in the operation of parts will be in contact with the parts of the periodic pulse force, making the observed vibration signal contains periodic components.

Fourier transform has been extensively studied in fault diagnosis [1]. Literature [2] combines the resonant demodulation technique and the time-domain characteristic value to analyze the bearing failure. Wavelet transform uses the compressed or extended Gaussian function to obtain the window function to analyze the signal, with diversity, information sensitivity, and information extraction reproducibility [3]. In the previous research on the fault diagnosis method of bearing, the periodic time-varying characteristics of the vibration signal are not fully considered and

D. Yan · X. Wei (✉)

State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing, China
e-mail: xkwei@bjtu.edu.cn

utilized. The cyclostationary uses the periodic time-varying characteristics of the vibration signal to analyze the operating state of the bearing, and reflects the true condition, revealing the physical nature of the fault. In this paper, cumulative sum (CUSUM) error cumulative graph and the spectral correlation density function are combined to determine the early fault point of the bearing based on the error accumulated point [4, 5].

The rest of this paper is organized as follows. Section 2 introduces the basics of second-order cycle statistics in detail. In Sect. 3, the early fault point of the bearing is identified. In Sect. 4, the fault characteristics of the bearing are analyzed based on the spectral correlation density combination and other methods are compared with this method. The conclusions are drawn in Sect. 5.

2 Cyclical Stationary Theory

2.1 Cyclic Autocorrelation

The time-dependent autocorrelation function of signal $x(t)$ is [5]

$$R_x(t, \tau) = E\left\{x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)\right\} \quad (1)$$

where τ is the delay factor, $E\{\bullet\}$ is statistical average, and $*$ is complex conjugate. If the signal period is T , the sample can be used instead of the statistical average, time-varying autocorrelation as follows:

$$R_x(t, \tau) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x\left(t + nT + \frac{\tau}{2}\right)x^*\left(t + nT - \frac{\tau}{2}\right) \quad (2)$$

The formula (3) is developed with a Fourier series the coefficient as follows:

$$R_x(\tau, \alpha) = \frac{1}{T} \int_{-T/2}^{T/2} R_x(t, \tau) e^{-j2\pi\alpha t} dt \quad (3)$$

where $\alpha = m/T (m \in Z)$ is cycle frequency, put (1) into the formula (3), cycle autocorrelation function is

$$R(\tau, \alpha) = \langle x(t + \tau/2)x^*(t - \tau/2)e^{-j2\pi\alpha t} \rangle_t \quad (4)$$

2.2 Spectral Correlation Density Function

Spectral correlation density (SCD) function is one of the main cycle statistics, which is the Fourier transform of the cyclic autocorrelation function with respect to the delay [5], expressed as

$$S(f, \alpha) = \int_{-\infty}^{\infty} R_x(\tau, \alpha) e^{-j2\pi f\tau} d\tau \quad (5)$$

where f is spectral frequency.

2.3 Spectral Correlation Density (SCD) Function Based on CUSUM

As a test point of the control algorithm, CUSUM is mainly focused on traffic monitoring, stock analysis, and to ensure products and service specifications and other fields. By summing up the small offset of the process in order to achieve amplification, to improve the sensitivity of the small offset test and realize good effect on small fluctuation detection [6].

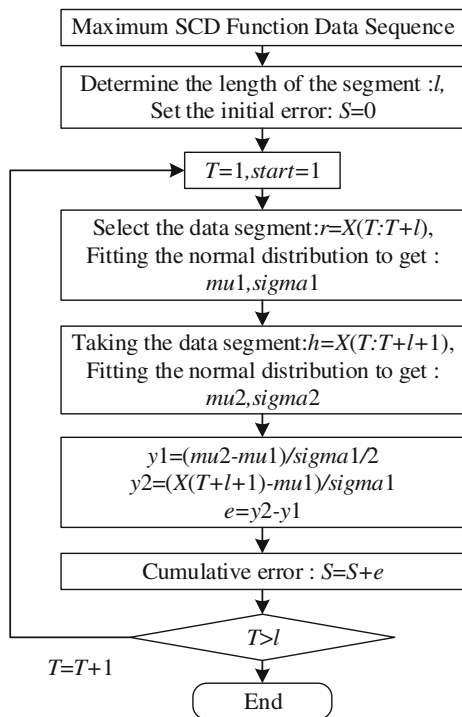
When the fault is detected by the time-domain eigenvalue, eigenvalue is easily submerged in the noise and can only be detected when the fault reaches a certain degree. In order to determine the fault point of the bearing earlier, the spectral correlation density function is used as the basis for measuring the fault, and the CUSUM method is used to detect the early fault time. When the maximum value of the spectral correlation density function changes in the waveform, the density distribution of the random probability also changes, and the fault time can be judged according to the sudden point of the curve. The CUSUM control chart flow in conjunction with the spectral correlation density function is shown in Fig. 1.

3 Case Analysis

3.1 Early Fault Point Identification

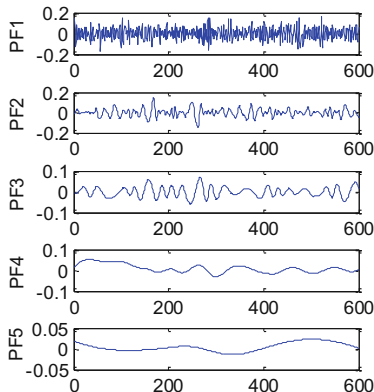
The outer race fault data which is from the bearing vibration test platform of Cincinnati University is taken as an example to verify the validity of the algorithm. Each bearing has 16 rollers. And the pitch diameter is 2.815 in., ball diameter is 0.331 in., contact angle is 15.17°, the rotation speed is 2000 r/min, sampling frequency is 20 kHz, and sampling time interval is 10 min [7].

Fig. 1 Processing of spectral correlation density function based on CUSUM



The local mean decomposition (LMD) is used to decompose the bearing signal, and then the single component can be selected by the largest energy proportion; as shown in Fig. 2, the signal has been decomposed to 5 PF components, and the first PF has the biggest energy ratio, which is shown in Fig. 3. Then, calculate the characteristic values of the first PF, the characteristic values are shown in Fig. 4. It can be seen that the moment of small fluctuations in the four values is different, and at 118 h, there is large fluctuation in all the values. Indicating that the bearing has been in a more serious fault state.

Fig. 2 Signal LMD decomposition



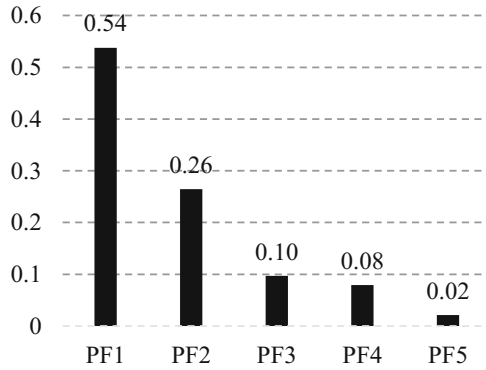


Fig. 3 Energy ratio

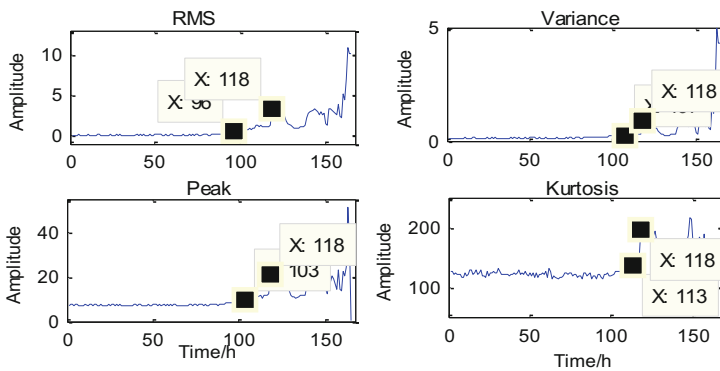


Fig. 4 Time-domain characteristic values

In order to judge the early fault time of the bearing, the spectral correlation density function is calculated at the characteristic frequency in the period of 0–118 h. The characteristic frequency is shown in Table 1, and the fluctuation in SCD value of different frequencies is shown in Fig. 5. Combined with the CUSUM theory, the cumulative error of SCD is calculated, as shown in Fig. 6. It is shown that 89.5 h is the sudden point of the curve error accumulated value. This cumulative error is reduced by 89.67 and 89.5 h, it is said that 89.67 h is the early failure time.

In the literature [8], the early fault point of the same data that is identified by the wavelet entropy method is 96.7 h. Therefore, the method in this paper can detect the fault time early.

Table 1 Characteristic frequency of bearing

Characteristic frequency	Outer race frequency	Inner race frequency	Frequency	Cage frequency	Rotation frequency
Frequency value (Hz)	236.4	296.9	139.9	14.8	33.3

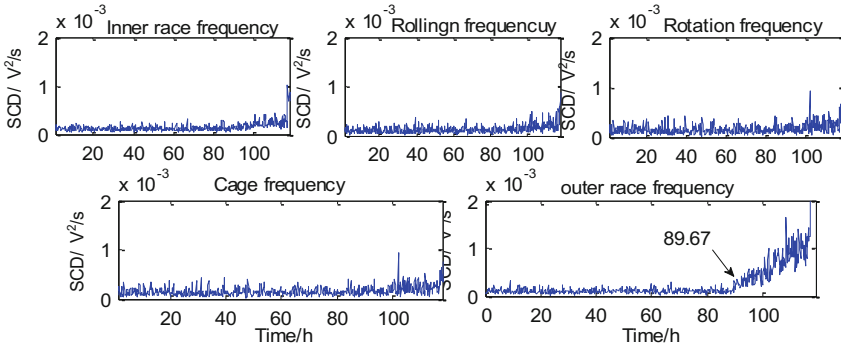


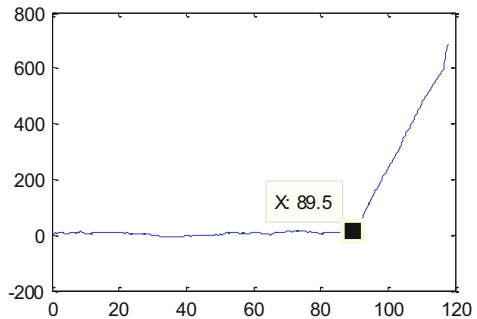
Fig. 5 Life time SSCD of characteristic frequency

3.2 Fault Detection Based on Spectral Correlation Density Combined Slice

The data at 89.67 h is analyzed. In order to reduce the computational complexity of the spectral density function, using the slice analysis method to carry out the bearing fault detection.

The actual characteristic frequency of the bearing signal has some deviation from the frequency of theoretical formula. Therefore, by choosing fault frequency as P_{cen} , the frequency range is set as $[P_{cen}-3 \text{ Hz}, P_{cen}+3 \text{ Hz}]$, and the interval is 1 Hz. The SSCD values of each fault frequency are calculated, and they are shown in Fig. 7. Figure 7f is the combination of the highest value from Fig. 7a to e. Figure 7f shows that the SSCD value of the outer race fault characteristic frequency is higher than others. It is determined that the bearing outer race has failed.

Fig. 6 Accumulation of SSCD by CUSUM



In order to verify the anti-noise ability of the combination of spectral correlation density, this method is used to diagnose fault signal with noises. Taken the SSCD frequency of rolling fault as example, the results are shown in Fig. 8. From the slice diagram, it can be seen that the spectral correlation density function is blurred due to the influence of noise, and the result of the combined is same as the result of without noisy. Therefore, it is shown that the spectral correlation density slicing method is suitable for the weak fault diagnosis when it is noisy.

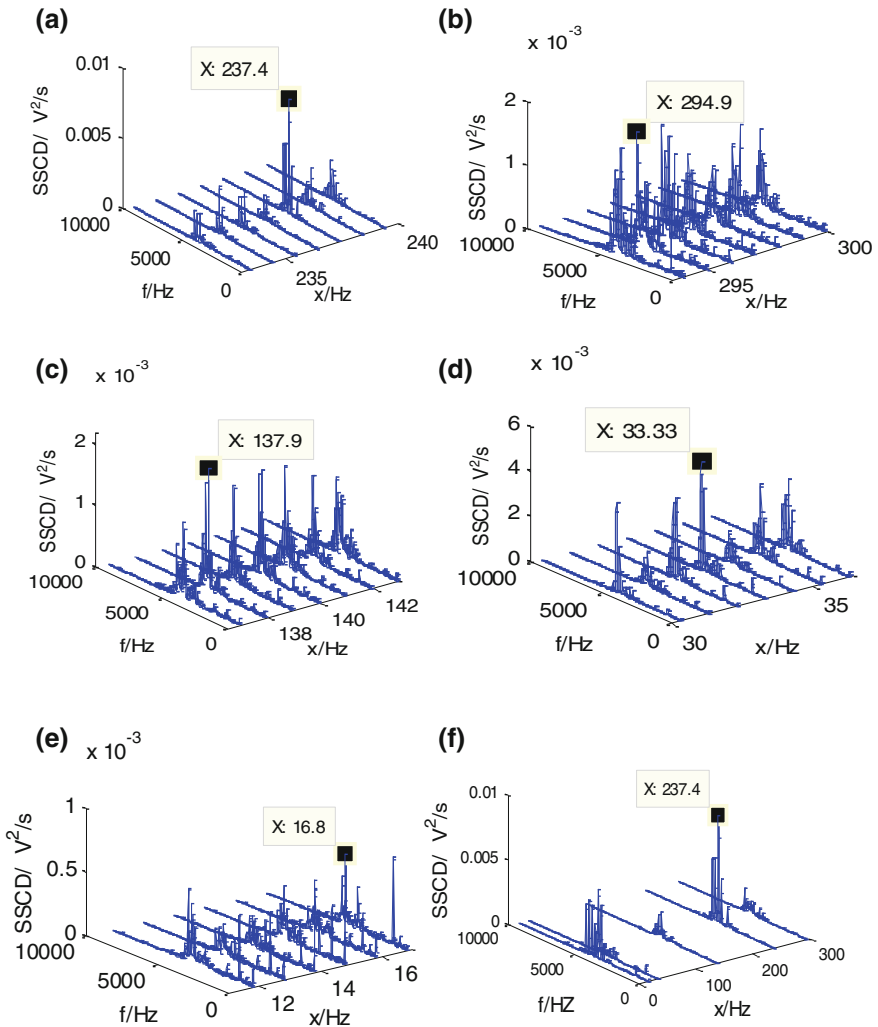


Fig. 7 According to the order, the pictures are as follows: **a** SSCD frequency of outer fault, **b** SSCD frequency of inner fault, **c** SSCD frequency of rolling fault, **d** SSCD frequency of rotation, and **e** SSCD of cage fault **f** SSCD combination

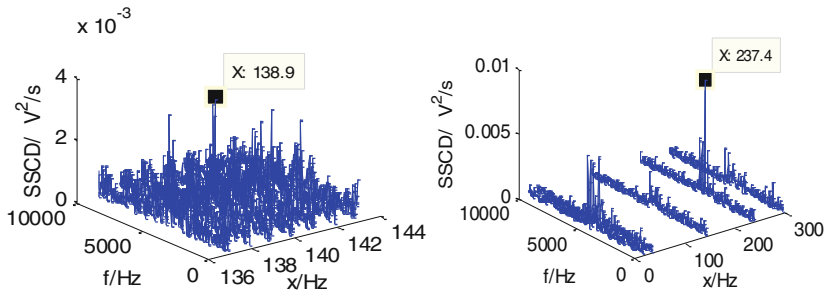


Fig. 8 a SSSCD frequency of cage fault and b SSSCD combination

4 Conclusions

Based on the basic knowledge of cyclic stationary theory, this paper proposes a CUSUM cumulative error method based on spectral correlation density function to identify early fault points. Compared with other methods using the same group of bearing data, it is shown that the spectral correlation density function can detect weak faults in advance. Based on the combination of spectral correlation density function, the different fault characteristics are analyzed, which reduces the computational cost. Compared with the FFT and Hilbert transform, it proves that the combined slice method has the ability of anti-noise.

Acknowledgements This work is also partly supported by State Key Lab of Rail Traffic Control & Safety (Contract No. RCS2016ZT006). This work is also partly supported by National Key R&D Program of China (Contract No. 2017YFB1201201).

References

1. Lopez-Ramirez M, Romero-Troncoso RJ, Morinigo-Sotelo D et al (2016) Detection and diagnosis of lubrication and faults in bearing on induction motors through STFT. In: International Conference on Electronics, Communications and Computers, pp 13–18
2. Lu Y, Hu X, Zheng Y (2010) Instrument for bearing fault diagnosis based on demodulated resonance technology. In: International Symposium on Precision Engineering Measurements and Instrumentation. International Society for Optics and Photonics, 75445E-75445E-6
3. Teng W, Ding X, Zhang X et al (2016) Multi-fault detection and failure analysis of wind turbine gearbox using complex wavelet transform. *Renewable Energy* 93:591–598
4. Gardner WA (1988) Exploitation of spectral correlation in cyclostationary signals. In: Fourth Assp Workshop on Spectrum Estimation and Modeling, IEEE, pp 1–6
5. Bi G (2007) Research on feature extraction of rolling bearing and gear weak fault based on cyclostation. Shanghai Jiaotong University, Shanghai (Chinese)

6. Luo L (2002) On the spectral correlation of several issues-II physical function and application effect. *Electron Confront* 84(3):18–25 (Chinese)
7. Qiu H, Lee J (2006) Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J Sound Vib* 289:1066–1090
8. Feng F (2012) Early fault diagnosis technology of bearings based on wavelet correlation arrangement entropy. *J Mech Eng* 48(13): 73–79

Research on Performance Degradation Prediction of Urban Rail Vehicle Suspension System

Tengteng Wang, Xiukun Wei and Limin Jia

Abstract Along with the rapid development of urban rail transit, the security and reliability of rail transit vehicle system draw more and more attention. Suspension system plays a key role to support the car body, bogies, and wheels; therefore, it is vitally important to implement overall performance attenuation prediction, which can establish the maintenance plan based on the equipment status and reduce the cost. In this paper, a method of performance attenuation prediction about whole urban rail vehicle suspension system is proposed. The performance attenuation prediction model is established by utilizing LSSVR, of which the parameters are optimized based on SPSO. Compared with traditional methods, the SPSO-LSSVR model has higher prediction efficiency and prediction accuracy. In addition, considering the actual application, the prediction effect is analyzed when the vehicle body and bogie acceleration are model inputs, and a comparative study is carried out in the horizontal and vertical directions, which verifies the feasibility of the method used in the field engineering.

Keywords Suspension system · Performance attenuation prediction
LSSVR · SPSO

1 Introduction

The railway vehicle suspension system is made up of a series of springs and dampers and locates among the car body, bogies, and wheels, which affects many aspects, such as riding comfort, train operation stability, and buffer action [1].

Nowadays, there are generally three kinds of theoretical methods of performance attenuation prediction about mechanical equipment, which are mechanics-based method, probability and statistics method, and data-driven method [2]. Data-driven

T. Wang · X. Wei (✉) · L. Jia
State Key Laboratory of Railway Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: xkwei@bjtu.edu.cn

methods have been widely studied as the main methods in the field of performance degradation prediction, which is suitable for solving complex equipment system with complex physical laws and more uncertainties. In [3], SVM is applied to the fault classification of elevator vibration signals, and the results show that the characteristics of the small sample learning problem and the fast computation ability of SVM have great advantages in the application of real-time systems.

SVR is improved on the basis of the SVM and has been extensively studied. However, when the standard SVR algorithm is used to solve the QP problem, the number of training samples affects the dimension of the variables at a large extent, which leads to excessive computation [4]. In order to overcome the above defects, the LSSVR is used to establish the prediction model, and the SPSO algorithm is used to optimize the parameters of the model, which could effectively improve the prediction efficiency and accuracy. On the other hand, in the previous studies, the vertical acceleration of the vehicle body, the relative dynamic load of the wheel sets, and the dynamic displacement are taken as the indices to reflect the performance degradation of the suspension system, which are the inputs of the model. But in practice, it is difficult to measure these scalar quantities directly, and indirect calculation using formula will cause a large error. Considering the actual application, acceleration sensors could be installed in the car body and bogies to obtain their vibration information in the direction of lateral and vertical, through the analysis of the vibration information, the characteristics of the performance degradation of the suspension system can be obtained.

This paper is organized as follows. In Sect. 2, the principle of performance degradation prediction of suspension system is introduced. In Sect. 3, LSSVR and SPSO algorithms are briefly presented. In Sect. 4, the SPSO-LSSVR model is compared with traditional models in the MATLAB simulation environment, and the prediction effect of vehicle body and bogie acceleration as model input is analyzed. Finally, some conclusions are given in Sect. 5.

2 The Principle of Performance Degradation Prediction of Suspension System

The flowchart of the performance degradation prediction of suspension system is shown in Fig. 1. This section mainly introduces the fault simulation platform of suspension system and the selection of the performance degradation index as the input of the prediction model.

In this paper, the experimental data come from the fault simulation platform urban rail vehicle suspension system. First, using SIMPACK to build urban rail vehicle model, the model parameters are from CRRC Zhuzhou Institute. SIMPACK is a professional multibody dynamics software [5]. With building vehicle model, an interface between SIMPACK and SIMULINK is put up, and faults of different levels and components based on the integrated simulation platform are simulated. The integrated simulation platform is sketched in Fig. 2.

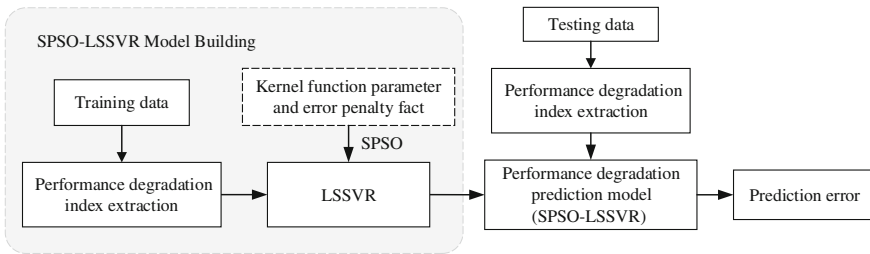


Fig. 1 Main steps of performance degradation prediction for vertical suspension system

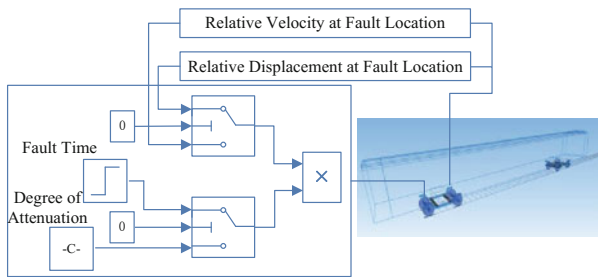


Fig. 2 Fault simulation platform of urban vertical suspension system

3 A Brief Review of the Prediction Algorithms

3.1 Least Squares Support Vector Regression

LSSVR replaces the inequality constraints in the standard SVR with equality constraint condition, so that the original quadratic programming problem is transformed into the problem of solving linear equations, which reduces the computational complexity [4].

For the given m training samples $Data = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, LSSVR model can be described as follows:

$$\min_{w, b, \xi} f(w, \xi) = \frac{1}{2} \|w\|^2 + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 \tag{1}$$

$$s.t. \ y_i = w^T \phi(x_i) + ba + \xi_i, \quad i = 1, 2, \dots, m$$

Using Lagrange multiplier method, the linear equations can be described as

$$\begin{bmatrix} 0 & I^T \\ I & \Omega + C^{-1}E \end{bmatrix} \times \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{2}$$

where $I = [1, 1 \dots 1]^T$; E is a unit matrix with m dimension; $\alpha = [\alpha_1, \alpha_2 \dots \alpha_m]^T$; $y = [y_1, y_2 \dots y_m]^T$; $\Omega_{ij} = \phi(x_i)^T \phi(x_j) K(x_i, x_j)$ is the kernel function which is satisfied Mercer conditions. Consolidating above formulas, the final LSSVR model is

$$f(x) = \sum_{i=1}^m \alpha_i K(x_i, x_j) + b \quad (3)$$

The prediction performance of the LSSVR model mainly depends on the choice of the penalty factor and kernel parameter, the traditional grid search method may lead to overfitting and takes too long when the parameter range is large. This paper proposed an improved PSO algorithm to optimize the penalty factor and kernel parameter, which improves the searching efficiency of global optimization.

3.2 Particle Swarm Optimization with Dynamic Adjustment of Inertia Weight Based on Similarity (SPSO)

In this section, based on similarity, a method of collection calculation is proposed, which is used to randomly initialize the position of the particle, control the diversification of PSO, and improve the ability of global exploration. The principle of standard PSO algorithm is referred to [6].

The similarity of two particles is defined as

$$s(i, j) = \begin{cases} 1, & d(i, j) \leq d_{\min} \\ 1 - \left[\frac{d(i, j)}{s_{\max}} \right]^\alpha, & d_{\min} \leq d(i, j) \leq d_{\max} \\ 0, & d(i, j) \geq d_{\max} \end{cases} \quad (4)$$

where d is the Euclidean distance for two particles; α , d_{\min} , and d_{\max} are fixed constants; the range of values of s is $[0, 1]$, and the bigger value means the higher similarity level of two particles. The updating of particle velocity depends heavily on the value of inertia weight w , compared with standard PSO, the basic idea of SPSO is that the inertia weight of particle i is not only smaller as the number of iterations increases but also dynamically adjusted with the similarity of the current optimal particle. The calculation formula of inertia weight w is

$$\omega = \omega_{\max} - s(i, g)(\omega_{\max} - \omega_{\min}) \quad (5)$$

$$\omega_i = \omega_{\min} + (\omega - \omega_{\min}) \cdot \left(\frac{t_{\max} - t}{t_{\max}} \right)^\beta \quad (6)$$

By dynamically adjusting the inertia weight, particles with similar similarity to the current optimal particle are searched in a small neighborhood with a smaller inertia weight, and the particles with lower similarity have higher searching power in the global range with larger inertia weight, so that the algorithm is not easy to fall into the local optimum.

4 Performance Degradation Prediction of Suspension System

Experimental data are derived from the fault simulation platform which is built in Sect. 2. In this paper, the coefficients of the overall suspension system are attenuated by 0–30%, and 101 sets of sample data (with 0.3% intervals) are used. For 101 groups of samples, 20 groups were taken as test samples at intervals, and the other 81 groups were training samples.

4.1 Performance Degradation Prediction of Suspension System Based on Acceleration, Dynamic Load, and Dynamic Displacement

In previous studies, the vertical acceleration of the vehicle body, the relative dynamic load of the wheel sets, and the dynamic displacement of the vehicle body and bogie are usually used to characterize the performance of the suspension system. The corresponding relationship between these indicators and the attenuation of suspension components is shown as follows.

As Fig. 3 shows, in the early stage of suspension component attenuation, the indicators remain at a low level, which means the suspension system has good vibration absorption and vibration reduction performance. In the middle and later period, the growth rate of each index is increasing, indicating that the performance of suspension system is deteriorating rapidly. The SPSO-LSSVR model is built with training samples, forecasting the testing samples, and the results of LSSVR model and SPSO-LSSVR model are shown as follows.

As Fig. 4 shows, the LSSVR model has high fitting accuracy for training samples, and the mean square error (MSE) is 5.36. But for testing samples, the MSE is 28.87, especially in the later period of performance degradation, the predictive error is large. The computational time is 63.47 s.

It can be seen from Fig. 5 that the MSE of SPSO-LSSVR for training samples is 9.12, which is slightly worse than LSSVR. But for testing samples, the MSE is 15.11, which is lower than LSSVR, it means that SPSO-LSSVR model has better

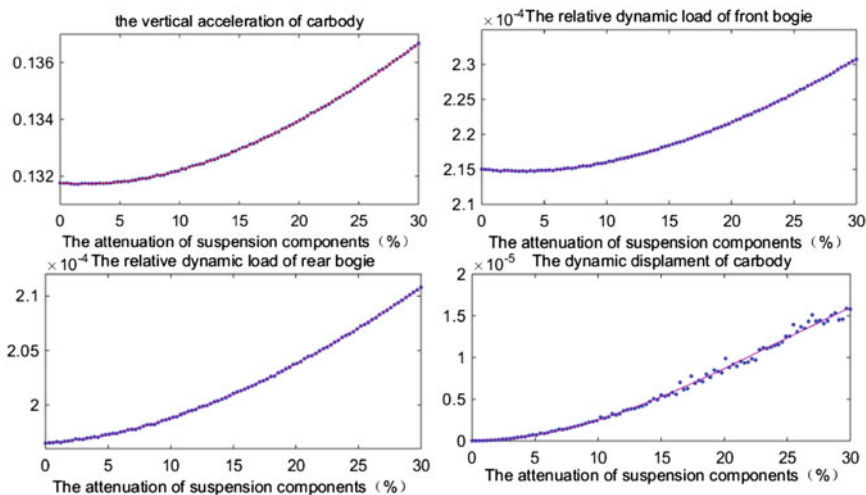


Fig. 3 The relationship between these indicators and the attenuation of suspension components

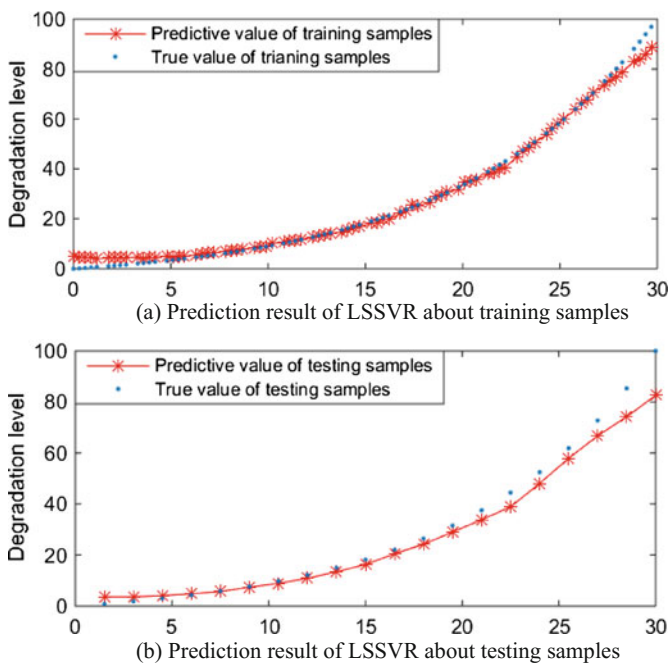


Fig. 4 Prediction results of LSSVR model

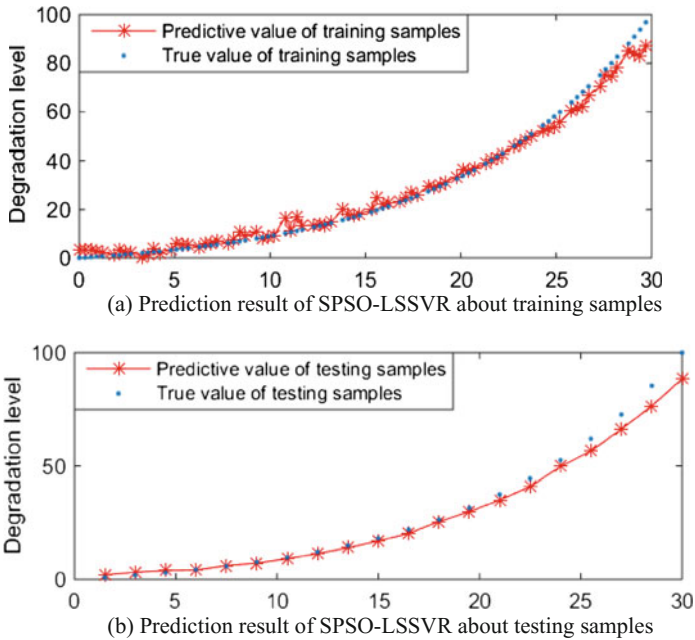


Fig. 5 Prediction results of SPSO-LSSVR model

generalization ability and higher prediction accuracy. In addition, the computational time is 39.63 s, which means the SPSO algorithm greatly improves the prediction efficiency of the model.

4.2 Performance Degradation Prediction of Suspension System Based on Acceleration of Car Body and Bogies

Considering the actual application, this section focus on analyzing the prediction effect of the SPSO-LSSVR model only when the input is the acceleration of the car body and bogies in direction of lateral and vertical.

In Fig. 6, the MSE of training samples is 15.46, the MSE of testing samples is 27.63. Compared with performance degradation prediction based on acceleration, dynamic load, and dynamic displacement, although the model prediction effect based on the acceleration of the car body and bogies is a little worse, taking into account the feasibility of the actual project, this method can provide reference with higher value.

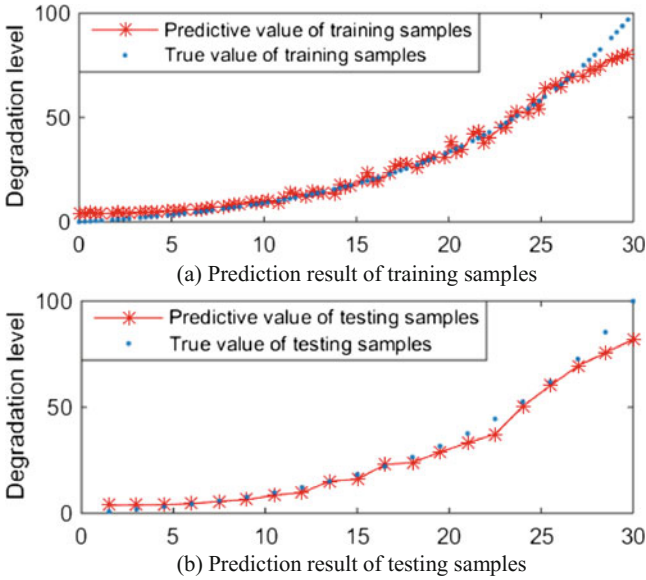


Fig. 6 Prediction results of SPSO-LSSVR model based on acceleration of car body and bogies

5 Conclusion

A method for performance degradation prediction of suspension system is proposed in this paper. The performance attenuation prediction model is established by utilizing LSSVR, of which the parameters are optimized based on SPSO. The simulation results show that compared with traditional prediction model, the SPSO-LSSVR model has higher prediction efficiency and prediction accuracy. In addition, according to the actual application, the prediction effect of vehicle body and bogie acceleration as model input is analyzed, which verifies the feasibility of the method used in the field engineering. Although this paper has made some initial results, there are still some problems that need to be studied further like field experiments instead of simulation studies.

Acknowledgements This work is partly supported by State Key Lab of Rail Traffic Control & Safety (Contract No. RCS2016ZT006) and National Key R&D Program of China (Contract No. 2017YFB1201201).

References

1. Liu H (2013) Research on the fault diagnosis of urban rail vehicle suspension system. Beijing Jiaotong University
2. Zhang X (2011) Review of life prediction for mechanical major equipments. *J Mech Eng* 47(11):100
3. Loutas TH, Roulias D, Georgoulas G (2013) Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic E-support vectors regression. *IEEE Trans Reliab* 62(4):821–832
4. Chang CC, Lin CJ (2007) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2(3, article 27):389–396
5. Zhou SX, Tao YZ (2013) SIMPACK 9 instance tutorial (volume one). Beijing Union Press, Beijing
6. Shi Y, Eberhart R (1998) A modified particle swarm optimizer. In: *IEEE world congress on computational intelligence*. Ahchorage, AK, USA, pp 69–73

Safe Location of High-Speed Maglev Train

Jianfeng Liu, Hongze Xu and Gegerile

Abstract This paper presented a new kind of architecture for the high-speed maglev train safe location. First, the function relationship of the operation control system of German maglev train was expounded, then the structure and function of train safe location were analyzed emphatically. Two sets of vehicle security computers were used in the new kind of safe location system architecture which were both connected with the wireless communication system, it enforced the completeness of the system structure and implemented real hot standby of the two sets of security computers, which improved the reliability of the system greatly.

Keywords High-speed maglev train · Train control system · Train safe location

1 Introduction

At present, in the field of maglev train transportation, Japan and Germany are the two main countries which have entered engineering test.

Japan started to research maglev technology in 1962, they achieved model train suspension driving for the first time in 1972, and they built Yamanashi test line in the 1990s [1]. The operation control system includes the subsystems such as traffic control, safety control, driving control, equipment monitoring, vehicle control, speed and position detection, and communication [2].

In the 1990s, Germany launched the applied technical test of high-speed maglev system in the Emsland test line. The operation control system of the test line is a three-layer structure which has four subsystems [3–5].

The subsystem includes the central control system, the decentralized control system, and the vehicle operation control system. The connection and data transmission between the three subsystems are realized through a communication network [6].

J. Liu (✉) · H. Xu · Gegerile
Beijing Jiaotong University, Beijing 100044, China
e-mail: 15120220@bjtu.edu.cn

In 2003, China introduced German technology and studied the German maglev traffic control system functions and core technology. Finally, China built the world’s first commercial maglev line in Shanghai [7].

“The 13th Five-Year Plan” national key development plan “modern rail transportation special” includes the speed of 600 km per hour high-speed maglev project and the speed of 200 km medium speed maglev project.

2 The Function Relationship of Germany Maglev Train Operation Control System

This chapter analyzes the function relationship of the maglev train control system in detail.

In Fig. 1, the function of the operation control system is represented by the ellipses, the peripheral components of the operation control system are represented by the rectangles, and the relationship and information interaction are represented by the arrows. The operation and display module is responsible for receiving the input command of the operator and transferring the command into driving parameter, then transmitting the driving parameter to other functional modules. The

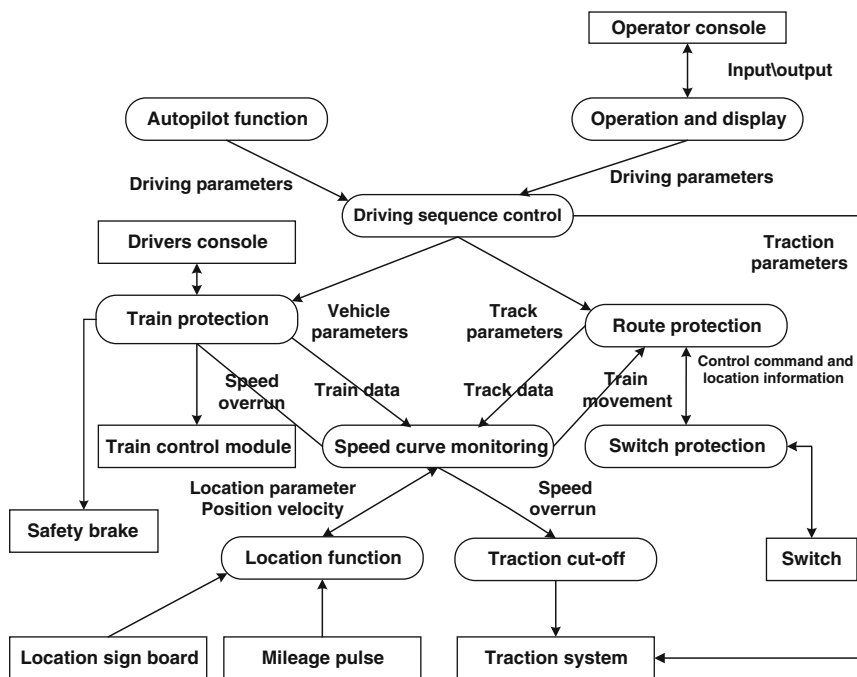


Fig. 1 Function relationship of operation control system

function of driving sequence control is to handle the driving parameters, then transmit the vehicle parameters to train protection and transmit track parameters to route protection. Route protection sets monitoring route according to track parameter. Therefore, route protection reads the switch position from the switch protection. Train protection monitors and controls the state of all trains, through train control module to manage the suspension and landing of trains. The positioning system determines the current position and speed of the train by means of the location signboards and the increment of the mileage pulse, achieves the safe location function of the operation control system, and provides the information to speed curve monitoring.

Speed curve monitoring is the core protection function of the operation control system, it calculates the allowable velocity limit based on train data and track data. The train position and velocity information are used to monitor whether the train speed exceeds the speed limit. If the current speed exceeds the limit, the overrun report will be given to the traction cutoff and train protection, traction cutoff will cutoff the traction power supply safely. When the speed exceeds the maximum limit, the train protection will activate the train eddy current braking according to the actual situation [8].

3 The Safe Location of the Train

Train safe location is used to determine the location and speed of the train by receiving the positioning parameters from the speed curve monitoring function.

These parameters include the identification information about the location signboards and their absolute position on the track. At the same time, it also includes the information about the position of the train, the train location state, the train speed, and the driving direction which is related to train, the driving direction which is related to track, and the direction of the train.

The positioning system can provide the following fault safety information: train location state, train position, train speed, driving direction which is related to train, and driving direction which is related to lines.

3.1 The New Structure of Positioning System

The positioning system of the maglev train operation control system provides safe train positioning data for the whole operation control system, the function of the positioning system is divided into the decentralized positioning system and the vehicle positioning system (Fig. 2).

DSC is the abbreviation of decentralized security computer, DTC is the abbreviation of decentralized transmission computer, VSC1/2 is the abbreviation of vehicle security computer, and VTC is the abbreviation of vehicle transmission

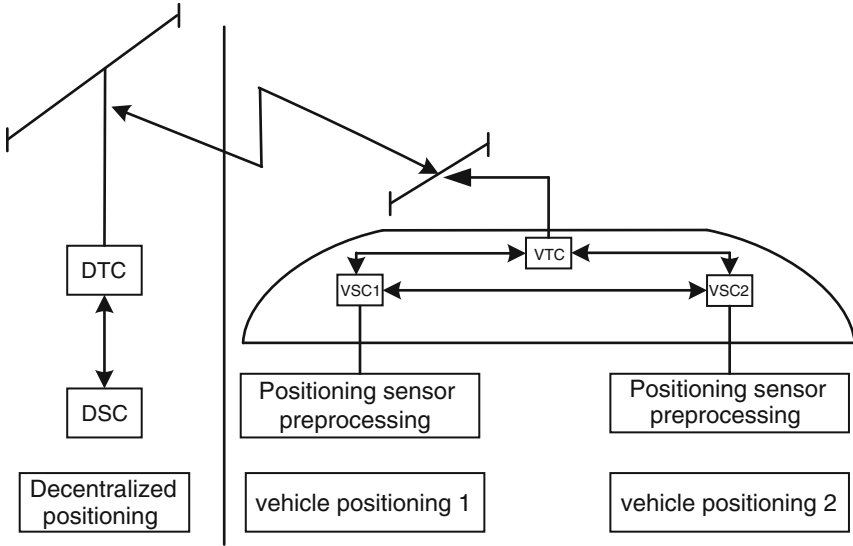


Fig. 2 Maglev train positioning system

computer. Under normal circumstances, the main function of the vehicle operation control system is completed by VSC1, while the VSC2 is in the hot standby state. DSC achieves decentralized position, VSC1 achieves vehicle position 1, and VSC2 achieves vehicle position 2. Under normal circumstances, the vehicle operation control system only uses the vehicle positioning 1. If there is something wrong with the vehicle position 1, vehicle position 2 is used to replace vehicle position 1, in order to provide the data needed for the train. The vehicle position sends the location information to the decentralized position through the 38 GHz wireless channel. Then, DSC manages the data to achieve safe location. In the German system, only VSC1 communicates with the VTC, in the new system VSC2 also communicates with the VTC. It makes the system structure more complete, realizes two sets of security computer really hot standby and improves the system reliability.

3.2 Decentralized Safe Location

The main features of decentralized safe location are the following functions (Fig. 3).

The management of initial and reference data module in DSC will be used to transmit the command of the train login and changes about driving direction to the VSC.

At the time of converting the control area, the synchronization of the positioning module with the adjacent decentralized control system is produced. So it is

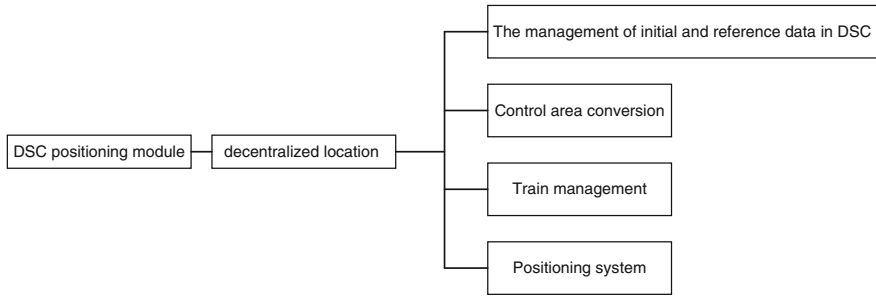


Fig. 3 DSC positioning system function diagram

necessary to communicate data by a control area conversion command. The internal train management module of decentralized positioning module maintains a list of trains, wherein the train in the control area of the decentralized control system is registered.

The positioning data determined by VSC is transmitted to DSC. After receiving the positioning data, the DSC processes the positioning data within the positioning function module. The positioning data which has been processed is provided to the speed curve monitoring module.

3.3 Vehicle Safe Location

Maglev vehicle positioning system has a total of eight positioning channels, which are located at both ends of the train and there is no similar positioning device in the middle car (Fig. 4).

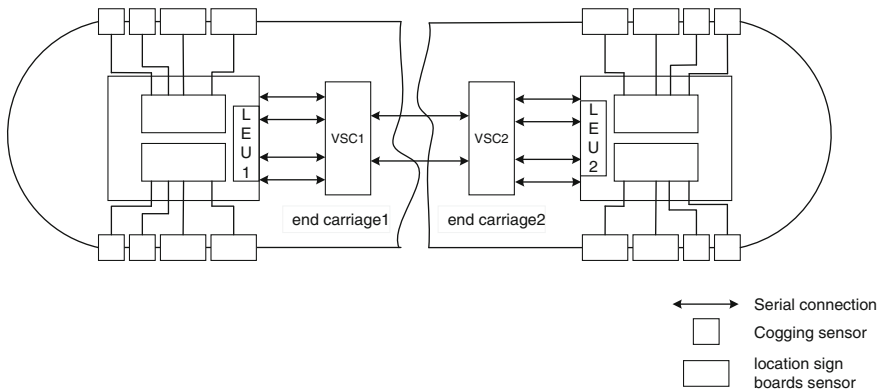


Fig. 4 The connection between vehicle safety computer and LEU

Both ends of the train are equipped with four cogging and four positioning mark sensors; a cogging sensor and a positioning mark sensor are connected to a location electronic unit (LEU) installed in the vehicle as a group to form a positioning channel. Therefore, each LEU of the carriage provides four independent positioning channels to the VSC. The VSC positioning software module will handle the positioning data from these positioning channels, and then generate a secure location data that can be used by the system [9].

The internal state management of the positioning system is mainly concentrated in the VSC1 positioning module, the whole state consists of nine basic states, and the details are shown in Table 1.

When the train location test command is executed, the positioning system is forced to turn to “state 1”, then starting the initialization check, before this step, the positioning system can be in any other state. When system completes the location initialization check, the system enters “state 2”, waiting for insert command to start “insert operation”.

3.4 The Sequence of Location Initialization Check

The initialization of the LEU in VSC1/2 is divided into five steps. First, the vehicle safety computer informs the LEU to get into the test mode; then, LEU returns a confirmation signal to VSC; vehicle safety computer starts ROM testing, RAM testing, and processor testing; after completion of these tests, the vehicle safety computer informs the LEU to switch to the normal operating mode; after the completion of the conversion of the LEU, the positioning data message is sent to the vehicle safety computer periodically.

3.5 Inserted Operation of the Train

In the decentralized control system, DSC will notify the DCC when the train has logged in. DCC will control the train to execute the “insert operation”. When the

Table 1 Internal state management of positioning system

Number	Illustration
0	Basic state after system starts
1	Initialization check
2	Initial check success
3	The train has been inserted and is ready for “insert operation”
4	The train starts to check position channel after passing the first positioning mark
5	The train completes “inserts operation” successfully, positioning system is secure
6	Recoverable positioning system is not secure
7	Nonrecoverable positioning system is not secure
8	Positioning system failure

train logs in the OCS system, the train management system can identify the initial login of the train, according to the internal state of the system, the location of the train will be corrected after the completion of the “insert operation”. The execution and monitoring sequence of the insert operation is controlled by the speed curve monitoring module.

4 Conclusion

High-speed maglev train positioning system is an important part of the maglev train operation control system. From the point of view of determining the position and speed of the train, the safe location can be regarded as the “eye” of the operation control system. From the point of view of providing the information of the train positioning state, the train position, the train speed, the train running direction, and so on, it can be regarded as a part of the control and safety protection technology.

References

1. Morishita K (2000) Train traffic control system on the Yamanashi Maglev test line. *Comput Railw VII* 7:641–649
2. Motoharu O (2002) Japan’s superconducting maglev train. *IEEE Instrum Meas Mag* 3:9–15
3. Guang Y, Zhenmin T (2006) Study on the operation control system framework for highspeed maglev train. *China Railw Sci* 06:68–69 (in Chinese)
4. Huaqing L (1995) German maglev vehicle. University of Electronic Science and Technology of China Press
5. Guang Y (2007) Study on high speed maglev train optimum speed curve and its tracking control. Beijing Jiaotong University (in Chinese)
6. He N, Long Z, Xue S (2013) Modeling and optimal design of relative position detection sensor for high speed maglev train. *Sens Actuators Phys* 15:24–32
7. Jian C (2015) The development trend of maglev transportation and its enlightenment abroad. *China Assoc Sci Technol* 7:291–292 (in Chinese)
8. Xiangming W (2003) Maglev train. Shanghai Science and Technology Press (in Chinese)
9. Han W (2007) Study on communication network of location and speed detection system for high speed maglev vehicle. Southwest Jiaotong University (in Chinese)

Dynamic Performance Evaluation of Beijing Subway Airport Line Vehicle with Domestic Primary Suspension and Secondary Suspension

Xi Li and Yuan Zhang

Abstract This paper introduced the method and result of Beijing subway airport line vehicle dynamic performance evaluation with domestic primary suspension and secondary suspension. Test method included test basis, test purpose, test conditions, test method, test data collection and analysis. The dynamic performance test was carried out in empty vehicle condition (AW0) with the highest speed limit as 110 km/h. The result showed that when the test vehicle ran on the straight and curve line below the highest speed limit, all dynamic indexes include wheel lateral force, and derailment coefficient and load reduction rate are within the limits specified in the specification.

Keywords Linear motor radial bogie vehicle · Domestic primary suspension and secondary suspension · Dynamic performance evaluation · Dynamic performance test

1 Introduction

Beijing subway airport line is an important rail line which connecting downtown and Capital International Airport. Vehicle of the subway line is jointly created by China CNR Corporation Limited CRC and Bombardier, and its prototype vehicle is New York John Fitzgerald Kennedy International Airport express vehicle. Beijing subway airport line was built and opened in 2008. With the operation time gradually increased and the backward of foreign products and services, long procurement and repair cycle, service is not timely and other problems had become increasingly prominent, which seriously affected the daily maintenance and operation of the vehicle. To address this problem, Beijing Mass Transit Railway Operation Corporation LTD carried out a series localization work of airport line vehicle components, especially domestic work in primary suspension and secondary

X. Li (✉) · Y. Zhang

Subway Operation Technology Centre, Beijing Mass Transit Railway
Operation Corporation LTD, Beijing 102208, China
e-mail: bjtulx@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_14

137

suspension. Dynamic performance test is a necessary test which could confirm whether the vehicle meets the project requirements of safety, stability and comfort [1]. It is commonly used in new vehicle validation process of different subway lines. As the airport line is the first subway line vehicle with linear motor and radial bogie in China, the dynamic performance of vehicle is very different from other subway lines in Beijing [2–5], and dynamics performance test result of initial operations is not very clear.

Therefore, in order to confirm the operation safety and evaluate the situation of vehicle dynamic responses in safety, stability and comfort aspects, after domestic research work, dynamic performance test of vehicle with domestic products especially domestic axle spring and air spring was carried out in empty condition. Dynamics performance test was carried out on the basis of GB/T5599-1985-Railway vehicles specification for evaluating the dynamic performance and accreditation test (hereinafter referred to as the specification) [6].

2 Dynamic Performance Test Conditions

2.1 Equipment

The following equipment and sensors were mainly used in the test:

- (1) Capacitive accelerometer with model 3713B112G and measuring range ± 2 g.
- (2) Plate instrumented wheel set with intermittent peak-to-peak value measurement mode, bridge and correction patch according to the ‘specification’.
- (3) LMS dynamic signal test system with 1–10,000 times magnification, measuring range ± 10 V and 24 bits sampling. It is used for amplification, filtering, sampling, recording and processing of the test signal, placed on the testing vehicle, connecting the sensors by four core shielded cable.

2.2 Vehicle and Marshalling

QKZ5 type subway vehicle marshalling is a total of four vehicles, the order is MC car (A1-No. 1071), M car (B1-No. 1072), M car (B2-No. 1073) and MC (A2-No. 1074), and vehicle marshalling is shown in Fig. 1.

Dynamics performance test vehicle is an MC train (A1) with domestic primary suspension and secondary suspension. Structure of the vehicle bogie is shown in Fig. 2 includes (1) axle, (2) primary suspension, (3) secondary suspension, (4) swing bolster (5) bogie frame, (6) disc brake equipment, (7) tachometer, (8) current collector, (9) magnetic track brake equipment and (10) linear induction motor (LIM) propulsion device.

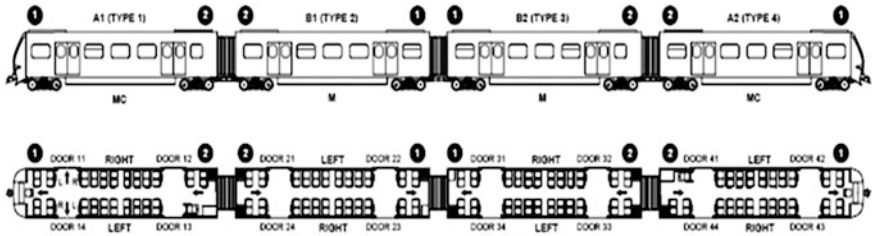


Fig. 1 Test vehicle marshalling

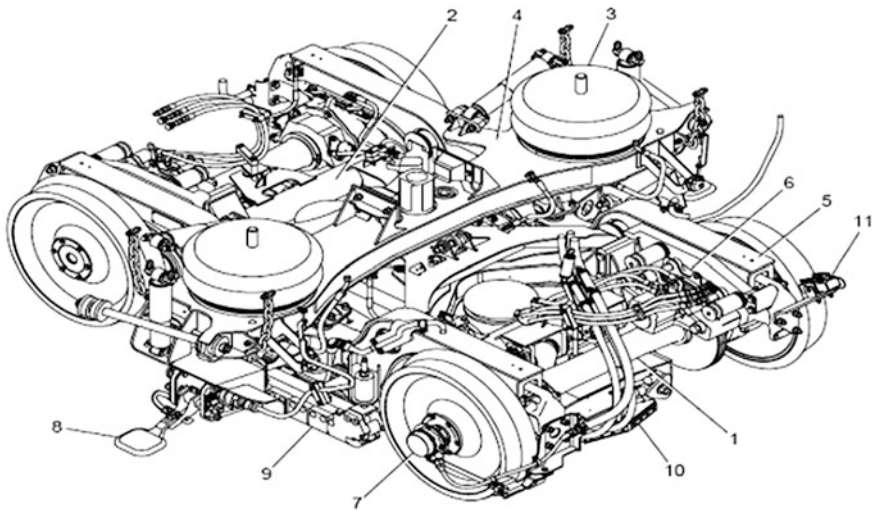


Fig. 2 Structure of test vehicles bogie

2.3 Line, Speed and Conditions

The test was carried out on the Beijing subway airport line with interval length of approximately 21.4 km, on the straight and curve line, the minimum radius of curve line is 160 m, multiple operations and single test mode.

According to the conditions of test line and operation, highest speed limit of the test is 110 km/h. The test speed was divided into several speed stages below the maximum speed, the difference of each speed stage is 10 km/h and the minimum straight-line testing speed is 60 km/h. The maximum speed of test vehicle pass through curve line according to operation allows and can achieve speed.

The dynamic performance test vehicle was carried out in empty vehicle condition (AW0).

3 Dynamic Performances Test and Evaluate Method

3.1 Main Test Items

- (1) Wheel-rail force test. The instrumented wheel set was installed in MC train (A1), as shown in Fig. 3. Replaced test wheel set to instrumented wheel set that after calibration. There are four wheel-rail force measuring points of each instrumented wheel set, in order to test vertical and lateral vehicle wheel-rail force of left and right side (Fig. 4).
- (2) Vehicle body vibration acceleration. According to the ‘specification’, four acceleration measurement (two vertical and two lateral acceleration sensors) points were installed on the test vehicle body ground to test vehicle body vibration acceleration, as shown in Fig. 5.

Fig. 3 Installation position of instrumented wheel set

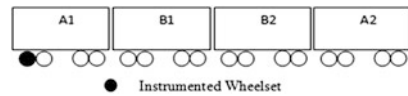
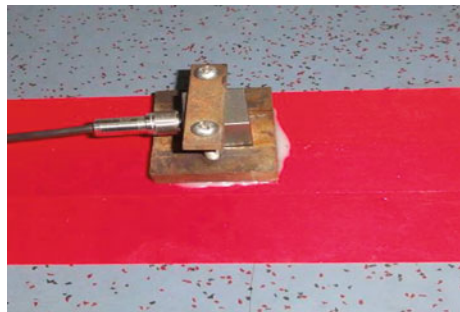


Fig. 4 Wheel-rail force measuring point



Fig. 5 Acceleration measuring point of test vehicle body



3.2 Evaluation Index

According to the ‘specification’, test data can be processed and evaluated by following indexes.

- (1) Stability index W was calculated by the acceleration signal value. Vertical stability index and lateral stability index was calculated by different frequency correction coefficient, but same rating was adopted.

$$\begin{aligned} W < 2.5 & \quad \text{Excellent level} \\ 2.5 \leq W < 2.75 & \quad \text{Good level} \\ 2.75 \leq W < 3.0 & \quad \text{Qualified level} \end{aligned}$$

- (2) Vehicle lateral force index H evaluated by lateral wheel interaction force which calculated by wheel-rail lateral force of instrumented wheel set is

$$H \leq 0.85 \left(10 + \frac{P_{st1} + P_{st2}}{2} \right)$$

According to test wheel-rail vertical force, the limit value of MC vehicle first wheel set is 27.46 kN when the vehicle is in empty conditions.

- (3) Derailment coefficient $\frac{Q}{P}$ calculated by the wheel-rail vertical and lateral force obtained from the instrumented wheel set is

$$\begin{aligned} \text{1st limit } \frac{Q}{P} & \leq 1.2 \text{ (Safety standard)} \\ \text{2nd limit } \frac{Q}{P} & \leq 1.0 \text{ (High Safety standard)} \end{aligned}$$

First limit standard is chosen to determine whether the test vehicle is qualified or not.

- (4) Wheel load reduction rate $\frac{\Delta P}{P_{avg}}$ calculated by the wheel-rail vertical force obtained from the instrumented wheel set is

$$\begin{aligned} \text{1st limit } \frac{\Delta P}{P_{avg}} & \leq 0.65 \text{ (Safety standard)} \\ \text{2nd limit } \frac{\Delta P}{P_{avg}} & \leq 0.60 \text{ (High Safety standard)} \end{aligned}$$

First limit standard is chosen to determine whether the test vehicle is qualified or not.

In the above indexes, W is vehicle stability index, H is axle lateral force, P_{st1} and P_{st2} are left and right static wheel weight of the same wheel, Q is wheel-rail lateral force, P is wheel-rail vertical force, ΔP is wheel load reduction rate, and P_{avg} is average weight of left and right wheel.

3.3 Test Process

First, decomposed the bogie, replaced the instrumented wheel set, confirm the vehicle meet the technical and operational requirements.

Second, complete test equipment setup such as sensors linking and device debugging. The line test was followed, test speed was divided into several speed stages below the maximum speed, the difference of each speed stage is 10 km/h and the minimum straight-line testing speed is 60 km/h. The maximum speed of test vehicle pass through curve line according to operation allows and can achieve speed. During the test period, test data of the vehicle ran through straight and curve line were recorded by test equipment.

Finally, the test vehicle was evaluated by evaluation indexes according to the collected data.

4 Dynamic Performances Test Result Analysis

According to ‘specification’ requirement, wheel-rail vertical vibration acceleration, vertical and lateral wheel-rail force of test vehicle were obtained. After data processing, the result can output index value above and calculate lateral stability index, vertical stability index, axle lateral force, derailment coefficient and wheel load reduction rate.

4.1 Vibration Acceleration

At operation condition with 110 km/h highest speed and different velocity levels, test vehicle ran through straight and curve line. The maximum vertical vibration acceleration of test vehicle first centre plate measuring point was 0.21 g, the average maximum values range from 0.03 to 0.11 g, and the RMS values range from 0.01 to 0.04 g. The maximum vertical vibration acceleration of test vehicle second centre plate measuring point was 0.17 g, the average maximum values range from 0.02 to 0.10 g, and the RMS values range from 0.01 to 0.04 g. The maximum lateral vibration acceleration of test vehicle first centre plate measuring point was 0.13 g, the average maximum values range from 0.03 to 0.08 g, and the RMS values range from 0.02 to 0.05 g. The maximum lateral vibration acceleration of test vehicle second centre plate measuring point was 0.15 g, the average maximum values range from 0.03 to 0.08 g, and the RMS values range from 0.02 to 0.05 g.

4.2 Stability Index

Stability index was calculated by vibration acceleration of test vehicle. At operation condition with 110 km/h highest speed and different velocity levels, test vehicle ran through straight and curve line. The maximum vertical stability index of test vehicle first centre plate measuring point was 2.35, which belongs to excellent level. The maximum vertical stability index of test vehicle second centre plate measuring point was 2.23, which belongs to excellent level. The maximum lateral stability index of test vehicle first centre plate measuring point was 2.73, which belongs to good level. The maximum lateral stability index of test vehicle second centre plate measuring point was 2.71, which belongs to good level (Figs. 6, 7, 8 and 9).

Fig. 6 Maximum values of vertical stability index on straight line

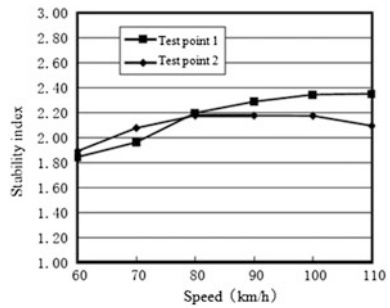


Fig. 7 Maximum values of lateral stability index on straight line

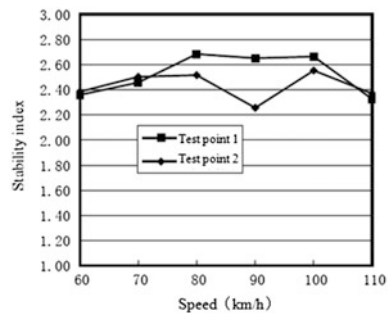


Fig. 8 Maximum values of vertical stability index on curve line

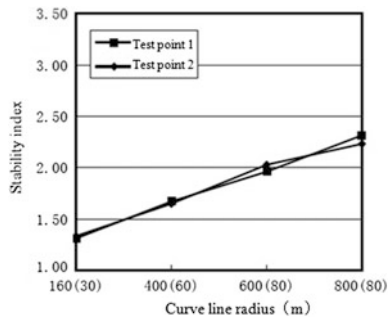
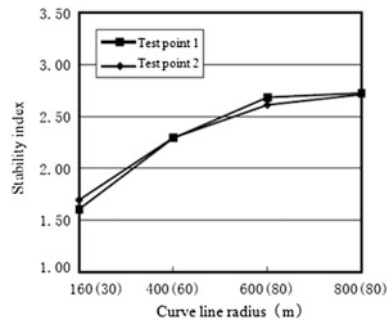


Fig. 9 Maximum values of lateral stability index on curve line



4.3 Wheel Lateral Force

Wheel lateral force was evaluated by lateral wheel interaction force which calculated by wheel-rail lateral force of instrumented wheel set. At operation condition with 110 km/h highest speed and different velocity levels, test vehicle ran through straight and curve line. The maximum wheel lateral force of test vehicle first test point was 17.37 kN, the average maximum values range from 5.37 to 9.99 kN, and the maximum value is less than ‘specification’ limits (Figs. 10 and 11).

Fig. 10 Maximum values of lateral wheel interaction force on straight line

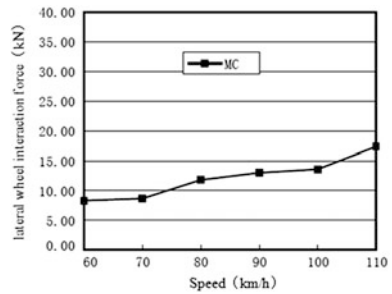
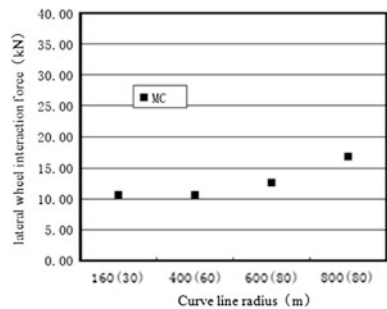


Fig. 11 Maximum values of lateral wheel interaction force on curve line



4.4 Derailment Coefficient

Derailment coefficient was calculated by the wheel-rail vertical and lateral force obtained from the instrumented wheel set. At operation condition with 110 km/h highest speed and different velocity levels, test vehicle ran through straight and curve line. The maximum wheel set derailment coefficient of test vehicle first test point was 0.38, the average maximum values range from 0.22 to 0.26, and the maximum value is less than ‘specification’ second limit (1.0) (Figs. 12 and 13).

4.5 Wheel Load Reduction Rate

Wheel load reduction rate was calculated by the wheel-rail vertical force obtained from the instrumented wheel set. At operation condition with 110 km/h highest speed and different velocity levels, test vehicle ran through straight and curve line. The maximum wheel load reduction rate of test vehicle first test point was 0.64, and the average maximum values range from 0.15 to 0.43. The maximum value is less than ‘specification’ first limit (0.65) (Figs. 14 and 15).

Fig. 12 Maximum values of derailment coefficient on straight line

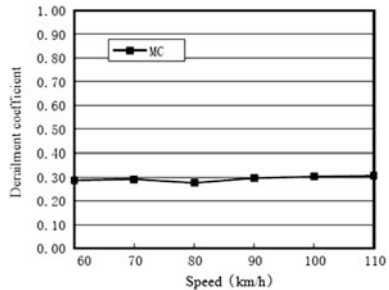


Fig. 13 Maximum values of derailment coefficient on curve line

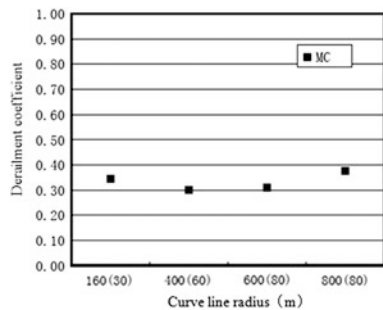


Fig. 14 Maximum values of wheel load reduction rate on straight line

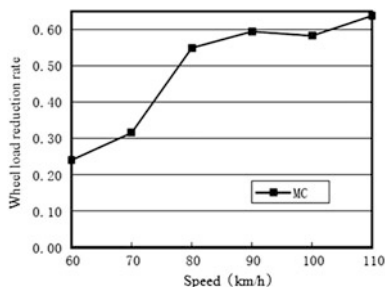
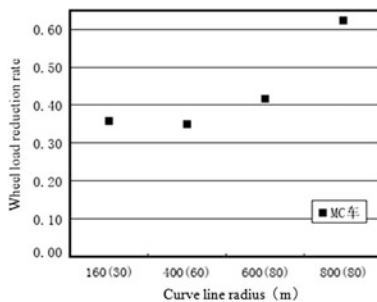


Fig. 15 Maximum values of wheel load reduction rate on curve line



5 Conclusion

The dynamic performance test of the QKZ5 type vehicle (No. 107) was carried out on the Beijing subway airport line. The maximum operating speed is 110 km/h. According to the above results, the following conclusions can be drawn:

First, when the test vehicle ran on the straight and curve line below the highest speed, the maximum value of vertical vibration acceleration is 0.21 g; the vertical stability index is 2.35, which belongs to excellent level. The maximum value of lateral acceleration is 0.15 g; the lateral stability index maximum value is 2.73, which belongs to good level.

Second, when the test vehicle ran on the straight and curve line below the highest speed, the lateral force of vehicle wheel and axle, derailment coefficient and load reduction rate are all within the limits specified in the specification.

Above all, when the test vehicle ran on the straight and curve line below the highest speed 110 km/h, all dynamic indexes are within the limits specified in the specification. Dynamics performance of Beijing subway airport line vehicle with domestic primary suspension and secondary suspension meets the operational requirements. The vehicle can be used in the normal operation.

Acknowledgments This paper is supported by National Key R&D Program of China (2016YFB1200402).

References

1. Luan X (1990) Automatic data processing method and error analysis of vehicle dynamics performance test. *Rolling Stock* 9:34–40 (in Chinese)
2. Yang L (2000) An analysis and discussion on radial truck vehicle with linear motor. *J Shanghai Tiedao Univ* 21(2):81–86 (in Chinese)
3. Feng Y, Wei Q (2004) Safety analysis of the linear motor wheel-rail system. *China Saf Sci J* 14(8):7–11 (in Chinese)
4. Yu Z (2004) Technical advantages and development of linear motor vehicle operation mode. *Modern Urban Transit* 1:52–57 (in Chinese)
5. Li F, Fu M, Huang Y (2003) Development and dynamic characteristics of radial bogies. *J Traffic Transp Eng* 3(1):1–6 (in Chinese)
6. GB/T5599-1985, Rail vehicle—specification for evaluation the dynamic performance and accreditation test (in Chinese)

Analysis of Traffic Impedance Effects of Highway Rescue for Passenger Train Accidents

Bing Wang, Xiexing Shi, Wenzhe Wang, Huiping Sun and Adila

Abstract In the less developed areas of China, the rescue of the injured personnel in passenger train accidents mainly relies on the implementation of highways, because of the effects of adverse traffic conditions, leading to rescue traffic impedance, long rescue time, and low efficiency of rescue. This paper, taking Xinjiang of China as the example, through extracting the related characteristics of highway rescue for passenger train accidents, analyzes the influence of rescue traffic impedance and proposes the methods of reducing the time for the highway remote rescue.

Keywords Passenger train accidents · Highway rescue · Rescue characteristics
Traffic impedance · Rescue time

1 Introductions

There are a vast territory and few railway and highway lines in the less developed areas of China, and traffic and transportation are restricted by natural environment. Although the train departure frequency is relatively low in these areas, once the passenger train accidents occurred, it is likely to cause heavy casualties. The accidents have the characteristics of temporal randomness and spatial randomness, due to the lack of measures of implementation for rescue of the injured personnel in passenger train accidents, poor dispatching technology in the less developed areas of China, resulting in large difficulty for rescue and high casualty ratio. According to the relevant regulations of the laws on emergency rescue and investigation and

B. Wang · X. Shi (✉) · W. Wang · H. Sun · Adila
Xinjiang University, No. 1230 Yan'an Road Urumqi, Xinjiang,
People's Republic of China
e-mail: 920989292@qq.com

X. Shi · H. Sun · Adila
Xinjiang Zheng Yang Traffic Planning and Design Institute,
No. 152 Xinhua Road Urumqi, Xinjiang, People's Republic of China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_15

handling of railway traffic accidents in China, after the occurrence of the passenger train accidents, it should be immediately carried out to rescue the injured personnel on the scene of the accident, and the government at the place where the accident occurred is responsible for the medical rescue work. Currently, restricted by many conditions, the rescue of the injured personnel in passenger train accidents in China's less developed areas still relies mainly on the implementation of highways, because of the sparse highway network, less rescue resources, and long distance for rescue, easily affected by random bad traffic conditions in this kind of area, making the rescue show the situation of highway remote rescue, and hence leading to long rescue time as well as low rescue efficiency.

This paper, citing Xinjiang of China as the example, through extracting the characteristics of network distribution of railways and highways, as well as rescue characteristics, analyzes the influence of traffic impedance of the rescue of the injured personnel in passenger train accidents and then proposes the methods of reducing the rescue time.

1.1 Characteristics of Network Distribution of Railways and Highways

1.1.1 Density of Network of Railways and Highways

According to recent statistics of China Statistical Yearbook, the distribution of China railway mileage and highway mileage in different regions is as shown in Fig. 1.

Among them, the distribution density of railways and highways in less developed areas such as Xinjiang, Qinghai, and Tibet is obviously lower than that in other areas. Xinjiang is located in the northwest of China, accounting for 1/6 of the

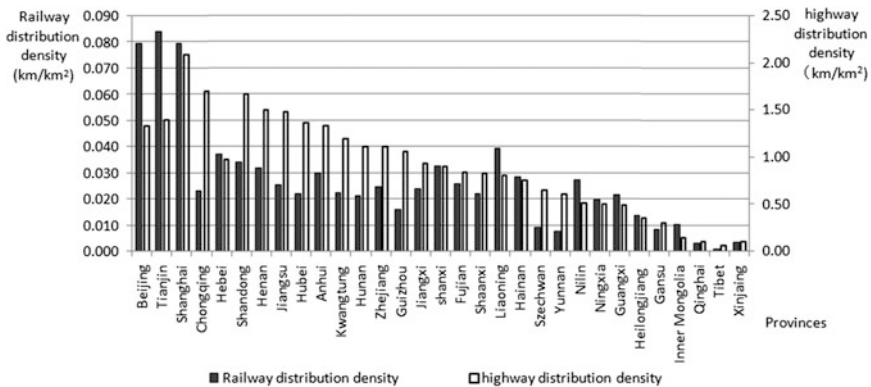


Fig. 1 Contrast map of distribution density of railway and highway in China

country's land area. There are railway lines in the autonomous region such as Lanzhou-Xinjiang Railway, Lanzhou-Xinjiang High-speed rail, Kruger Railway, Xinjiang-Tibet Railway, etc., and international railway lines such as Central Asia Railway, China-Pakistan Railway, etc. The railway distribution density reaches 0.003 km/km^2 , below the average level of the railway distribution density with 0.01 km/km^2 in China. Xinjiang is under the construction of the "Three transverse, Two longitudinal, Second ring, Eight channels" highway grid Bureau. The current highway density is 0.11 km/km^2 —and also below the national average level density of 28.30 km/km^2 ; grade highways in Xinjiang are mainly secondary highway or below, lack of high-class highways; the number of network nodes and optional paths are less, and the lines and travel time between nodes are long.

1.1.2 Association of the Layout of Network of Railways and Highways

Xinjiang railway lines and highway lines are similar, and the location of intersection nodes is approximately coincident; railway and highway are on the line, or formed a cross. The distance between the two construction land is not far. Once the passenger train accidents occurred, rescue resources will come to rescue relying on the highway. Meanwhile, there is a lot of Gobi and deserts in Xinjiang. When the spot of passenger train accidents is away from the highway, it also provides certain conditions for the rescue vehicles to develop new shortest routes at the right time.

1.2 Remote Rescue Characteristics

1.2.1 Accident Characteristics

Passenger train accidents, referring to the accidents such as collision or derailment, etc., in the course of running, which affect the normal traffic safety and cause casualties, with the traits of contingency, sudden, sociality, and unpredictability, are liable to cause a mass casualty. The occurrence of accidents is very likely to affect the other trains which are on the same line stop and to limit the speed of trains which are on the adjacent lines, resulting in a large area of late train and other derivative events. The injury types of injured personnel in passenger train accidents, mostly because of sudden deceleration and stop of the train, are fracture, visceral injury, and traumatic brain injury, which are all under the local "120" medical rescue and disposal capacity category.

1.2.2 Distribution of Rescue Resources

It is often to abbreviate Chinese medical rescue organization as “120”. In this paper, an ambulance (equipped with a group of medical rescue personnel and a set of rescue and medical rescue equipment) is used as a rescue unit. According to the investigation, “120” rescue resources at all levels in Xinjiang are equipped with 1–3 rescue units. One or two people who are seriously injured can be treated by each rescue resources. “120” medical rescue resources are mainly distributed central urban area of each city and county. The number of rescue resources per square kilometer is only 0.000072, and the distance between every rescue resources is far.

1.2.3 Traffic Characteristics of Rescue Highways

The highways in Xinjiang, with less highway network node, long distance between nodes, and few optional paths, are mainly secondary highway or below and are basically two-way two lanes. There is not enough room for the vehicle to turn round. The design and construction of highway are influenced by geography, geology, and construction conditions, and the partial alignment is difficult to optimize, affected by the special weather such as snow, fog and wind, and the freezing of cold pavement. The speed limit level is generally not high, and the driving safety and reliability are low. Traffic characteristics of rescue highways cause high traffic impedance for rescue for passenger train accidents and form strong constraints for rescue benefit.

1.2.4 Rescue Process

After the occurrence of passenger train accidents, the phases of the rescue process for the injured under the existing relevant emergency rescue mechanisms and plans are as shown in Fig. 2.

- (1) Alarm responses, with time T_a . By the railway traffic control department and alarm to interact with the local rescue dispatching agency to provide an accurate basic situation of traffic accident identification, identifiable locations, type and extent of the casualties, injured, and the scene involved in the rescue of the number of information.

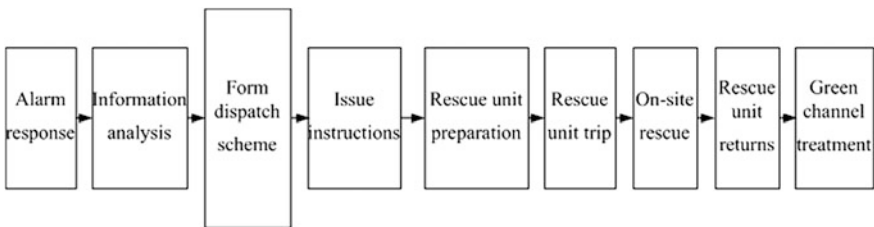


Fig. 2 Flowchart of rescue process

- (2) Information analysis, with time T_i . Count the number of injured persons and types of injuries and conduct manual selection of rescue units and round-trip routes; select rescue demands resource conditions of green channel, etc.
- (3) Form dispatch scheme, with time T_d . Manually determine the optimal rescue resources and rescue units and their round-trip routes. Put forward rescue resources preparation conditions of green channel, etc.
- (4) Issue instructions, with time T_o . The dispatching command is issued to the relevant rescue resources by the rescue dispatching agency.
- (5) Prepare rescue unit, with time T_r . According to the dispatching instruction, set up rescue forces.
- (6) Rescue unit trip, with time T_g . By paths and speed limit instructions, the rescue unit goes to the scene.
- (7) On-site rescue, with time T_f . Implement on-site rescue and temporary medical rescue.
- (8) Rescue unit returns, with time T_b . Send the wounded back by paths and speed limit instructions.
- (9) Green channel treatment, with time T_t . Access to the green channel and implement medical rescue.

Set the total time for the highway remote rescue process for passenger train accidents as T , then

$$T = T_a + T_i + T_d + T_o + T_r + T_g + T_f + T_b + T_t \quad (1)$$

Main affecting factors of each stage time and their relationship is complex, with a multidisciplinary association. At present, greatly reducing the total time of each stage is difficult to achieve. The rapid selection of rescue units and the effective guidance of their round-trip routes are the key issues to effectively shorten the rescue time.

1.3 Analysis Traffic Impedance Effects of Highway Rescue

1.3.1 Causes of Traffic Impedance

Xinjiang has the temperate continental climate, with a long cold winter and a short hot summer. The temperature difference between the northern and southern Xinjiang is great, and the temperature in day and night varies obviously in different regions. There exist famous “Anxi Wind Pool”, “Thirty-mile Wind Zone”, and “Hundred-mile Wind Zone”. All grades of highway are mainly located in alluvial plain and Gobi area, and the terrain is relatively flat, part of which is in the mountains and hilly areas. Highways are affected greatly by special climate, weather, and road surface conditions, and the traffic impedance of each route is complicated because of the randomness of those effects. The limited optional paths make the decision of the shortest route of rescue become a difficult problem.

The highway rescue is affected randomly by special weather such as heavy snow, dense fog, gusty wind, and icy pavement which do harm to the thoroughfare of rescue vehicles. By statistics analysis, single adverse conditions and influences are as follows: wind affecting driving stability, rain affecting friction coefficient and visibility, snow affecting friction coefficient and visibility, fog affecting friction coefficient and visibility, icy pavement affecting friction coefficient, and sandstorm affecting visibility and driving stability; coupling adverse conditions and influence: the wind with rain affecting driving stability, friction coefficient and visibility, the wind with snow affecting driving stability, friction coefficient and visibility, the rain with snow affecting friction coefficient and visibility, ice and snow-road surface affecting friction coefficient, etc. The main influence elements of various unfavorable traffic conditions are visibility and friction coefficient.

1.3.2 Effects of Traffic Impedance on Rescue Benefit

The establishment of rescue capability function R , referring to the number of serious injuries that the medical rescue process can save, is established. If we set the number of people who are injured seriously in the passenger train traffic accident as H , it should be subjected to

$$R \geq H \quad (2)$$

Setting the original rescue capacity of rescue units as a constant R_0 , affected by the random traffic impedance of the round-trip of rescue units, generating a trip cut function R_f and a return cut function R_b for rescue capacity, then

$$R = R_0 - (R_f + R_b) \quad (3)$$

Trip cut function R_f and return cut function R_b for rescue capacity can be described as

$$R_b = \sum F(X_{bi}) \text{ and } R_f = \sum F(X_{fi}) \quad (4)$$

Among them, X_{fi} and X_{bi} are the traffic impedance formed by the main influence factors of the single or bad coupling conditions on each section of the round-trip paths.

In order to meet Eq. (2) requirements, the rescue capability of a single rescue unit is obviously insufficient in the face of the passenger train accidents characteristics, and the number of n optimized rescue units should be selected for group rescue. Then,

$$R = \sum R_n \quad (5)$$

1.4 Technical Means to Reduce Rescue Time

Under the condition of limited passenger train accident rescue means in the less developed areas, it is difficult to dispatch accurate dispatching information quickly because of the lag of rescue scheduling technology. This paper puts forward the main technical measures to reduce the time for highway remote rescue in the current objective condition of passenger train accident rescue:

- (1) Optimize and select the best rescue unit and its round-trip paths. Construct dynamic information collection system for main influencing factors of bad road conditions. Because the optional route for rescue units near accident spot to come to rescue is less, the method of exhaustion is used to calculate and compare the rescue capability function R , so as to make the rescue capability maximize, then decide and select the optimal rescue unit rescue route, and format rescue unit round-trip paths planning and instructions to guide rescue unit to carry out rescue.
- (2) Establish dispatching system of highway remote rescue for passenger train accidents. Using the various dispatching command platforms built in different regions of Xinjiang to develop embedded scheduling software of highway remote rescue for passenger train accidents, and to establish a dispatching system of highway remote rescue for passenger train accidents based on GIS-T in Xinjiang. Once the passenger train accidents occur, the system automatically selects the best rescue unit, establishes the rescue unit round-trip route planning program, and issues rescue-related instructions.

1.5 Epilogues

This paper, through the analyzing of the related characteristics of the highway remote rescue for passenger train accidents, explores and researches traffic impedance of highway remote rescue and its effects on rescue benefit, and then puts forward the method of reducing the rescue time. Because there are few literatures in this study, it is expected that this paper can provide references for further research on passenger train accidents and other railway accident rescue techniques. With the rapid development of “The Belt and Road” international development strategy, railway and highway integrated transport pattern will be formed quickly among China and relevant nations. In this pattern, the research and development of new technologies and new theories for all kinds of traffic accident rescue under objective limited conditions will also be carried out in depth.

Acknowledgements This paper is funded by the science and technology project of Xinjiang Transportation Department in 2017.

References

1. The State Council of the people's Republic of China. Regulations on emergency rescue and investigation and handling of railway traffic accidents. <http://www.gov.cn/zhengce/content/2008-03/28/content3725.htm> (in Chinese)
2. National Bureau of statistics of the People's Republic of China (2016) China statistical yearbook. China Statistics Press, Beijing (in Chinese)
3. Ministry of railways. Emergency rescue rules for railway traffic accidents. <http://www.110.com/fagui/law365082.html> (in Chinese)
4. State Railway Administration. Report on safety supervision of railway transportation in 2013. <http://www.nra.gov.cn/xwzx/xwdt/xwlb/201405/t201405226185.shtml> (in Chinese)
5. State Railway Administration. Railway Safety Bulletin in 2014. <http://www.nra.gov.cn/jgzf/ajjc/zfdt/201504/t2015042713279.shtml> (in Chinese)
6. State Railway Administration. Railway Safety Bulletin in 2015. <http://www.nra.gov.cn/xwzx/xwdt/xwlb/201603/t2016031722146.shtml> (in Chinese)
7. State Railway Administration. Railway Safety Bulletin in 2016. <http://www.nra.gov.cn/xwzx/xwdt/gdxw/201703/t2017032335994.shtml> (in Chinese)
8. Wang B, Chen W, Shi X (2015) Process mechanism analysis of remote first aid and rescue capability modeling for emergencies on sparse road network. In: 15th COTA international conference of transportation professionals: efficient, safe, and green multimodal transportation, CICTP 2015. American Society of Civil Engineers (ASCE), Beijing, China, pp 3236–3241
9. Wang B (2016) Case study of traffic improvement design in western remote cities. People Communications Press (in Chinese)

Application of Multi-resolution State Domain Method in State Identification of Train Motor Rolling Bearings

Xinan Chen, Limin Jia, Yunxiao Fu and Yong Qin

Abstract Based on the research of motor rolling bearings of the urban rail train, we propose a kind of multi-resolution state domain (MRSD) method to identify the working condition of the rolling bearing and apply principal component analysis (PCA) method on fault classification. First, the preprocessed signal is decomposed by wavelet packet algorithm to obtain wavelet packet coefficients, and then the correlation coefficient with the original signal is calculated. On the foundation of the correlation coefficient, multi-resolution correlation entropies (MRCE) are extracted as feature parameters. Then, we utilize PCA approach to get Descartes set which represents bearing status, that MRSD, to classify the different-sourced motor bearing condition intelligently. Finally, the method is certified effectively and feasibly by the experimental result to discriminate the state of train motor bearings.

Keywords Correlation entropy · Multi-resolution state domain
Principal component analysis · Rolling bearing · State identification

1 Introduction

Rolling bearing is the crucial mechanical part of rotor-rolling bearing system. Its working condition directly affects the working reliability of the whole system [1]. Metal roller surface of bearing is smooth, and their back and forth movement between the inner race and outer race. Lubrication metal surface is useful to reduce friction [2]. When the bearing element rolls damaged surface points, it will produce mutations impulse force [3]. Damage fault has the basic characteristics such as strong sudden, high risk, difficult to recognize the early symptoms and so on [4].

X. Chen · L. Jia (✉) · Y. Fu · Y. Qin
The State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, No. 3 Shang Yuan Cun,
Hai Dian District, Beijing, China
e-mail: jialm@vip.sina.com

The key technology of rolling bearings fault diagnosis includes feature extraction and state identification [5]. The energy or energy entropy of the decomposition of wavelet transform or EMD is extracted as the characteristic parameter [6]. Although the characteristic parameters are widely used, the diagnostic efficiency needs to be improved [2]. State identification is another essential process for rolling bearing fault diagnosis, such as artificial neural network [7], hidden Markov model [8], and support vector machines [9]. However, these algorithms all have fixed structure so that optimal decisions cannot be preset parameters, and therefore, the recognition accuracy would be impacted. This paper analyzes the condition trend of the bearing from the perspective of domain. Wording process is shown in Fig. 1.

2 State Domain Model

Domain thought as a method to solve multivariate measure mathematic problem primarily used in electric system safety analysis. Assuming the feature set is Q and the state set is S , and every vector element of S is a mapping of Q , $Q \rightarrow S$ owns n dimensions, which are, respectively, composed of m -dimensional irrelative state vectors and n -dimensional irrelative feature vectors. This Descartes set satisfies the basic property of Hilbert space, and it is defined as a subspace of Hilbert space, which is named of Euclidean space. The mathematic description is as follows:

1. $\Lambda = F(\Omega)$. $\text{dimension}(\Lambda) = m$, $\text{dimension}(\Omega) = r$, $m < r$. Λ is the feature fusion space, while Ω is the feature space. Every vector in Λ is one to many mapping result of Ω .
2. State space description. $\text{dimension}(S) = n$, s_i is i th state element, the mathematic expression is $S(n) = \{s_i | i \in [1, n] \cup i, n \in N^+\}$.
3. $\Lambda(m) = \{q_i^j | i \in [1, n] \cup j \in [1, m] \cup i, j, m, n \in N^+\}$. This mathematic expression gives a description of feature fusion space. q_i^j is an element of space Λ and expresses j th feature elements of i th state element.

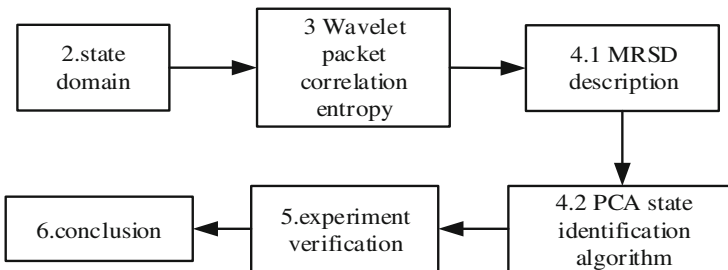


Fig. 1 Article outline

4. Feature element belonging to i th state element. $s_i = \{q_i^j \times q_i^k | i, j, k, \in N^+\}$, so, $S = \{q_i^j \times q_i^k | j, k, \in N^+\}$ where $q_i = \{q_i | i \in N^+\}$. State space is a collection of Descartes sets composed of feature fusion space. The mathematic model represents the relation between the feature space and the state space, which called state domain. It is applicable for the system that meets the above terms.

3 Multi-resolution Correlation Entropy Algorithm

Multi-resolution correlation (MRC) is obtained by calculating the correlation between the original signal and the coefficients of wavelet packet decomposition. In the wavelet packet frequency, if bearing signals in different states have obviously distinct correlations to the original signal, then this character can be used as one element in identifying bearing states. Supposed that primary signal λ has n times of sampling points. Wavelet packet transform is used to reach the l level of decomposition, and this level has $t = 2^l$ frequency bands l_1, \dots, l_t . The correlation entropy of corresponding some frequency is calculated as Eq. 1:

$$\left\{ \begin{array}{l} I_{l_i} = \sum_{k=1}^n l_i^k - (\sum_{k=1}^n l_i^k)^2 / n \\ I_{\lambda\lambda} = \sum_{k=1}^n \lambda^k - (\sum_{k=1}^n \lambda^k)^2 / n \\ I_{l_i\lambda} = \sum_{k=1}^n l_i^k \lambda^k - (\sum_{k=1}^n l_i^k \sum_{k=1}^n \lambda^k) / n \\ \gamma_i = I_{l_i\lambda} / \sqrt{I_{l_i} I_{\lambda\lambda}} \end{array} \right. \quad (1)$$

where l_i^k is k th time-frequency coefficient in i th sub-frequency of level l , and λ^k is k th sampling point of primary signal. $I_{\lambda\lambda}$ represent the variance of primary signal while $I_{l_i\lambda}$ represents the covariance of frequency band l_i with original signal, and γ_i is the correlated coefficient index of single frequency to original signal.

Correlation coefficient set $\Omega = \{\gamma_i | i \in (1, 2, \dots, 2^l)\}$ for all the frequency bands to the original signal is considered as 2^l dimensions multi-resolution space as well as the whole features set of devoted for bearing identification.

Information entropy signifies the uncertain measure of signal. In stochastic nonlinear signal processing, information entropy can characterize probability measure of state, where the entropy indicates the uncertain degree between coefficients of wavelet packet with the original signal. Multi-resolution correlation entropy (MRCE) contains this uncertain information, so that the results tend to be more stable in the case of state recognition accuracy being not affected by probability error. The only aspect needs to be concerned is that the correlation entropy without the relevant of positive or not only contains the correlation information, and the sign is determined by correlation coefficients. The equation is as follows:

$$\begin{cases} p_i = |\gamma_i| / \sum |\gamma_i| \\ H'_i = p_i \ln p_i \\ H_i = \gamma_i H'_i \\ i \in [1, 2, \dots, 2^l] \end{cases} \quad (2)$$

In which p_i expresses the probability distributions of some wavelet packet sub-frequency, while H'_i stands for Shannon entropy of some wavelet packet sub-frequency, and H_i represents the meaning of MRCE.

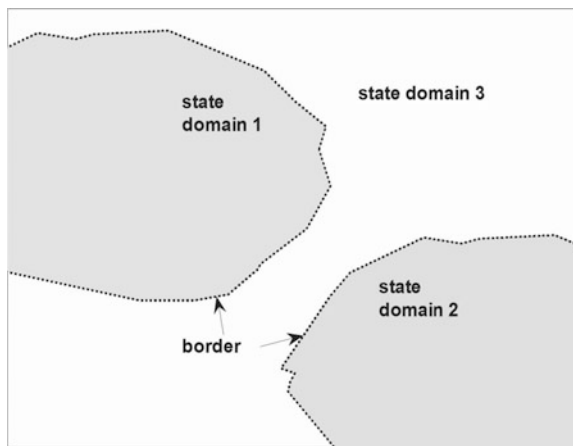
H_i is the improvement of correlation and contains correlation and uncertain information. There are 2^i MRCE in total after n levels wavelet packet decomposition, which means feature space is of 2^i dimensions.

4 Multi-resolution State Domain

4.1 Description of Multi-resolution State Domain

Multi-resolution means the multi time–frequency space which is gained by wavelet packet transform, while state domain explains a two-dimensional state space that is represented by a group of Descartes feature sets. Different states occupy different regions, and the case is shown in Fig. 2. MRSD applied to identify the working conditions of bearing takes the information of multi-resolution as the coordinate of state domain. The state domain should have a closed border that is used to distinguish which state the tested point belongs to. There are three steps to calculate the domain border.

Fig. 2 Status domain schematic



- Step 1: Center of gravity method to calculate each state training samples.
- Step 2: Compute the vertical of connection between two states' center of gravity.
- Step 3: To intercept the state domain border based on the principle that domain boundaries vertex should become the intersection of the vertical of every two states.

4.2 PCA State Identification Algorithm

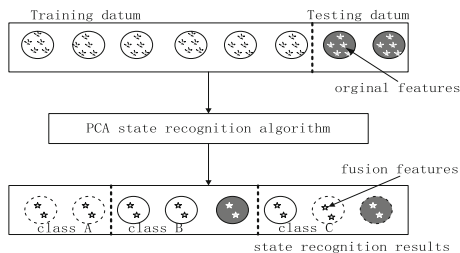
Principal component analysis (PCA), also known as principal component analysis, is aimed to apply dimensional reduction thought to transform the multiple indexes to few comprehensive indicators. The proportions that the first principal component and the second main component occupied constructed the Descartes set of state domain. Disordered training and testing samples as input set can be classified by uncorrelated parametric which is produced by PCA coordinate transform, then the classified result in foundation of samples state domain comes out to realize the function that recognizes the bearing condition intelligently. The philosophy of PCA discriminating approach is shown in Fig. 3.

5 Experiment Analysis

Here, the status field is used to describe two-dimensional Descartes set characterized by rolling state. State identification experiment process is as follows:

- Step 1 Collecting vibration acceleration signals of rolling bearings, and making noise cancelation preprocessing to get original vibration signal, where sampling frequency is 12 kHz and speed is 1745 r/pm. The inner race fault depth and the outermost end of the outer ring fault depth are both 0.178 mm, while the ball pitting failure depth is 0.14 mm.
- Step 2 Extracting each 30 samples of the roller fault, the inner race ring fault, the outer ring fault and no-fault state and calculate MRCE (four level decompositions) of all samples. Choose one sample in each state to show

Fig. 3 PCA identified principle schematic



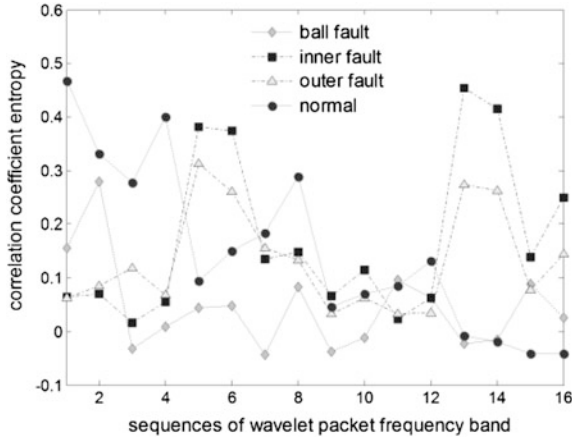


Fig. 4 Correlation entropy distribution of wavelet packet

the correlation entropy distribution of each frequency band which is shown in Fig. 4.

Step 3 Using PCA algorithm to make space conversion to MRCE and extracting the first two principal components with the maximum variance to construct two axes of two-dimensional state domain. 120 sets of training datum that consist of 30 samples evenly from four states (ball failure, inner race failure, outer race failure and no failure) are extracted, meanwhile, another 40 group datum is selected as test samples (Fig. 5), multi-resolution intelligent recognition result is shown by the status field in Fig. 7.

In Fig. 6, subfigure (a) shows the layout of security domain and failure domain with black points security and golden failure. Subfigure (b) gives the states identification domain layout, where red points stand for normal state and orange outer

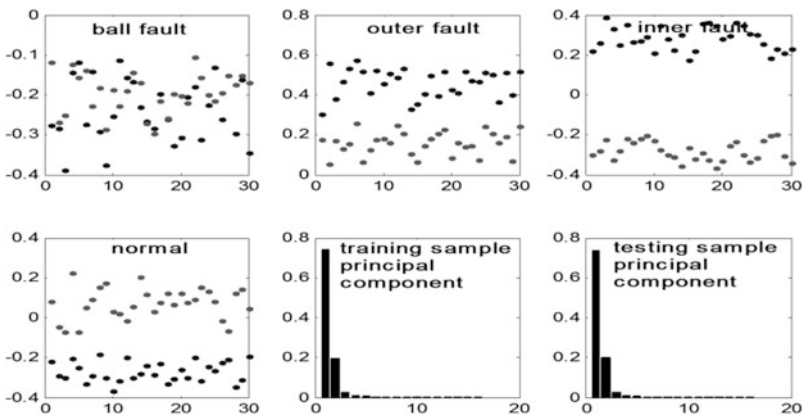


Fig. 5 Sample principal component distribution under different conditions

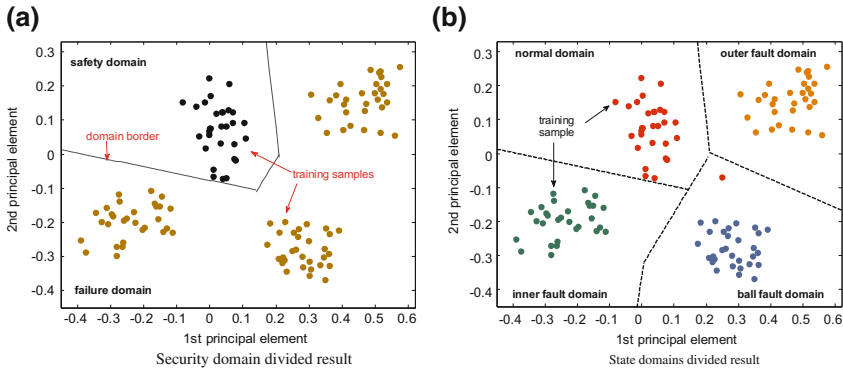


Fig. 6 The generation of state domain

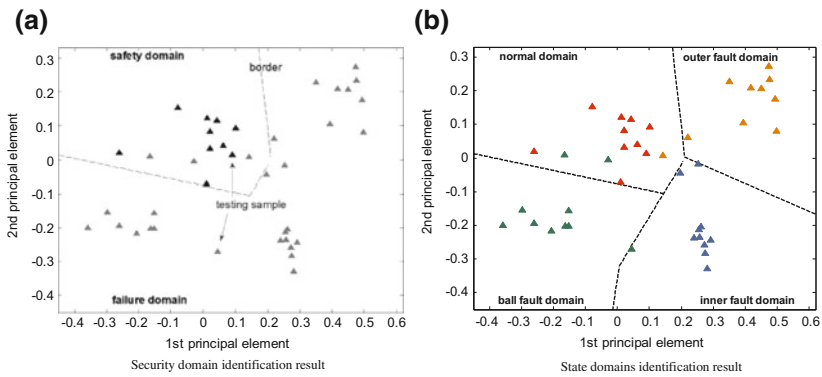


Fig. 7 State domain identification result

expresses fault state, with navy blue inner race fault state and dark green ball fault state. Subfigures (a) and (b) of Fig. 7 show the state differentiating results. The first subgraph is the result of security state recognition and the second one appears all the states distinguished result. The statistical result shown in Table 1 reflects that applying MRCE in MRSD to identify rolling bearings safety state.

Table 1 Security state identification result

Sample state	Right testing amount/total amount	Accuracy (%)
Security	30/30	100
Non-security	87/90	96.7

6 Conclusion

This paper depicts the space model of characteristics—state, and within the framework proposes a method of the state domain identification. Based on wavelet packet transform, we combine the MRCE and PCA algorithm to get the elements of space. Then, MRSD is constructed in two-dimensional Euler space to complete the status identification of testing samples, and the results prove the effectiveness of the proposed method. Method of domain boundary determination is still in researching stage at present. In the future, clustering and optimization algorithms will be analyzed in focus, and a large number of experimental data will be summarized, in which the two research aspects are aimed at efficient calculating method of domain boundaries are obtained.

Acknowledgements The authors would like to give appreciation to the editor and anonymous reviewers. This work was supported by National Key Research and Development Program of China (No. 2016YFB1200505).

References

1. Wang B, Li H, Xu B (2013) Extraction of performance degradation feature for motor bearings based on multi scale morphological decomposition spectrum entropy. *Bearing* 8:43–47
2. Al-Bugharbee H, Trendafilova I (2016) A fault diagnosis methodology for rolling element bearings based on advanced signal pretreatment and autoregressive modelling. *J Sound Vib* 369:246–265
3. Luo Z, Xue X, Wang X (2005) Study on the method of incipient motor bearing fault diagnosis based on wavelet transform and EMD. *Chinese Soc Electr Eng* 25(14):125–129
4. Li Y, Liang X, Xu M, Huang W (2017) Early fault feature extraction of rolling bearing based on ICD and tunable Q-factor wavelet transform. *Mech Syst Sig Process* 86(Part A):204–223
5. Zhang Y, Qin Y, Jia LM Research on methodology of security region estimation of railway system operation safety assessment. In: *Proceedings of world congress on engineering and technology*. Shanghai, pp 803–807
6. Peng Z, Peter WT, Chu F (2005) A comparison study of improved Hilbert-Huang transform and wavelet transform: application to fault diagnosis for rolling bearing. *Mech Syst Signal Process* 19(5):974–988
7. Li X, Fu Y, Jia L (2012) Fault diagnosis of railway axlebox bearing based on wavelet packet and neural network. *Appl Mech Mater* 226–228(1):749–755
8. Jiang H, Chen J, Dong G (2016) Hidden Markov model and nuisance attribute projection based bearing performance degradation assessment. *Mech Syst Signal Process* 72–73:184–205
9. Zhang Y, Zuo H, Bai F (2013) Classification of fault location and performance degradation of a roller bearing. *Measurement* 46(3):1178–1189

Analysis of EMUs Vehicle Body Voltage Caused by Pantograph-Catenary Off-Line Arc

Ying Wang, Fengyang Gao, Xiuqing Mu, Qiyao Li, Xiaoqiang Chen and Shiwen Liang

Abstract Aiming at the inevitable pantograph-catenary off-line arc when electric multiple units (EMUs) run, the CRH3 EMUs, which operate at China Chongqing–Lichuan line, are selected as the research subject. Vehicle body voltage caused by the pantograph-catenary off-line arc is analyzed. First, the structure of the vehicle body and CRH3 EMUs grounding system are studied. Second, based on PSCAD/EMTDC software, electric parameters of traction network and CRH3 EMUs are ascertained, then a coupling model of vehicle-grid system is established, and an extended pantograph-catenary arc model is utilized to simulate pantograph-catenary off-line when vehicles run. Through the pantograph-catenary off-line model, vehicle body overvoltage distribution is analyzed. Finally, aiming at the CRH3 EMUs grounding system, corresponding methods are proposed to suppress vehicle body overvoltage.

Keywords Electric multiple units (EMUs) · Pantograph-catenary off-line arc
Vehicle body · Overvoltage suppression

1 Introduction

With the rapid development of electrical railways, more and more attention is paid to the safe and steady operation of the EMUs. overvoltage phenomenon is always caused by raising or dropping pantograph, passing neutral section, and pantograph-catenary off-line [1], which will result in serious fluctuation of vehicle body potential and overvoltage between vehicle body and rail [2].

The reason for producing vehicle body surge overvoltage about Japanese AC vehicle was analyzed in [3, 4]; combined with experiment, improved methods were discussed to avoid extreme overvoltage phenomenon. For the phenomenon of

Y. Wang (✉) · F. Gao · X. Mu · Q. Li · X. Chen · S. Liang
School of Automation & Electrical Engineering, Rail Transit Electrical Automation
Engineering Laboratory of Gansu Province, Lanzhou Jiaotong University,
No. 88 West Anning Road, Lanzhou, China
e-mail: wangying01@mail.lzjtu.cn

vehicle body overvoltage and its uneven distribution, existing vehicle body grounding method of the CRH380BL EMUs was improved [5]. The equivalent circuit analysis model about the process of dropping pantograph was built in [6, 7], characteristics of vehicle body overvoltage were analyzed, the effect, which result from grounding resistance, transformer, voltage phase at the dropping pantograph moment, on overvoltage amplitude was studied. In [8], the electric circuit model was established to analyze the transmitting and reflecting characteristics between different car bodies, as well as experiments were carried out in field.

Generally, many researches on the vehicle body overvoltage are carried out, but the impact of the pantograph arcing is hardly studied. In this paper, based on China Chongqing–Lichuan line, the CRH3 EMUs vehicle body overvoltage is studied first. Second, with actual pantograph-catenary arc phenomenon taken into account, the CRH3 EMUs coupled vehicle-grid system considering the pantograph-catenary off-line arc is built. Third, the specific voltage waveform and distribution law in different vehicle bodies are analyzed. Finally, several measures of restraining overvoltage are put forward, which have theoretical and practical significance for the safe operation of the CRH3 EMUs in Chongqing–Lichuan line.

2 Equivalent Modeling

When high-speed train collect or stop collecting current from catenary, the arc will occur because of the pantograph-catenary off-line, this transient phenomenon spread easily in the vehicle body and cause overvoltage. For CRH3 EMUs, overvoltage will through the high-voltage cable core on vehicle roof, coupled with the shielding layer; then, the overvoltage can enter every vehicle body through the shield grounding and spread among the vehicle bodies. Thus, vehicle body voltage will obviously increase instantly.

There are two main kinds of vehicle body grounding methods, i.e., working grounding and protection grounding. Grounding layout of CRH3 EMUs vehicle body is depicted in Fig. 1. The working grounding is set at Nos. 2 and 7 vehicle bodies, and the protection grounding is set in the middle vehicle, i.e., Nos. 4 and 5 vehicle bodies. In order to analyze the CRH3 EMUs vehicle body overvoltage under the pantograph-catenary off-line arc, corresponding coupled vehicle-grid model should be established.

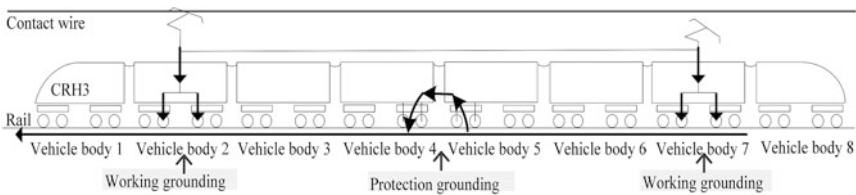


Fig. 1 Vehicle body and grounding diagram of CRH3 EMUs

2.1 Determination of Parameters

The electric parameters of coupling model mainly include the parameters of traction network and vehicle body impedance of the CRH3 EMUs. The model used in this paper is based on Chongqing–Lichuan line railway traction supply system and the running CRH3 EMUs.

First, the traction network is mainly formed by three parts: the traction substation, the catenary, and the rail. During the process of establishing Chongqing–Lichuan line, single-phase 27.5 kV AC voltage and equivalent substation impedance are taken into account. In order to calculate the equivalent impedance, the actual traction transformer parameters are listed in Table 1.

Based on the parameters, the equivalent impedance of traction transformer is gained as $0.177 + j9.97 \Omega$. Applying with multiconductor method [9], the parameters of Chongqing–Lichuan traction network are shown in Table 2.

In addition, CRH3 EMUs vehicle body parameters mainly consist of two parts: the high-voltage cables and the vehicle body. The high-voltage cable is used to connect two pantographs on Nos. 2 and 7 vehicle bodies. The cable is mainly composed of cable core and grounding shield layer, and the whole high-voltage cables can be equivalent to the T circuit. According to the actual cable parameters, resistance inductance and capacitance of unit length cable are obtained as $1.213 \times 10^{-5} \Omega/\text{m}$, $2.197 \times 10^{-4} \text{mH}/\text{m}$, and $4.1162 \times 10^{-4} \mu\text{F}/\text{m}$. Considering the distance between two pantographs as 75 m, it can be calculated that the cable resistance, inductance, and capacitance are 0.90975 m Ω , 0.0165 mH, and 0.061743 μF . Direct working grounding method is utilized in CRH3 EMUs [2], and the specific parameters are shown in Table 3.

Table 1 Parameters of Chongqing–Lichuan railway traction transformer

Parameters	Value	Parameters	Value
Rated capacity	50,000 kVA	Load loss	146.488 Kw
Rated voltage	220/2 \times 27.5 kV	Short-circuit impedance	16.48%
Rated current	227.27/909.08 A	Short-circuit current	0.23%
No-load loss	43.566 kW		

Table 2 Parameters of Chongqing–Lichuan railway traction network

Parameters	Value	Parameters	Value
Contact wire type	GLCN-250	Rail type	60 kg/m
Equivalent radius	8.92 mm	Equivalent radius	12.79 mm
Unit resistance	0.149 Ω/km	Unit resistance	0.135 Ω/km
Messenger wire type	GJ-100	Track gauge	1.435 m
Equivalent radius	5.64 mm	Contact wire height of	6 m
Unit resistance	1.45 Ω/km	Structural height	1.3 m
Line relaxation	0.6 m	Earth resistivity	$10^{-4} \Omega \text{ cm}$

Table 3 Parameters of CRH3 EMUs vehicle body

Parameters	Value (mΩ)	Parameters	Value
Carbon brush resistance	5.65	Unit length resistance	0.009 mΩ
Wheel-rail resistance	0.2	Vehicle-rail unit capacitance	1.672×10^{-11} F
Vehicle ground resistance	5.85	Body connection resistance	11.2 mΩ

2.2 Establishment of Vehicle-Grid Model

Based on the former work, the electrical parameters of traction network and CRH3 EMUs vehicle body are already determined. Then, the vehicle-grid system simulation model, which can produce the pantograph-catenary off-line phenomenon, can be established at PSCAD/EMTDC platform. The whole model is shown in Fig. 2. Where pantographs are installed on Nos. 2 and 7 vehicle bodies, respectively, but only pantograph on No. 2 vehicle body collect current from catenary, working grounding is located in Nos. 2 and 7 vehicle bodies, and protection grounding is at Nos. 4 and 5 vehicle bodies. In CRH3 EMUs, low-voltage equipment is always equipped with overvoltage protection device, so these low-voltage equipment can be ignored when analyzing voltage distribution in every vehicle body.

In Fig. 2, the model mainly includes the simplified equivalent catenary model, the high-voltage cable model, the vehicle body impedance model, the EMUs working grounding model, the protection grounding impedance model, and the rail impedance model. Note that the pantograph-catenary off-line is realized through the breaker and the extended pantograph-catenary arc model [1]. The specific parameters are as same as the former parameters determination.

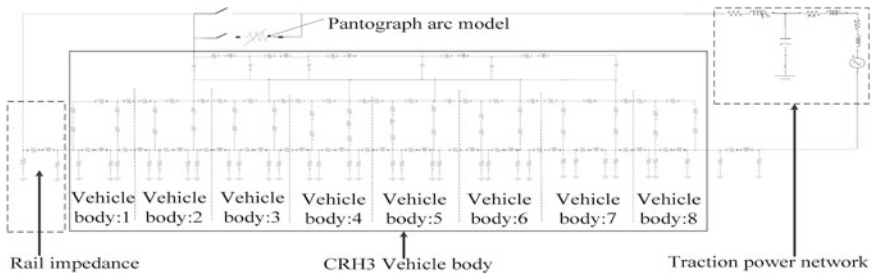


Fig. 2 Coupled vehicle-grid system in pantograph-catenary off-line arc

3 Analysis of CRH3 Vehicle Body Voltage

Based on vehicle-grid simulation model in Fig. 2, the vehicle body voltage variation caused by the pantograph-catenary arc of every vehicle body is obtained. Setting two measure points at every vehicle body during the specific simulation test, and these two measure points locate at the head and end part of each vehicle body, respectively, after obtaining two measured voltages, take the larger values as vehicle body voltage. Specific test results are shown in Fig. 3. The simulation result is shown in Fig. 3a. In [8], the actual measurement result is presented in Fig. 3b. These two comparison results confirm the validity of simulation model in Fig. 2.

It can be seen that head vehicle body voltage increase steeply when pantograph-catenary arc occurs in Fig. 3a. The head vehicle body overvoltage increase to 2428 V, but the continued time is short. Figure 3c shows the overall voltage distribution of the eight vehicle bodies. It can be observed that the voltage value is low in the middle and high voltage locate at two sides. This is because the centralized grounding method is adopted in CRH3 EMUs vehicle body, protection grounding is set in Nos. 4 and 5 vehicle bodies. The corresponding instantaneous maximum overvoltage is lower than others. In addition, no grounding points are set in the CRH3 EMUs head and end vehicles. So, the overvoltage in Nos. 1 and 8 vehicle bodies is a superposition of traveling wave and reflected wave. It also results in some voltage increase in adjacent vehicle body. Based on China railway standard TB/T3021-2001, the maximum overvoltage that electronic equipment in

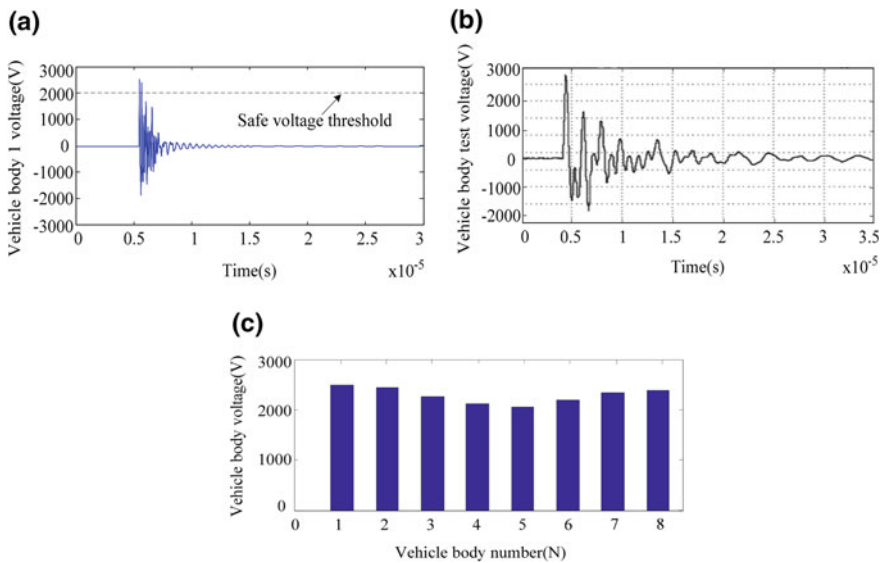


Fig. 3 CRH3 vehicle body voltage distribution in pantograph-catenary off-line arc: **a** head vehicle body voltage through simulation, **b** vehicle body voltage through measurement, and **c** vehicle body voltage distribution from No. 1 to No. 8

vehicle is 2000 V [3], but the most of vehicle body voltage are higher than 2000 V; it will lead to instantaneous attack to the reliable operation of the CRH3 EMUs.

4 Suppression Measures of CRH3 Vehicle Body Overvoltage

After the voltage distribution of every CRH3 vehicle body, when the pantograph-catenary off-line arc occurs, is analyzed. The following part of this paper will analyze the overvoltage suppression of the vehicle body by optimizing the grounding layout and optimizing the grounding method.

On the one hand, we can optimize vehicle body grounding layout. Here, two kinds of grounding layout are mainly considered. First, we can set protection grounding in the CRH3 vehicle body which contains working grounding. Second, the protection grounding for every CRH3 vehicle body can be set.

Figure 4 shows the result of setting protection grounding in the CRH3 vehicle body which contains working grounding. It can be found that, after protection grounding is adding to the vehicle body (Nos. 2 and 7 vehicle bodies) which contains working grounding, the maximum overvoltage is obviously suppressed because of the increase of leak current point number. It can be observed that the peak voltage of Nos. 2 and 7 vehicle bodies is extremely low, and the maximum overvoltage of Nos. 1 and 8 vehicle bodies is about 950 V.

Figure 5 reveals the result of setting the protection grounding for every CRH3 vehicle body. In Fig. 5, after adding protection grounding for the CRH3 head and end vehicles (Nos. 1 and 8 vehicle bodies), the suppression result for overvoltage is

Fig. 4 Comparison of CRH3 vehicle body overvoltage distributions

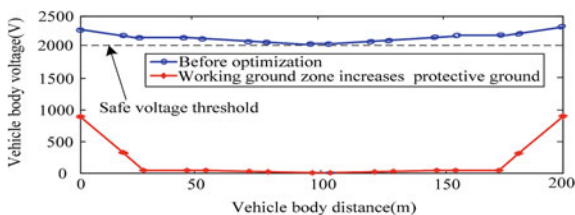
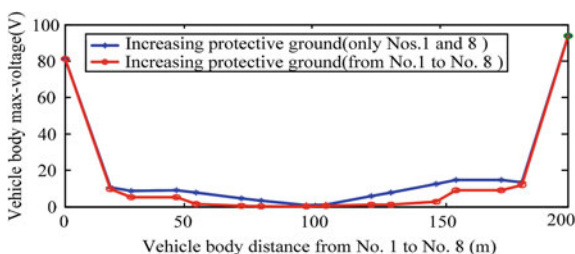


Fig. 5 Comparison of CRH3 vehicle body overvoltage distributions



demonstrated by the blue curve, the maximum overvoltage are just 81.6 and 93.8 V, respectively. In Fig. 5, the red curve demonstrates the suppression effect of the overvoltage in the CRH3 traction vehicle (Nos. 1, 3, 6, and 8). It can be found that the two methods have almost same remarkable suppression result for vehicle body overvoltage suppression.

On the other hand, we can optimize vehicle body grounding modes. Vehicle body voltage distribution is analyzed when different inductances were applied to the work grounding. Here, take inductance as 10, 1, 0.1, and 0.01 mH to test vehicle body voltage. The results are displayed in Fig. 6.

In Fig. 6, blue bar is the vehicle body voltage distribution (i.e., Fig. 3b) before optimizing grounding methods. Red bar represents the vehicle body voltage distribution after adopt optimizing grounding methods. According to the comparison between Figs. 3 and 6b, it can be gained that when the inductance value is big (grounding inductance is 10 mH) in working grounding, the vehicle body overvoltage is not suppressed, but become higher. But when the grounding inductance value is small (grounding inductance is 1, 0.1, or 0.01 mH), the overvoltage peak value of every vehicle body can be effectively reduced. And Fig. 6 shows the vehicle body voltage all stay below the safe value, i.e., 2000 V. Thus, the influence caused by vehicle body overvoltage on equipment inside the vehicle is reduced.

Also, we can parallel small resistance at the working grounding point to reduce the equivalent grounding resistance. Here, 0.5, 0.005 Ω resistance in the simulation to test vehicle body voltage can be selected. The comparison results before and after optimization are shown in Fig. 7.

In Fig. 7, the blue bar represents each CRH3 vehicle body voltage distribution before optimization (i.e., Fig. 3b), and the red bar is the voltage distribution after

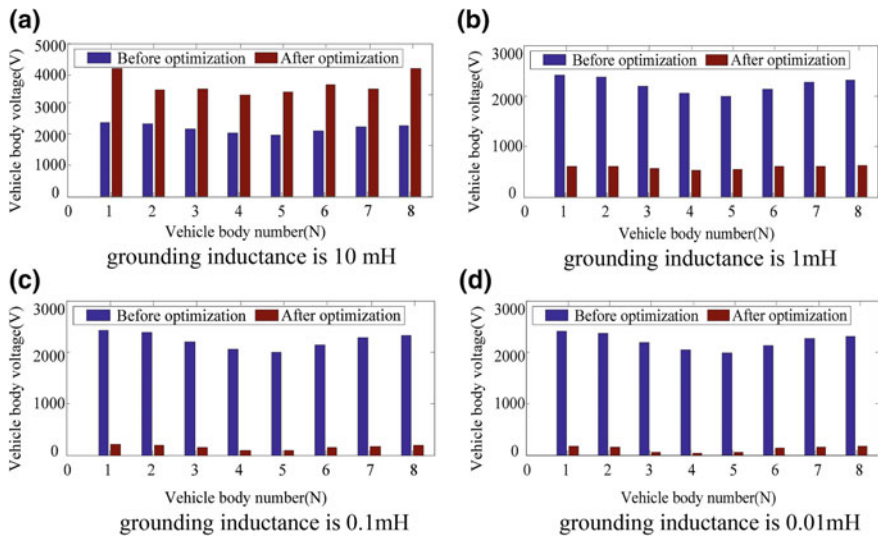


Fig. 6 Comparison of vehicle body overvoltage under different grounding inductances

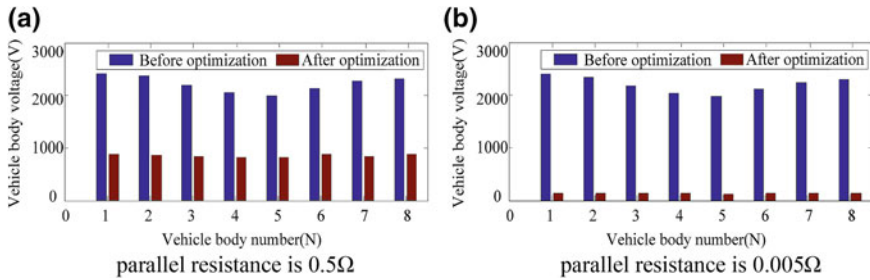


Fig. 7 Comparison of vehicle body overvoltage under different grounding resistances

optimization. As shown in Fig. 7, when small resistance is parallel with carbon brush in working grounding. The vehicle body voltage shows a decrease trend in each vehicle (stay below the safe value 2000 V). Especially in the case that 0.005 Ω resistance is adopted, compared with the result before optimization, every vehicle body overvoltage significantly decreased.

Through the above analysis, it can be gained that vehicle body overvoltage can be effectively suppressed by optimizing grounding methods. Adopting small inductance in working ground points or paralleling resistance to the working grounding is effective and reasonable.

5 Conclusion

China Chongqing–Lichuan line vehicle body model is completely established, and pantograph-catenary off-line arc of CRH3 EMUs is taken into account in the model. Based on vehicle body voltage distribution, two kinds of methods are obtained to suppress the vehicle body overvoltage. One is to increase the number of protection grounding point, and the other is to improve work grounding characteristics. The comparison results show that the proposed suppression methods are reasonable to avoid CRH3 EMUs vehicle body overvoltage in the pantograph-catenary off-line arc occurs at Chongqing–Lichuan line.

Acknowledgements This work was supported by the National Natural Science Foundation of China (51767013), the National Key Research and Development Program of China (2017YFB1201003), the Science and technology research and development plan of China Railway Corporation(2017J012-A), the Natural Science Foundation of Gansu Province Education Department (2017A-020), and the Science and Technology Program of Gansu Province (17JR5RA083).

References

1. Wang Y, Liu Z, Mu X et al (2016) An extended Habedank's equation-based EMTP model of pantograph arcing considering pantograph-catenary interactions and train speeds. *IEEE Trans Power Delivery* 31(3):1186–1194
2. Huang K, Liu Z, Zhu F et al (2017) A systematic EMTP impedance modeling scheme aimed at train body in high-speed railway. *IEEE Trans Transp Electrification* 3(1):272–283
3. Hatsukade S (2009) Reduction method of surge voltage on AC railcar body. *Q Rep RTRI* 50(2):70–75
4. Hatsukade S, Fujino K, Kanno S (2013) Suppressing over-voltage on carbodies of a high speed EMU. In: *Proceeding of 2013 15th European conference on power electronics and applications*, Lille, France, pp 1–9, 2–6 Sept 2013
5. Li X, Wang J, Wu H et al (2013) Optimization research of the earthing design for CRH380BL EMU. *Electr Locomotives Mass Transit Vehicles* 36(5):19–26 (in Chinese)
6. Yang S, Cao B, Gao G et al (2015) Analysis of surge overvoltage of dropping pantograph for high-speed EMUs. *J China Railway Soc* 37(7):46–50 (in Chinese)
7. Huang k, Liu Z, Wang Y et al (2016) Analysis on railcar's body over-voltage distributional characteristics considering operating conditions of high-speed railway station. *J China Railway Soc* 38(9):38–45 (in Chinese)
8. Nie Y, Hu X (2013) Research on rising pantograph surge overvoltage for high-speed EMUs. *Electric Drive Locomotives* 12(4):9–11 (in Chinese)
9. Mu X, He Z, Wang Y et al (2015) Analysis of harmonic current model in traction power supply network. In: *Proceedings of the 2015 international conference on electrical and information technologies for rail transportation*, Zhuzhou, China, vol 377, pp 199–207, 17–19 Oct 2015 (in Chinese)

Safety Analysis of Train Control System Based on Colored Petri Nets and System-Theoretic Process Analysis

Shaoqiang Hu, Daohua Wu and Huashen Wang

Abstract This paper proposes a new method called formalSTPA which combines Colored Petri Nets (CPN) and System-Theoretic Process Analysis (STPA). Using formal method to extend the STPA includes establishing the system hierarchical control structure and transforming it into CPN model. First we use ASK-CTL formula to test and verify the CPN model, and then Programming Language (ML) is used to identify the cause of hazards and realize safety analysis of system functions. ML can be used to identify hazards effectively. In this paper, the formalSTPA is applied in Chinese Train Control System Three Level (CTCS-3). The result shows that the formalSTPA is suitable for the safety analysis of CTCS-3.

Keywords CTCS-3 · STPA · CPN · Safety risk analysis

1 Introduction

As an important part of system safety discipline, the safety risk theory considers system safety from the perspective of risk which represents the latest concept and development direction of the system safety engineering. The basic idea of safety risk theory is to identify various hazards carefully and analyze the safety risk level scientifically to understand and master the risk, and then take effective control actions to eliminate or reduce harmfulness of hazards [1]. The traditional method of safety risk identification is mainly based on the event chain model. It describes that the accident contains the root event and a series of intermediate events, and the relationship between all the model elements is linear. Second, the traditional analysis method is used to ensure system safety when the design is completed,

S. Hu · D. Wu (✉) · H. Wang
Beijing Key Laboratory of Security and Privacy in Intelligent Transportation,
National Research Center of Railway Safety Assessment, Beijing Jiaotong
University, No. 3 Shang Yuan Cun, Haidian District, Beijing 100044, China
e-mail: wudaohua@bjtu.edu.cn

which is not suitable for the complex system and the cost will increase in follow-up system modification procedure [2].

The structure of train control system is complex, and the subsystem is interdependent, many tasks are executed at the same time, and the process is complicated. If simply apply the traditional methods of safety risk identification, it is easy to ignore the hidden hazards in the process of software execution and state transition which affects the comprehensiveness and accuracy of identification results. Leveson [3] proposed a new method called System-Theoretic Accident Modeling and Process (STAMP) in 2002, and then proposed the System-Theoretic Process Analysis (STPA) based on STAMP in 2011. STPA can be used to analyze the safety of complex systems. Ouyang [4] used STAMP to analyze the accident of 4.28 Jiaoji Line in China and simulated the message propagation process. The author only established hierarchical control structure model, but it cannot simulate the direction of information transmission and dynamic process. Liu [5] used formal method to extend STPA, established the extended UML model, and transformed it into PHAVer model. The reachable set is used to analyze the PHAVer model to identify the factors leading to inadequate control. However, the analyst should be familiar with the system and has rich practical experience. Spiegel [6] combined STAMP with CPN, presented a formalSTAMP, and applied formalSTAMP to Wenzhou 7.23 accidents to verify the advantages of this method. But he did not come up with a specific risk analysis log. Wang [7] combined STPA with CPN and used CPN tools to generate reachable graphs. The reachability graph is used to analyze hazards. Finally, take the release scenario of temporary speed restriction as an example and generate the specific hazard log of this scenario. The emphasis of this paper is to identify hazards using the hierarchical control structure model; therefore, the model has not been verified.

The emphasis of this paper is to use the CPN hierarchical control structure to identify system hazards and verify the consistency of model function with actual system. CPN supports branch temporal logic formula of ASK-CTL which can be effectively applied in formal verification. In this paper, we can find a more credible and more accurate path by using ML to trace the cause of hazard. Compared with the original control structure model of block diagram, the formal method is more rigorous, and CPN not only standardizes model but also reflects the interaction between subsystems.

2 Basics

As a kind of system model language, Petri nets can not only describe system structure but also describe the dynamic behaviors and state transition. Petri nets have rigorous mathematical definition and visually graphical expression [8]. The system analysis method of CPN has many advantages over other modeling methods. For complex systems, Colored Petri Nets can be layered and refined gradually.

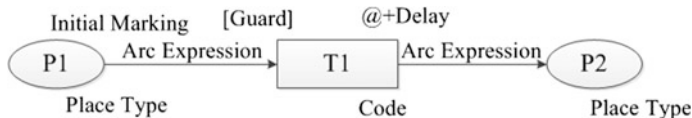


Fig. 1 An abstract CPN model

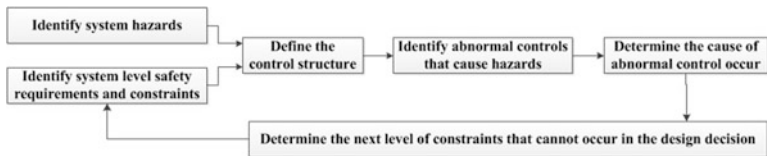


Fig. 2 The main step of STPA

The description of transition and place of CPN is shown in Fig. 1. The place represents a possible state and the transition represents a probable action. The arc function is used to indicate that the marking is increased or consumed because a transition has occurred. The *guard* is a function which is a Boolean expression to determine whether a transition is allowed or restricted that triggers a transition.

STPA considers safety from the perspective of system. The goal of STPA is to identify inadequate controls that may lead to risk and determine the relevant safety constraints. According to the use of different time, it provides necessary information and documentation to ensure that the safety constraints are implemented in the system design and operation. The main steps of STPA are shown in Fig. 2.

The basic idea of model verification is to represent the system behaviors by state transition and describe the system characteristic with modal temporal logic. However, in order to improve CPN verification capabilities, Cheng et al. [9] combine CPN with model verification technique and then introduce the ASK-CTL into CPNTools. ASK-CTL is explained by the model state space and it has a recursive definition of state formula and transition formula; we can explain state sequence and transition sequence by state space.

3 FormalSTPA

The model of classical hierarchical control structure is block diagram, but the description ability of block diagram is insufficient. In addition, the whole analysis process and results are literal description; it is difficult to describe the specific behavior and control process systematically. The model is not executable and lacks the support of data structure; it is short of adaptability and portability of other platforms. The flexibility of CPN model can describe the dynamic behaviors and

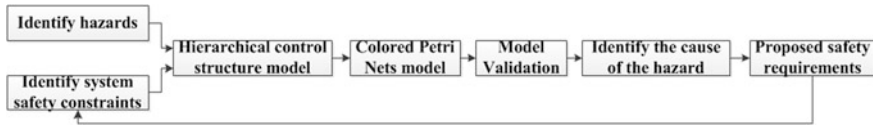


Fig. 3 The hazard identification process of formalSTPA

state changes [10]. Therefore, this paper considers transforming block diagram into CPN model. Under the premise of ensuring the model correct, ML is used to find the basic reason of hazard. The general framework of hazard identification method based on formalSTPA is shown in Fig. 3.

- Identify hazards: Obtain system hazards which can cause equipment damage, environment pollution, and even casualties.
- Identify safety constraints: After identifying the system hazards, the safety constraints can be identified, which prevent these hazards from occurring.
- Hierarchical control structure model: According to safety constraints, a preliminary control structure is designed to meet the system requirement, and then the control structure should be improved in design refinement.
- Petri nets model: The control structure model is transformed into CPN model. CPN describes the system behaviors and details of internal interaction. The model conforms to the hierarchical control mechanism proposed by STPA.
- Model verification: CPN supports branch temporal logic formula of ASK-CTL which can be well applied in formal verification.
- Identify the cause of hazards: Use the state space function to find the state node containing malfunction and use ML to find the reachable path from initial state node to fault state node.
- Propose safety control requirement: Put forward safety constraints to ensure the system safety.

4 An Application to the CTCS-3

4.1 CPN Hierarchical Control Structure Model of CTCS-3

The advantage of STPA lies in the comprehensive consideration of the accident cause. Leveson [3] proposed the hierarchical control structure in 2002, and each layer imposes a constraint on the lower layer and the upper layer safety constraint controls the lower layer behaviors. The hierarchical control structure model in this paper consists of the top-level model of train control system, the process model, and the control subsystem model. Figure 4 describes the typical control structure of high-speed train control system.

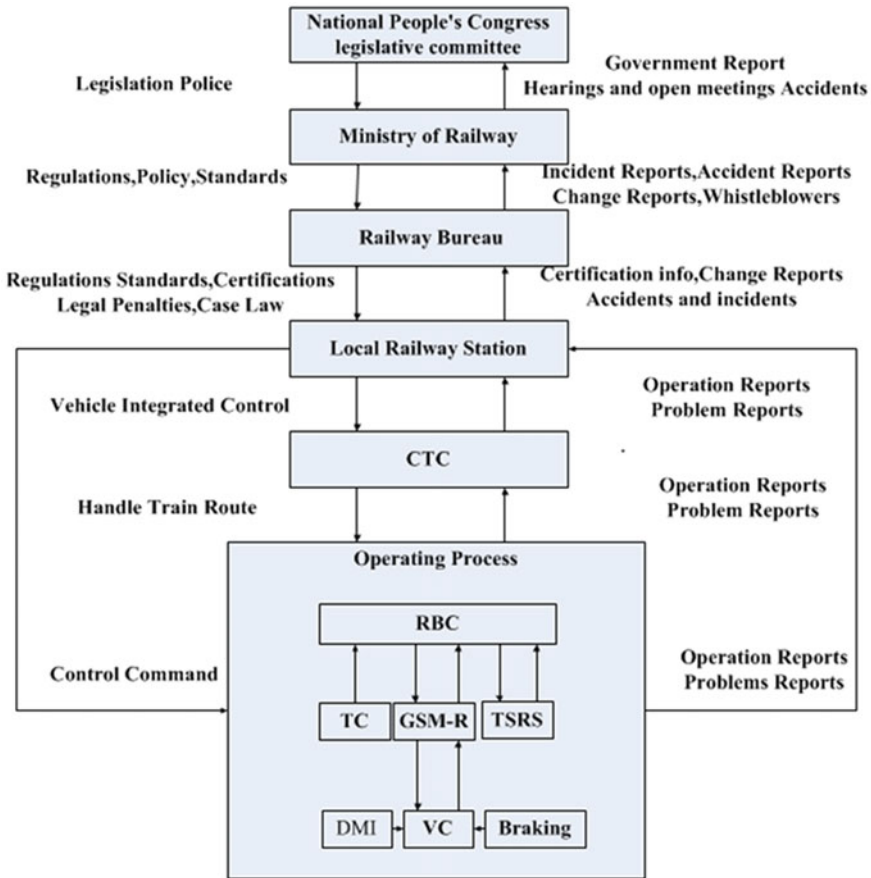


Fig. 4 Typical control structure of CTCS-3

CTCS-3 is based on the wireless communication and it is the highest level of train control system in operation. Track Circuit (TC) is used to detect train occupancy, and the balise is used to locate trains. Temporary Speed Restriction Server (TSRS) issued temporary speed restriction information to Radio Block Center (RBC). When RBC receives movement authority request and location reports from Vehicle On-Board Controller (VOBC), Movement Authority (MA) is generated according to the available train route data provided by the Computer Interlocking (CI), then the route data and temporary speed restriction will be sent to Vital Computer (VC) through GSM-R wireless communication platform. VC calculates target distance monitor curve according to received control data and braking performance. By comparing the train speed with the monitor curve, the braking system will ensure train safety when train is overspeed and provide the necessary feedback information to driver [11]. The train operation model is shown in Fig. 5.

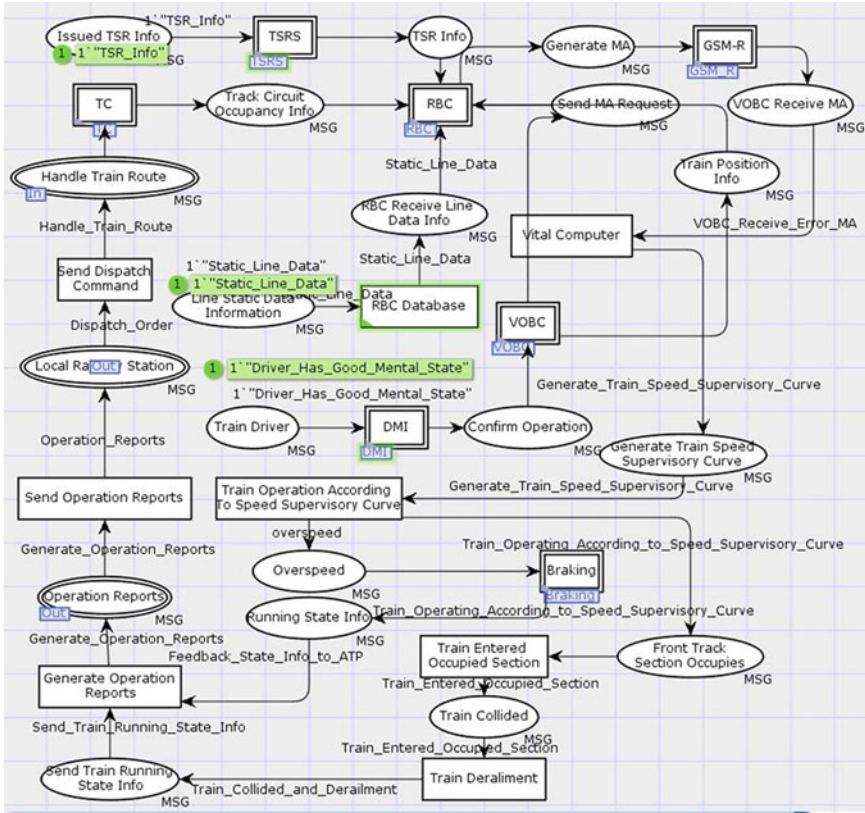


Fig. 5 The process model of CTCS-3

The CPNTools has a toolbox that calculates the model state space and generates state space reports. It supports a variety of standard queries and we only need to write a simple function instead of writing complex codes. For the CPN model, the boundedness properties, home properties, liveness properties, fairness properties, and reachability properties are inspected by writing simple functions. If the CPN model marking is liveness, any transition may remain in a state that may be triggered. If the CPN has dead marking, it is necessary to discuss whether the model has deadlock and then prove that the dead marking is safety state and reasonable. The test results of the model properties are shown in Table 1.

Table 1 Model performance verification and results

Model properties	Verification results
Liveness properties	List of dead markings: 13,572
Dead marking	use code “print(NodeDescriptor 13,572)” to get details
Home properties	The initial marking is not a home state
Boundedness properties	All places are bounded

4.2 Analysis of the Causal Factors

The concept of safety is that there is no unacceptable risk, and there is no absolute safety. Generally speaking, the main types of railway accidents are rear-end collision, head collision, and train collision with the surrounding buildings and derailment. For the scenarios involved in this paper, rear-end and overspeed are more likely to cause the train to derail. As a result, system-level hazards can be reasoned forward by these two types of accidents [10]. In this paper, the dangerous state is defined in Table 2, and these states will cause the train overspeed or derailment.

Through simulation and dynamic verification, all the possible scenarios of system model are executed step-by-step, and the display result is compared with the actual system. The model’s reachability graph is large and it is inconvenient to list all of them, therefore, ML is used to identify model hazards and the path to unsafe state. The summary information is stored in a file called DangerousState, and the results are shown in Table 3. We can see all places state and token information which reach unsafe path. Finally, the token information contained in each place and the transition directions are compared with actual system, and the model is consistent with the actual system. In addition, we can observe tokens passing process and state transition direction through the reachability graph, so as to trace the potential cause of the dangerous state. If all transitions are no longer authorized to trigger at any time, the marking is called dead marking. Since the dead marking includes the model’s safe and unsafe state, you can first view the dead marking node that represents the unsafe state and then search for the path to unsafe state. There is one dead marking in this model, namely, node 13,572 which can be seen from Table 1. We can search the path to the dead marking after explicating the specific content of the dead markings. Table 3 gives a specific algorithm for calculating the path to unsafe state. This paper summarizes the results of the implementation due to length limitations.

Table 2 Definition of dangerous state

Place name	Marking
Train Derailment	Train Collided and Then Derailment
Track Circuit Occupancy Information	Error TC Occupancy Info
Link Failure	Link Failure
BTM Failure	BTM Failure
Switch not Transmit into Specified State	Switch not Transmit into Specified State
Switch Lost Indication	Switch Lost Indication
Train Enter the Occupied Block	Not Detect Train in Block
Data Communication Network is Abnormal	Data Communication Network is Abnormal
Generate MA	Generate Error MA
Overspeed	Overspeed

Table 3 An algorithm for searching the path to dangerous state

An algorithm for searching the path to dangerous state
ML codes
<pre> fun risk n = (Mark.TC'Train_Derailment 1 n = ["Train Collided and Then Derailment"] orelse Mark.TC'Switch_Lost_Indication 1 n = ["Switch Lost Indication"] orelse Mark.TC'Track_Circuit_Occupancy_Information 1 n = ["Error TC Occupancy Info"] orelse Mark.TC'Train_Enter_the_Occupied_Block 1 n = ["Not Detect Train in Block"] orelse Mark.TC'Switch not Transmit into Specified State 1 n = ["Switch not Transmit into Specified State"] orelse Mark.TSRS'Data_Communication_Network_is_Abnormal 1 n = ["Data Communication Network is Abnormal"] orelse Mark.RBC'Generate_MA 1 n = ["Generate Error MA"] orelse Mark.GSM_R'Link_Failure 1 n = ["Link Failure"] orelse Mark.Braking'Overspeed 1 n = ["overspeed"] orelse Mark.VC'BTM_Failure 1 n = ["BTM Failure"] fun DangerousState() = PredNodes(ListDeadMarkings(),fn n =>risk n,NoLimit); let val fid = TextIO.openOut"DangerousState2.txt" val _ = TextIO.output(fid,"Path to a dangerous state:\n") val _ = EvalArcs(ArcsInPath(1,hd(DangerousState())), fn a =>STRING.output(fid,st_BE(ArcToBE a))) in TextIO.closeOut(fid) end </pre>
Results: Path to a dangerous state
<ul style="list-style-type: none"> •"DMI'Driver_Has_Good_Mental_State 1: {msg53 = "Driver Not Concentrate On DMI"}" •"TSRS'Send_TSR_Info_to_RBC 1: {msg25 = "Data Communication Network is Abnormal"}" •"TSRS'Transmit_Failure 1: {msg25 = "Data Communication Network is Abnormal", msg26 = "Transmit Failure and RBC cannot Receive TSR"}" •"TC'Check_the_Occupancy_of_the_Block 1: {msg13 = "Block is Occupied"}" •"TC'Generate_Interval_Route_Info 1: {msg17 = "Switch not Transit into Specified State", msg74 = "Switch Lost Indication"}" •"RBC'Read_Relevant_Info 1: {msg19 = "Error TC Occupancy Info",msg26 = "Transmit Failure and RBC cannot Receive TSR"}" •"GSM_R'Transmit_MA_by_GSM_R 1: {msg42 = "Generate Error MA",msg44 = "GSM-R Breakdown"}" •"DMI'ATP_Implement_Emergency_Braking 1: {msg55 = "Driver Not Response", msg58 = "ATP Implement Emergency Braking"}" •"TC'Not_Detect_Train_in_Bloc 1: {msg13 = "Block is Occupied",msg15 = "Not Detect Train in Block"}" •"TC'Two_Train_Collision 1: {msg15 = "Not Detect Train in Block",msg16 = "Train Collided and Then Derailment"}" •"VOBC'Unable_to_Get_Train_Position_Info 1: {msg32 = "BTM Failure",msg34 = "Unable to Get Train Position Info"}" •"VOBC'Train_Stop 1: {msg34 = "Unable to Get Train Position Info",msg35 = "Train Braking until Stop"}"

As can be seen from Table 3, there are two factors that lead to dangerous state. On the one hand, the RBC is not able to receive temporary speed restriction information issued by the TSRS because of the interruption of the data communication network. The train is overspeed in some place due to no temporary speed restriction information. On the other hand, the track circuit shows error occupancy information and switch in wrong position; the train enters the occupied section and

Table 4 Safety constraints corresponding to unsafe control behavior

Unsafe control behavior	Safety constraints
Drivers do not have a good mental working condition	Ensure that the employees have a good sleep and mental state
The transmission of signal system data network is interrupted	The network states require monitoring in real time
The track circuit occupancy is abnormal	Periodic inspection to prevent the track circuit from being in a state of illness
The switch is not converted to the specified condition	The switch state monitoring system monitors the switch state information in real time, and finds out the abnormal and alarm in time to solve the problem of switch maintenance
When GSM-R malfunctions, the communication between the on-board equipment and ground equipment is interrupted	Optimize handover algorithm and improve coverage of wireless network
When balise transmission module malfunctions, it is unable to determine train position	Improve demodulation method of balise transmission module

then collision lead to derailment. The unsafe control behavior identified by the above algorithm can be transformed into the safety constraint associated with the train control system, as shown in Table 4. The system safety constraints are redefined to ensure the system safety. After the above analysis, it shows that the results are consistent with common sense and verify the rationality and practicability of formalSTPA method.

5 Conclusion

Because the description of the system control structure model is insufficient, and the description of block diagram is literal, it is difficult to describe the specific behavior and system process behaviors, and the model is not executable. But CPN model can describe the dynamic behavior and state transition due to its flexibility. Therefore, CPN and STPA are combined to propose a formal method called formalSTPA in this paper. The formal model of CPN is used to strengthen and supplement the hierarchical control structure model in the matter of form and function, then use the ASK-CTL to verify the system model. The application of CPN to STPA does not significantly alter the original analysis process of STPA. It can be seen that combining these two methods can realize the complementary advantages and reaches the reasonable and detailed description of the real system. We can write simple code to trace the path lead to the unsafe state. Finally, the train control system is taken as an example to identify the system hazards.

Compared with the traditional method of the chain events model, using the formalSTPA method to identify hazards can greatly reduce the reliance on expert experience in safety analysis. The safety analysis results are accurate and more objective while the efficiency is improved. But it is impossible to formalize each demand due to the narrow scope of formal methods. We can only reduce the design errors in the modeling process and avoid human factor. The state space explosion should be avoided. Further research should be carried out as follows:

- ① We should reduce the redundancy state of the model with state space explosion. In this process, it should be noted which states need to be deleted, and prove that the state space after deletion is equal to the state space before the deletion.
- ② Research on the reachable analysis algorithm. For the complex control system, the model structure is more complex, and the analysis process may cause the state space explosion. So we can research the reachable algorithm and develop a tool to automatically identify each path that may lead to accidents.
- ③ Consider the impact of more failures on the CTCS-3 level control system.
- ④ The method can be more widely applied to other systems in order to fully test the superiority and rationality of this method.

References

1. Zhang Y (2013) Study on the safety risk identification and analysis of train control system of high-speed railway. Southwest Jiaotong University, Chengdu, pp 2–10 (in Chinese)
2. Liu J (2015) A safety analysis method for high-speed railway train control system in requirements phase based on STPA. Beijing Jiaotong University, pp 5–7 (in Chinese)
3. Leveson NG (2011) Engineering a safer world: systems thinking applied to safety (engineering systems). MIT Press, Cambridge, MA
4. Min O, Liu H, Yu MH et al (2010) STAMP-based analysis on the railway accident and accident spreading: taking the China-Jiaoji railway accident for example. *Saf Sci* 48(5):544–555
5. Liu J, Tang T (2014) Functional safety analysis method of CTCS-3 level system based on STPA. *China Railway Sci* 35(5):86–95 (in Chinese)
6. Dirk S, Sebastian HR, Jan W et al Integration of petri nets into STAMP/CAST on the example of Wenzhou 7.23 Accident. *Control Autom Theory Transport Appl* 65–70
7. Wang R, Zheng W, Liang C et al (2016) An integrated hazard identification method based on the hierarchical colored petri net. *Saf Sci* 88:166–179
8. Colored Petri Nets (2008) Stochastic discrete event systems. Springer Berlin Heidelberg, pp 99–124
9. Rengaowa SA (2010) Research on the LDP protocol verification based on colored petri nets. Inner Mongolia University, Hohhot, pp 47–59 (in Chinese)
10. Wang R (2014) Design and application of the hazard identification method for socio-technical system based on the BFM-STPA. Beijing Jiaotong University, Beijing, pp 13–23 (in Chinese)
11. Chinese train control system three level. China Railway Publishing House (in Chinese)

Research on Reliability Evaluation of Traction Power Supply Equipment Considering Human Factors

Qianqian Yang, Sheng Lin and Ding Feng

Abstract Human error is one of the important factors that affect the reliability of traction power supply equipment. However, the existing reliability evaluation methods of traction power system and equipment often neglect the influence of human factors. This paper focuses on this issue, and the task scenes are classified into three kinds: equipment status monitoring type, equipment emergent repairing type, and equipment periodic maintaining type. First, FMEA is used to qualitatively analyze the human error in each scenario. Then, the human reliability analysis methods (HRA) of three scenarios are proposed to quantify the human error probability (HEP). Finally, the example demonstrated that the proposed methods can quantify the probability of HEP and characterize the influence of human error on the reliability of traction power supply equipment.

Keywords Human failure · HRA · FMEA · HEP · Traction power supply equipment operation reliability

1 Introduction

As a link between the power system and electric locomotives, it is essential to ensure the reliability of the traction power supply system. Many experts have proposed different reliability research methods in this field. Such as Weibull distribution, FMEA, and accident tree were used to quantify the reliability of traction power supply system [1]. However, the above literatures ignored human factors. Although Zhengyou He pointed out that reliable human operation is the key to ensure safe and stable operation of the railway traction power supply system [2], there still lacks deep research. Statistically, 10–15% accidents of the traction power system directly are caused by man-made mistakes [3]. So it is inevitable to consider

Q. Yang · S. Lin (✉) · D. Feng
School of Electrical Engineering, Southwest Jiaotong University,
Jiulidi Street, Jinniu District, Chengdu, Sichuan Province, China
e-mail: slin@home.swjtu.edu.cn

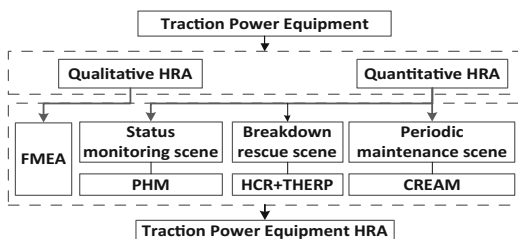
human-made hazards. Human reliability analysis (HRA) [4, 5] was proposed in the 1950s, and then Cox [6] first proposed the proportional hazard model (PHM) to estimate invalidation probability of mechanism facility in 1970. The first generation HRA is typically human cognitive reliability (HCR) [7] and technique for human error rate prediction (THERP) [8, 9], which provide new perspectives for assessing human errors. The most important representative of the second generation HRA such as cognitive reliability and error analysis method (CREAM) [10, 11] has made outstanding transformation of HRA. Based on these, Wang Yao [12] made improvement and offered a more simplified process than that of antecedent CREAM. This paper focuses on analyzing human factors of traction power supply system and first classifies the scenarios as the following three kinds: equipment status monitoring scene, equipment breakdown rescuing scene, and equipment periodic maintaining scene. Then, respectively, establish HRA models under the three scenarios. Finally, create an example to verify the feasibility and applicability of the models.

2 Qualitative Analysis of Human Reliability

The framework of traction power equipment HRA is shown in Fig. 1:

FMEA is used to qualitatively analysis human errors of equipment. First, list main devices of traction power system in the above three kinds of work conditions: equipment status monitoring condition, equipment emergent repairing condition, and equipment periodic maintaining condition. Then, respectively, sort out brief descriptions of potential human failure modes and consequences and use these as the basis for further quantitative analysis. In addition, put forward prevention and improvement methods to reduce the probability of human error occurrence.

Fig. 1 HRA chart of traction power supply equipment



3 Quantitative Models of Traction Power Equipment HRA

3.1 HRA Models in Equipment Status Monitoring Scene

Equipment status monitoring scene refers to the scenario which needs operators continuously inspect, monitor, commission, and other tasks. Long hours working makes personnel fatigue. At the same time, the human error probability (HEP) is also affected by other covariates such as cognitive level, operating environment, operating intensity, physiological factors, psychological factors and training level, etc. Proportion hazard model (PHM) is applied to evaluate HEP in this scene. It is assumed that each covariant motioned above has a product effect on the failure rate [13]. The failure rate function form of PHM is as follows:

$$h(t, Z) = h_0(t)\psi(Z) \tag{1}$$

$h_0(t)$ is the hazard function, which indicates human error varies with working hours, and the selection of reliability parameter distribution in applications is based on the actual system or equipment. The most commonly used one is the two-parameter Weibull distribution [14]. The function is given in its explicit form:

$$h_0(t) = \beta t^{\beta-1} / \alpha^\beta \tag{2}$$

where α is the characteristic life parameter; β is the shape parameter.

$\Psi(Z)$ is the connection function, which shows the influence of the six covariates: cognitive level, operating environment, operating intensity, physiological factors, psychological factors and training level on human error rate, where $Z = [Z_1, Z_2, Z_3, Z_4, Z_5, Z_6]$; the influence index value of each covariate Z_i ($i = 1, 2, 3 \dots 6$) can be $\{0, 1, 2\}$. The greater of its value, the more obvious negative effect on human reliability and vice versa. The expression is as below:

$$\psi(Z) = \exp\left(\sum \gamma \times Z\right) \tag{3}$$

where γ is the weight coefficient of Z_i .

The operational reliability function $R_h(t)$: If an operator starts a task from 0 and works to the time t constantly without any human errors. Then, we can get the expression of $R_h(t)$ as follows:

$$R_h(t) = \exp\left[-\int_0^t h(s, Z) ds\right] \tag{4}$$

Finally, the HEP can be described by the following formula:

$$HEP = 1 - R_h(t) \tag{5}$$

3.2 HRA Models in Equipment Emergent Repairing Scene

Equipment emergent repairing scene needs personnel to take emergency measures within the shortest time to eliminate failures and recover normal operation of devices rapidly when it comes to the traction power equipment failure.

THERP+HCR comprehensive model is used to quantitatively analysis HEP in emergency situation. The former is better at describing the errors that occur when operators perform a specific sequence of actions; the latter can better depict HEP when there is a limit to the completion time. In this scenario, if we only use THERP model, it is too rough to measure HEP precisely; if we only use HCR method, for specific operations, it is virtually impossible to reflect the distinction of various types of operation error characteristics. Hence, the best way is to combine the above two methods. The modeling process is shown in Fig. 2.

HCR is based on two basic assumptions:

1. The operational behavior of human is divided into three types: skill, rule, and knowledge. Table 1 shows different corresponding parameters of different behavior types.
2. The human error probability $P(t)$ follows the three-parameter Weibull distribution:

In which, α is the scale parameter, β is the shape parameter, γ is the position parameter, t is operation permission time, and $T_{1/2}$ is execution time.

$$P(t) = \exp \left\{ - \left(\frac{t/T_{1/2} - \gamma}{\alpha} \right)^\beta \right\} \tag{6}$$

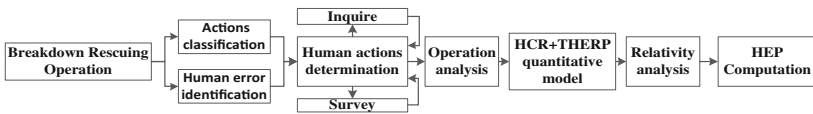


Fig. 2 HRA model in equipment emergent repairing scene

Table 1 Three kinds of behavior type parameter table in HCR model

Type of behavior	a	β	γ
SKILL	0.407	1.2	0.7
RULL	0.601	0.9	0.6
KNOWLEDGE	0.791	0.8	0.5

Operator’s execution time is generally determined by simulating. $T_{1/2}$ needs to be amended before using the formula because each group’s execution time varies by case. The behavioral correction factors considered in HCR are operating experience (K_1), psychological pressure (K_2), and man–machine interface (K_3), and the specific values are shown in Table 2. $T_{1/2}$ correction formula is as follows:

$$T_{1/2} = T_{1/2,n}(1 + K_1)(1 + K_2)(1 + K_3) \tag{7}$$

where $T_{1/2, n}$ is the average execution time (simulator training execution time).

THERP is based on the HRA event tree of serial or parallel sequence movements. The failure probability of a task is represented as basic human error probability (BHEP) which can be found in the well-established THERP forms. But the human error probability is much different due to personnel differentiation. Therefore, the performance shaping factors (PSFs) must be used to modify BHEP to obtain accurate HEP, and the PSF value is also prepared by experts according to the actual situation. Generally, the correction can be represented by the following general formula:

$$P = BHE * P(\text{PSF})_1 * P(\text{PSF})_2 \dots \tag{8}$$

3.3 HRA Models in Equipment Periodic Maintaining Scene

Equipment periodic maintenance scene refers to checking traction power supply equipment at regular intervals to ensure the reliable running.

The model in this scenario is based on CREAM. The core of the method is that human behaviors are not stochastic, instead hinged on the conditions in which they are requested to perform. We can obtain a more accurate and objective HEP through

Table 2 Correction factors of HCR model

Behavior type	Correction factor		Behavior coefficient
Operating experience	K_1	1. Expert, well trained	-0.22
		2. Average training level	0.00
		3. Novice, low training level	0.44
Psychological pressure	K_2	1. Serious emergency situation	0.44
		2. Potential emergency situation	0.28
		3. Best emergency situation	0.00
		4. Low emergency situation	0.28
Man–machine interface	K_3	1. Excellent	-0.22
		2. Good	0.00
		3. General	0.44
		4. Inferior	0.78
		5. Very inferior	0.92

the amendment of common performance conditions (CPCs). Table 3 shows the relationship between CPCs and expected effects, which indicates the quantification of human cognitive errors.

Where 1, 0, and -1 , respectively, represent three kinds of performance expectation effects, which quantify the impact of human cognitive errors. When a single CPC score is determined, its expectancy effect can be denoted as S_i 1 on behalf of improvement, -1 indicates reduce, and 0 means non-significance.

We first determine the cognitive behaviors and movements involved in specific operations according to the operational tasks, second figure out the possible cognitive failure mode, and finally refer to Table 4 so as to get the basic cognitive function failure probability HEP_0 .

Table 3 CPCs and expectancy effects

CPC	Level	Expectancy effect (S_i)
1. Adequacy of organization	Very efficient (9–10)	1
	Efficient (6–8)	0
	Inefficient (3–5)	0
	Deficient (0–2)	-1
2. Working conditions	Advantageous (8–10)	1
	Compatible (4–7)	0
	Incompatible (0–3)	-1
3. Man–machine interface	Supportive (9–10)	1
	Adequate (6–8)	0
	Tolerable (3–5)	0
	Inappropriate (0–2)	-1
4. Availability of plans	Appropriate (8–10)	1
	Acceptable (4–7)	0
	Inappropriate (0–3)	-1
5. Number of simultaneous goals	Fewer than capacity (8–10)	1
	Matching current capacity (4–7)	0
	More than capacity (0–3)	-1
6. Available time	Adequate (8–10)	1
	Temporarily inadequate (4–7)	0
	Continuously inadequate (0–3)	-1
7. Duty time	Daytime (6–10)	0
	Nighttime (0–5)	-1
8. Adequacy of training	High experience (8–10)	1
	Low experience (4–7)	0
	Inadequate (0–3)	-1
9. Crew cooperation quality	Very efficient (9–10)	1
	Efficient (6–8)	0
	Inefficient (3–5)	0
	Deficient (0–2)	-1

Table 4 Nominal values for cognitive function failures

Cognitive function	Generic failure type	Basic value
Observation (P_1)	Wrong object observed (P_{11})	0.0010
	Wrong identification (P_{12})	0.0700
	Observation not made (P_{13})	0.0700
Interpretation (P_2)	Faulty diagnosis (P_{21})	0.0200
	Decision error (P_{22})	0.0100
	Delayed interpretation (P_{23})	0.0100
Planning (P_3)	Priority error (P_{31})	0.0100
	Inadequate plan (P_{32})	0.0100
Execution (P_4)	Action of wrong type (P_{41})	0.0030
	Action at wrong time (P_{42})	0.0030
	Action on wrong object (P_{43})	0.0005
	Action out of sequence (P_{44})	0.0030
	Missed action (P_{45})	0.0300

The model hypothesis 1: The task environment is continuous, and any HEP value can be calculated by the human error rate prediction model.

The model hypothesis 2: There is a logarithmic relationship between the change of task conditions and the variation in HEP [10].

Then, set 2 variables: $I = \Sigma$ Improved total score, $R = \Sigma$ Reduced total score:

$$I = \sum_{i=1}^9 W_i \times S_i (S_i = 1, i = 1, 2, \dots, 9) \tag{9}$$

$$R = \sum_{i=1}^9 W_i \times S_i (S_i = -1, i = 1, 2, \dots, 9) \tag{10}$$

where W_i represents the weight of each CPC.

We divide expectancy effects of CPCs into three categories, the first type we set up only considers those which have improved effect on human operation, and the corresponding HEP is recorded as HEP_I . The second category only considers those which play diminished roles in human reliability, the corresponding HEP is recorded as HEP_R . The effect of the final kind on human reliability is not significant, that is, $S_i = 0$ ($i = 1, 2 \dots 9$), then the corresponding HEP is denoted as HEP_0 . The total human error probability is HEP. We can get the following function expressions from the above hypothesis 2:

$$HEP_I = HEP_0 \times 10^{k_1 I} \tag{11}$$

$$HEP_R = HEP_0 \times 10^{k_2 R} \tag{12}$$

Now we can establish a relationship between HEP_I , HEP_R , HEP_0 , I , and R :

$$HEP = (HEP_I - HEP_0) + (HEP_R - HEP_0) + HEP_0 \tag{13}$$

$$HEP = HEP_0 \times (10^{k_1 I} + 10^{k_2 R} - 1) \tag{14}$$

The method of undetermined coefficients can be applied to obtain the unknown parameters k_1 and k_2 . After we get the value of HEP_0 , it is assumed that the weights of the 9 CPCs should be equal to $1/9$. Specify 0.0001 as the minimum value of HEP_I when I takes the maximum value of $7/9$. The maximum value of HEP_R is 1 when R is at its minimum -1 .

4 Application

Scene 1: Main Transformer Status Monitoring (SM-HRA)

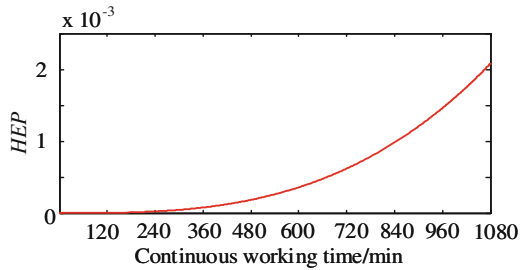
It is assumed that the staff starts the transformer monitoring work from $t = 0$. Suppose there is a context, and the six covariates impact indices and weights obtained by the analytic hierarchy process are shown in Table 5.

The parameters of $h_0(t)$ continue using historical data $a = 200$, $\beta = 3$ [15]. Final quantified HEP curve with the staff continuous working time is shown in Fig. 3 (The total length of the interval is 18 h). As is shown, in the first 393 min of continuous operation, human errors HEP is no more than 1×10^{-4} , which has maintained in a lower level, that is, personnel operating reliability is higher at this stage. So management departments should make flexible work shift system, and try to avoid continuously long time working by the same team.

Table 5 Indices and weights of the six covariates

Covariates	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
Indices	1	1	1	2	1	1
Weights γ	0.26	0.07	0.14	0.05	0.09	0.39

Fig. 3 HEP with staff continuous working time curve



Scene 2: Main Transformer Emergent Repairing (ER-HRA)

The transformer ignition accident occurs when it is running to 4 h (240 min), there are two stages of emergency treatment:

Stage 1: Time-limited emergency procedures.

- (1) Disconnect the high, low side switch, and secondary control power;
- (2) Extinguish fire immediately to stop the fire from spreading.

According to the regulations, $t = 10$ min, $T_{1/2, n} = 5$ min and we choose RULE type operators A, then figure out the human error probability P_1 of this stage.

Stage 2: Time-independent operations.

- (3) Open the oil discharge valve; (Failure rate: 0.0007 [12]);
- (4) Replace the lost equipment (Failure rate: 0.0007 [12]).

The human error probability at this stage denoted as P_2 is quantified by THERP model. The total HEP is shown in Table 6.

Scene 3: Main Transformer Periodic Maintaining (PM-HRA)

When it comes to 9 h (540 min), the transformer is just catching up with its minor repair. It will take 6 h to fulfill the maintenance work if the level of maintenance personnel skill is general. According to the former research of the specific steps and corresponding cognitive failure modes, we can get $HEP_0 = 0.0894$, $k_1 = 2.95$, $k_2 = 1.05$, at the same time, we can evaluate $I = 0.36$, $R = 0$ from the data listed in Table 7. Then, bringing these parameters into Eq. (14), we can get $HEP = 0.0894 \times (10^{-2.95I} + 10^{-1.05R} - 1) = 0.0078$

From what we have mentioned above, we ultimately get the following HEP simulation curve under the three scenarios as Fig. 4. The HEP of main transformer breakdown rescuing scene is about 0.0563, which is much higher than the HEP of 0.0078 in main transformer periodic maintaining scene.

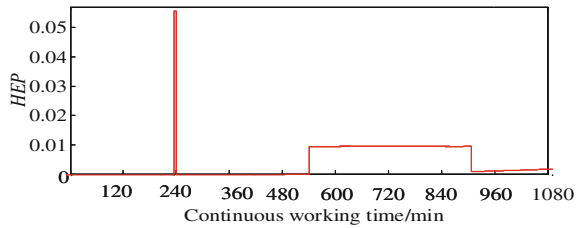
Table 6 Human error rate

Operator	a	β	γ	K_1	K_2	K_3	P_1	P_2	HEP
A (RULE)	0.601	0.9	0.6	0.28	0.28	-0.22	0.0549	0.0014	0.0563

Table 7 Weights and grades of CPCs

CPC	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
Grade	9	7	9	7	6	6	6	7	8
S_i	1	0	1	0	0	0	0	0	0
Weight	0.24	0.18	0.12	0.10	0.06	0.03	0.03	0.12	0.12

Fig. 4 HEP variation curve with working time in three operational scenes of the transformer



5 Conclusion

In this paper, a quantitative evaluation of human error in traction power supply equipment is realized by establishing the reliability models considering human factors. Through the analysis of examples, we can get: Try to keep the working time within a reasonable range to avoid crew fatigue operations under the equipment monitoring scene. In the meantime, it is necessary to improve the staff skill levels to ensure operational reliability for equipment rescuing operations. And it is also of great importance to improve the level of other factors to ensure secure running of facilities under the equipment maintenance scene. The calculation and simulation results show that the influence of various factors on human performance can be taken into account when we quantify the human error probability in specific contexts by the proposed models. It is of great significance to improve the reliability of the whole traction power supply system.

Acknowledgements This paper is supported by Science and Technology Development Project of China Ministry of Railways (2016J010-A), Science & Technology Project of Guangdong Province (2015B010128002), and Collaborative Innovation Centre of Zhuhai (approval numbers: 55560307).

References

1. Xie J, Wu J, Wu Y (2009) Modeling of reliability of traction power supply system based on genetic algorithm. *J China Railway Soc* 31(4):47–51 (in Chinese)
2. He Z, Feng D, Lin S, Sun X (2016) Research on security risk assessment for traction power supply system of high-speed railway. *J Southwest Jiaotong Univ* 51(3):418–429 (in Chinese)
3. Dhillon BS (1990) Applied reliability and quality. In: Springer series in reliability engineering
4. Hong M (2012) Discussion on contents and methods of transformer operation monitoring. *Sci Technol Vision* 9(26):417–501
5. Kim AR, Park J, Kim Y (2016) Quantification of performance shaping factors (PSFs)' weightings for human reliability analysis (HRA) of low power and shutdown (LPSD) operations. *Ann Nucl Energy* 101:375–382
6. Cox DR (1992) Regression models and life: tables, breakthroughs in statistics. Springer, New York, pp 527–541
7. Hannaman GW (1984) Human cognitive reliability model for PRA analysis. NUS-4531

8. Handbook of human reliability analysis with emphasis on nuclear power plant applications: A.D. Swain and H.E. Guttman Draft report for interim use and comment. Office of Nuclear Regulatory Research, US Regulatory Commission, Washington DC 20555, pp 440
9. Kirwan B, Kennedy R (1996) The validation of three human reliability quantification techniques—THERP, HEART and JHEDI: part II—results of validation exercise
10. Hollnagel E (2001) Cognitive reliability and error analysis method
11. Fujita Y, Hollnagel E (2004) Failures without errors: quantification of context in HRA. *Reliab Eng Syst Saf* 83(2):145–151
12. Wang Y, He X (2010) Quantitative analysis of cognitive errors and its application in PSA. *Nucl Power Eng* 27(4):54–58
13. Jiang ST, Landers TL, Rhoad TR (2006) Assessment of repairable-system reliability using proportional intensity model: a review. *IEEE Trans Reliab* 55(2):328–336
14. Muthanna KT, Sarkar A, Das K et al (2006) Transformer insulation life assessment. *IEEE Trans Power Deliv* 21(1):150–156
15. Bao Yingkai, Wen Yunfeng, Han Yuqi et al (2016) Analysis and modeling of human errors in power system reliability evaluation. *Power Syst Technol* 40(2):500–507 (in Chinese)

Bayesian Network-Based Reliability Analysis of High-Speed Train Bogie System

Shuting Zheng, Yong Qin, Li Wang, Limin Jia and Jinchu Zheng

Abstract The events in the fault tree of CRH2 high-speed train bogie system are regarded as nodes in Bayesian network (BN) based on the relationship between the fault tree and the BN. The bogie system is simulated based on BN with improved expectation maximization algorithm (EM). The weakness of the bogie system can be found through the calculation of the fault probability and the causation analysis. It provides the theoretical basis for drawing up the maintenance project and troubleshooting the system, and helps to make the maintenance decision.

Keywords High-speed train · Bogie system · Fault tree analysis
Bayesian network

1 Introduction

With the rapid development of the high-speed railway, the demand for high-speed trains in our country is also increasing. As the core subsystem of the high-speed train, whether the bogie system can work normally and steady affect the safety and reliability of high-speed train directly.

Scholars have proposed qualitative and quantitative analysis methods for train reliability such as fault tree analysis (FTA), failure model and effects analysis (FMEA), Petri nets, and Monte Carlo method. Liu et al. further studied the structural

S. Zheng · Y. Qin (✉) · L. Jia · J. Zheng
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: yqin@bjtu.edu.cn

S. Zheng · Y. Qin · L. Wang · L. Jia
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing, China

L. Wang
Beijing Engineering Research Center of Urban Traffic Information Intelligent
Sensing and Service Technologies, Beijing 100044, China

features and related functions of the electrical subsystem of the CRH3 high-speed train, and constructed the reliability diagram of the subsystem [1]. Zhang et al. researched the full cycle of CRH train using the method of FMEA and proposed a more efficient and scientific method to control the reliability [2]. Liming et al. gave a reliability analysis method by combining FTA and Monte Carlo method to study the urban rail door system [3].

In this paper, the reliability of the bogie system is analyzed and studied by BN with improved EM algorithm. First FTA and BN are introduced. Then, construct BN of the bogie system based on the fault tree model, and utilizing MATLAB to simulate and analyze the model. Finally, conclusions will be summarized.

2 Reliability Analysis Based on the Fault Tree and Bayesian Network

Failure reasons can be found in the fault tree model, and Bayesian network (BN) has a good performance in parameter learning which can update new information and optimize the network and parameters. The combination of the two methods can increase the accuracy of the reliability analysis of the bogie system. The specific method and process are shown in Fig. 1.

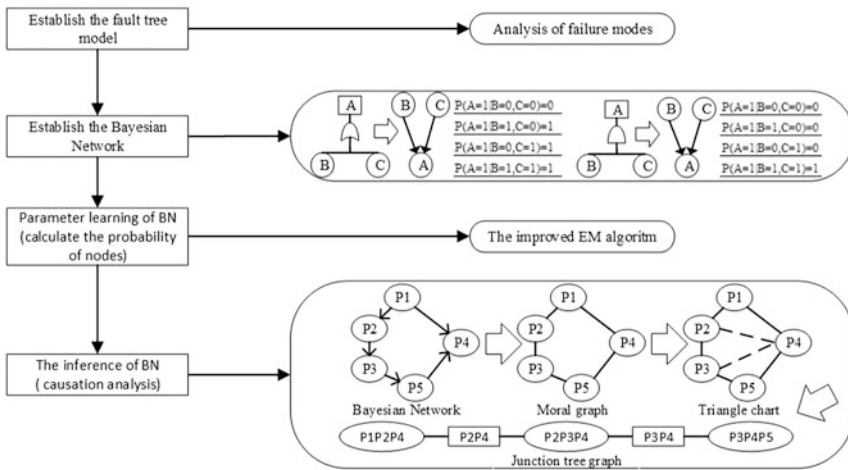


Fig. 1 Process of the reliability analysis based on the fault tree and BN

2.1 The Establishment of the Fault Tree Model

Fault tree model is composed of events and logical gates. Events are used to describe the fault conditions of components and systems. Logical gates are used to describe certain logical relationships. The cause of other events is called the bottom event, and the result of the bottom events is the top event or the intermediate event, which can result in the top event. The common events and logical gates of the fault tree are shown in Table 1.

2.2 Parameter Learning of BN Based on EM Algorithm

BN is a graphical representation which shows the probability relationship between variables, and the degree of mutual influence is a probability value [4]. It consists of conditional probability table (CPT) and directed acyclic graph (DAG) [5].

BN can be represented in $BN = (G, \theta)$, where $G = (N, E)$ represents a DAG with n nodes, where $N = \{B_1, B_2, B_3, \dots, B_n\}$ is a node set, and E means an edge set, showing the influence relationship between nodes. θ represents the CPT of each node, that is, the logical relationship between nodes, and it can be shown as $\theta_i = P(B_i|\pi(B_i))$, where $\pi(B_i)$ is a parent node set of B_i . On the premise of satisfying the above hypothesis, the structure of the BN is determined by the conditional independence assumption among the nodes as follows:











$$P(B_i|B_1, B_2, B_3, \dots, B_{i-1}) = P(B_i|\pi(B_i)), \quad i = 1, 2, 3, \dots, n \tag{1}$$

The joint probability distribution of N is

$$P(B) = \prod_{i=1}^n P\{B_i|\pi(B_i)\} \tag{2}$$

The main idea of the EM algorithm is to set an initial estimate θ_0 and optimize the estimate. Two steps are taken from the current estimate θ_t to the next estimate θ_{t+1} , expectancy calculations and maximization [6]. The expectancy calculation is to compute the likelihood expectation of current θ when D is given.

Table 1 The common events and logical gates of the fault tree

Top event	Intermediate event	Bottom event	Unproved event	Condition event
				
AND gate	OR gate	XOR gate	VOTE gate	Prohibiting gate
				

$$l(\theta|\theta_t) = \sum_l \sum_{x_l} \ln p(D_l, X_l|\theta)P(X_l|D_l, \theta_t) = \sum_{ijk} f_i(x_i = j, \pi_i = k) \ln \theta_{ijk} \quad (3)$$

where $f_i(x_i, \pi_i) = \sum_l p(x_i \pi_i | D_l, \theta_t)$ and X_l is a variable set, which is incomplete in the data set D_l . Maximize the current expected likelihood function value, and calculate the next estimate θ_{t+1} . For all θ , $l(\theta|\theta_{t+1}) \geq l(\theta|\theta_t)$ and $f_i(x_i, \pi_i) = \sum_l p(x_i \pi_i | D_l, \theta_t)$.

Therefore, the expectancy calculation is to compute each variable function $f_i(x_i, \pi_i)$ in proper order. The maximum is based on formula 4.

$$\theta_{ijk} = \frac{f_i(x_i = j, \pi_i = k)}{\sum_i f_i(x_i = j, \pi_i = k)} \quad (4)$$

The main idea of the improved EM algorithm is to speed up the calculation by modifying $f_i(x_i, \pi_i)$.

Define Y is a variable subset of BN, and the joint probability of its ancestors set $\pi(Y)$ is $p(\pi(Y)) = \prod_{x_i \in \pi(Y)} p(x_i | \pi_i)$

When the data set is incomplete, first calculate the function $l(\theta|\theta^t) = \sum_l \sum_{Z_l} \ln p(D_l, Z_l|\theta)P(Z_l, D_l|\theta^t)$, where Z_l is a variable of the original parent node set and its value is the missing data of D_l . And then after maximizing $l(\theta|\theta^t)$, estimate the next θ_{ijk}^{t+1} .

$$l(\theta|\theta^t) = \sum_{ijk} g_t(x_i, \pi_i) \ln \theta_{ijk} \quad (5)$$

$$\theta_{ijk}^{t+1} = \frac{g_t(x_i = k, \pi_i = j)}{\sum g_t(x_i = k, \pi_i = j)} \quad (6)$$

2.3 Bayesian Network Inference

BN inference can reason out the cause of failure according to the collected fault data based on the CPT and the DAG [7]. The greater the combination probability of the root nodes is, the better the root node can describe the components under the fault data. Junction tree algorithm, which is one of the exact reasoning algorithms [8], will be used in this paper. It has a better performance in calculating speed, and it applies to both simply connected networks and multiplies connected networks.

The implementation procedure of the junction tree algorithm is as follows:

Step 1, transform the BN into the undirected moral graph, ignoring the direction of the edges of BN.

Step 2, triangulate the ring, and the number of nodes of the ring is not less than four. Form a triangle chart by adding an undirected edge to noncontiguous nodes.

Step 3, a junction tree is joined by a scatter set consistent with group characteristics. Initialize the tree and give the probability value to each node to complete the establishment of the junction tree.

3 Simulation and Analysis of the Bogie System Model

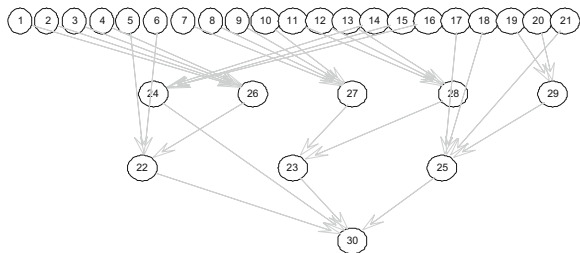
3.1 The Establishment of Fault Tree and Bayesian Network of the Bogie System

Four main functions of the bogie system are summed up according to the analysis of the structure and functions and the historical fault data. Establish the fault tree and transform it into the corresponding BN according to the mapping relation. Simulate the model based on MATLAB as shown in Fig. 2. The nodes X1–X21, Y1–Y8, and T are numbered 1–30, and the value of each node is 1, which means “normal”, or 2, which means “failure”.

The meaning of nodes in the BN of the bogie system is shown in Table 2, and each node stands for the state of the component or the system.

Proper conditional probabilities of nodes in the BN need to be confirmed. First, give the prior probabilities of root nodes, which depend on the fault data of the CRH2 high-speed train from January 2010 to December 2012. After the pretreatment, there is the number of 923 fault data in total. The failure rate of the bottom events in the fault tree can be obtained according to statistical analysis of high-speed train failure modes.

Fig. 2 The BN model simulated by MATLAB



3.2 The Parameter Learning of the Simulation Model

The improved EM algorithm converges faster than the traditional EM algorithm. As shown in Fig. 3a, the red line is the improved EM algorithm, and the blue line is the traditional EM algorithm. The computation time of the improved EM algorithm is 15–18% faster than that of the traditional one.

Training data are randomly generated, and learning space data are taken 600, 1600, and 2600. 5, 15, and 30% loss ratios are arranged in each data sample.

Table 2 The meaning of each node in the BN

No.	Node meaning	No.	Node meaning	No.	Node meaning
T	Bogie	X2	Axle	X12	Grounding device
Y1	Load steering	X3	Wheel	X13	Air spring
Y2	Buffer function	X4	Bearing	X14	Drawbar
Y3	Traction function	X5	Framework	X15	Traction motor
Y4	Braking function	X6	Coupling	X16	Gearbox
Y5	Wheelset device	X7	Height adjustment device	X17	Brake disk
Y6	Damping device	X8	Anti-hunting damper	X18	Brake
Y7	Auxiliary buffering device	X9	Lateral shock absorber	X19	Brake clamp
Y8	Caliper brake	X10	Vertical shock absorber	X20	Pressure cylinder
X1	Axle box body	X11	Tread cleaning device	X21	Speed sensor

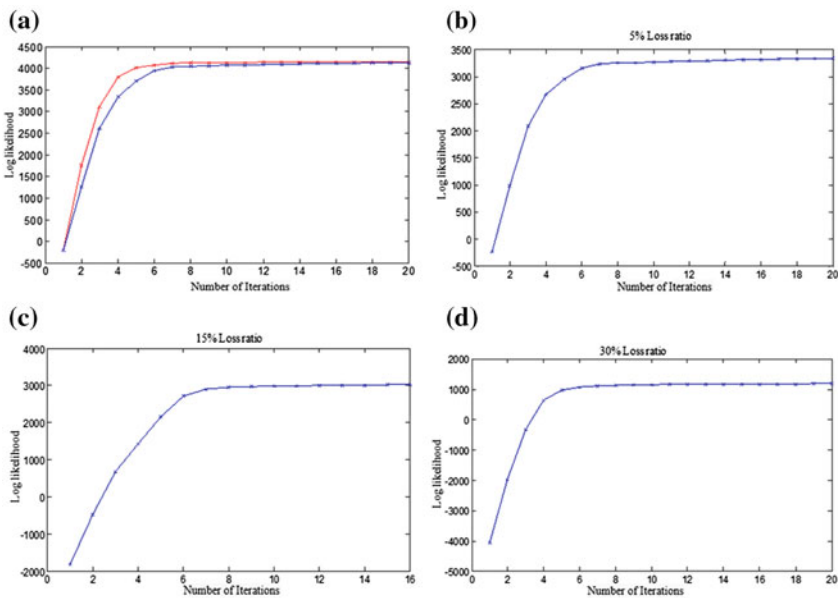


Fig. 3 The comparison of EM algorithm

Different iterative and convergence of log likelihood for different loss data are shown in Fig. 3b, c and d. Under the same data set, as the loss ratio of the data gets bigger, the log likelihood convergence value gets smaller. When the loss ratio is the same, the bigger the data set gets, the smaller the log likelihood convergence value gets. Therefore, the learning situation of parameter learning can be more realistic by reducing the proportion of data loss or increasing the data set.

3.3 Inference of Simulation Model

The BN of the bogie system is simulated by MATLAB, and the improved EM algorithm is used for the parameter learning. The causation analysis is carried out based on the junction tree algorithm. Compute the posterior probability of the 21 root nodes under the situation of “evidence{T} = 2”, and the results are shown in Fig. 4.

The causes of failure can be found quickly and precisely by causation analysis which can improve the efficiency of maintenance work and reliability of bogie system the when the bogie system fails to work steadily. The most probable mode of the induced bogie failure is “bearing damage” → “wheelset failure” → “load steering failure” → “bogie failure” or “axle crack” → “wheelset failure” → “load steering failure” → “bogie failure”. When the bogie system fails to work, the causes ranking higher which are also the weak parts need to be paid more attention. More attention is needed to these components during the design and maintenance period.

The following measures can be taken for the maintenance of the key components. Clean the surface of bearings before checking the appearance and use fine sandpaper to polish the scratches and remove rust. Oil stains and rust on the surface of the axle need to be cleaned before checking. The wheels need to be visually checked for any cracks or other problems. Each component of the bogie system has

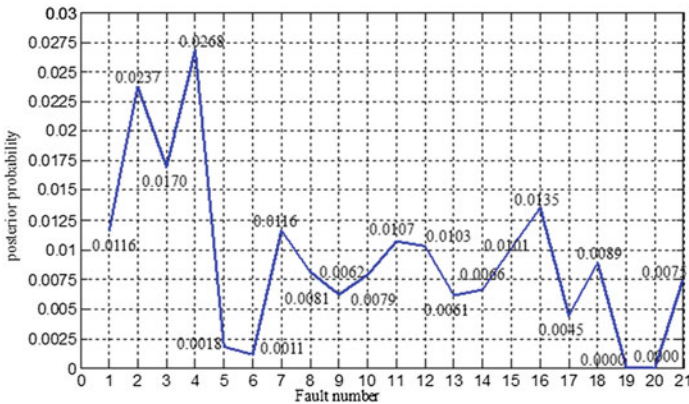


Fig. 4 The posterior probability of causation analysis

its own emphases during the design and maintenance period. It is important to improve the reliability of the bogie system by regular inspection and maintenance.

4 Conclusion

First, the fault tree method and BN theory are studied. Based on these methods, the fault tree model of the bogie system is established and transformed into a BN model simulated by MATLAB. And then the improved EM algorithm is proposed to increase the speed and precision of the calculation. Last, causation analysis of the bogie system is completed. The results show that the bearings, axles and wheels are the weak parts of the bogie system. A new method is provided for the reliability research of the bogie system.

Acknowledgements The authors would like to express thanks to the editor and anonymous reviewers for their help in revising the manuscript. This study is funded by the National Key Research and Development Program of China (2016YFB1200401 & 2016YFB1200505), National Key Technology R&D Program (NO.2015BAG12B01), National Natural Science Foundation of China (71701010), the Fundamental Research Funds for the Central Universities (2017JBM033), and Beijing Jiaotong University State Key Laboratory of Rail Traffic and Control and Safety (RCS2017K001 & RCS2016ZT016).

References

1. Liu JQ, Cui XG, Sun BC (2013) Research on reliability of high voltage electrical system for CRH_3. *J Railway* 35(06):22–27 (in Chinese)
2. Zhang WJ, Lan N (2013) Research on the reliability growth management techniques of high-speed train for whole life cycle. Springer, Berlin, Heidelberg
3. Liming Z, Guoqiang C, Jianwei Y et al (2010) Monte-Carlo simulation based on FTA in reliability analysis of door system. In: 2010 the 2nd international conference on computer and automation engineering (ICCAE). IEEE, pp 713–717
4. Yin XW, Qian WX, Xie LY (2008) Application of Bayesian network in reliability evaluation of mechanical systems. *J Northeast Univ, Sci Ed* 29(4):557–560 (in Chinese)
5. Yu DF, Sun ZL, Chang SH (2004) Research on uncertain reasoning based on Bayesian networks. (8):6–8 (in Chinese)
6. Liu JN (2007) Research on Bayesian network inference algorithm. Hefei University of Technology (in Chinese)
7. Wang FY (2012) Fault prediction of locomotive traction converter based on Bayesian network. Beijing Jiaotong University (in Chinese)
8. Friedman N (1998) The Bayesian structural EM algorithm. In: Proceedings of the fourteenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., pp 129–138

A Rail Corrugation Detection Method Based on Wavelet Packet Energy Entropy

Kuan Lang, Zongyi Xing, Wei Dong and Xudong Gao

Abstract Aiming at the problem of the rail corrugation detection about low efficiency and low success rate, a method based on wavelet packet energy entropy (WP-EE) was proposed. First, the rail corrugation and the train physical fitting models were established, and then a coupling dynamics model was developed by using SIMAPCK. Next, four-layer wavelet packet decomposition of the axle box vibration signal is performed, wavelet packet energy entropy and wavelet energy of each node are calculated. Finally, the simulation results were compared which reflect the information of rail corrugation.

Keywords Rail corrugation · Axis box vibration · Wavelet packet energy entropy

1 Introduction

As the key component of the train operation, the quality of track affects the comfort of passengers and the safety of train operation. In the daily operation process, the track continuously receives the impact and compression from the wheelset, which makes the track appear horizontal and vertical deformation. And serious deformation will cause great harm to train operation. Therefore, monitoring track quality and ensuring track safety is one of the maintenance items of Metro Company [1].

At present, all countries regard both the track inspection technology and the equipment for operating vehicles as the focus of current studies. In the 1996, Grassie [2] first proposed the vision of rail vehicle detection based on the idea of dynamic operation, through the analysis of the operation vehicle axle box vibration and other components of vibration signal. Hitoshi Tsunashima [3] found that the vibration signal in some frequency domain can fully characterize the track wave mill failure, through the acquisition of vehicle vibration signals and wavelet packet

K. Lang · Z. Xing (✉) · W. Dong · X. Gao
Nanjing University of Science and Technology, No. 200 Xiao Lin Wei,
Xuanwu District, Nanjing 210094, China
e-mail: xingzongyi@163.com

analysis based on these signals. Xining Cao [4] analyzed the vibration signals of the axle box collected by Hilbert transform and found that this method can roughly locate the track irregularities. Hongfeng Li [5] modeled the train by using SIMPACK and made a twice integration of the simulated vibration signals; it was found that the vibration signals of the axle box could well characterize the rail corrugation faults. Hongmei Shi [6] studied the algorithm of track stiffness and geometric irregularity based on train vibration signals, and the experimental results show that this method can identify and distinguish rail faults effectively.

This paper put forward the rail corrugation detection based on wavelet packet energy entropy. Through the analysis of axle vibration acceleration signal based on wavelet packet energy entropy, the simulation is operated by using SIMPACK model, and the simulation results show that the method based on wavelet packet energy entropy can determine the rail corrugation fault accurately.

2 Model Foundation

2.1 Simplified Model of Track Corrugation

The track corrugation is a periodic vertical irregularity, so the harmonic function can be used to approximately describe the track corrugation. Considering the diversity of wave grinding phenomenon, if only one wavelength of rail corrugation is considered, that can be just described by a single harmonic function, as shown in Fig. 1a. If the orbital wave grind with a variety of wavelengths is considered, it can be described by a harmonic function, as shown in Fig. 1b.

For the single harmonic function, the expression is

$$Z_o(t) = \frac{1}{2}a(1 - \cos \omega t) \quad 0 \leq t \leq \frac{L}{v} \tag{1}$$

For the multiple harmonic functions, the expression is

$$Z_o(t) = \frac{1}{2}a(1 - \cos \omega t) \quad 0 \leq t \leq \frac{nL}{v} \tag{2}$$

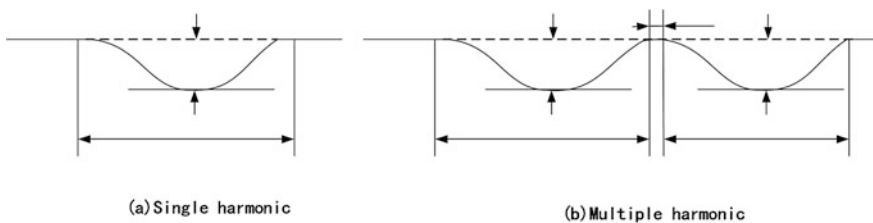


Fig. 1 Described by harmonic function

In the formula, L is the wave grinding wavelength, a is the wave grinding depth, and n is the stimulated wave number.

2.2 Simplified Lumped Parameter Model of Trains

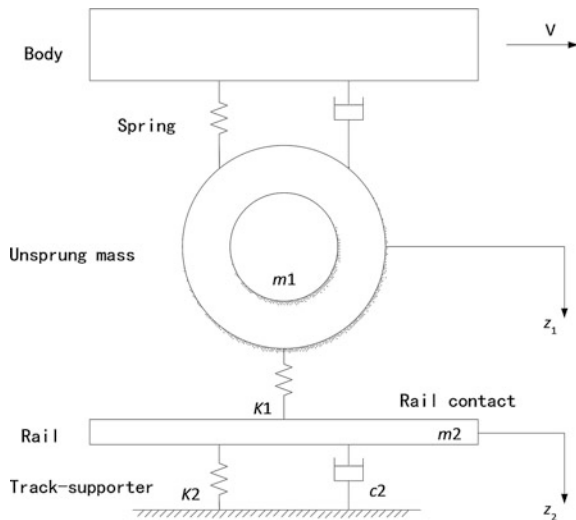
To facilitate the analysis of the change of vibration from the axle box, caused by rail corrugation, it needs to establish a set of lumped parameter model of train. Comparing with the vibration of the axle box caused by the wheel-rail section, there is no obvious vibration impact caused by the upper part of the spring, so the unsprung mass [7] is modeled emphatically. The model built is shown in Fig. 2.

In Fig. 2, we describe the upper part of the wheelset as the body and consider the secondary suspension with both stiffness and damping. The wheel-rail interaction is described as a wheelset with a certain mass, which is connected by the primary spring with stiffness. And the rail is described as a single degree of freedom system with stiffness and damping. In Fig. 2, V represents the speed, m_1 represents the unsprung mass, z_1 is the wheel vertical displacement in the static situation, K_1 is the vertical stiffness of the primary spring, m_2 represents the mass of rail, K_2 represents the vertical stiffness of the rail, and c_2 represents the rail damping. z_2 represents the vertical displacement of the rail in the static state. All parameters are down to the positive direction.

The vertical dynamic equation of wheel/rail is at the speed of v is proposed as Eq. (3)

$$\begin{cases} m_1 \ddot{z}_1 + K_1(z_1 - z_2) = 0 \\ m_1 \ddot{z}_2 + c_2 \dot{z}_2 + K_2 z_2 - K_1(z_1 - z_2) = 0 \end{cases} \quad (3)$$

Fig. 2 Lumped parameter model of train



Assuming the rail corrugation is η , then the rail static equilibrium displacement is $z_2 + \eta$ and the wheel–rail vertical dynamic equations at the speed of v are described by the following formulas:

$$\begin{cases} m_1 \ddot{z}_1 + K_1(z_1 - z_2 - \eta) = 0 \\ m_1 \ddot{z}_2 + c_2 \dot{z}_2 + K_2 z_2 - K_1(z_1 - z_2 - \eta) = 0 \end{cases} \quad (4)$$

Equation (4) is solved by ordinary differential equations:

$$\begin{cases} \frac{\ddot{z}_1}{\omega_1^2} + z_1 - z_2 = \eta \\ \frac{\ddot{z}_2}{\omega_2^2} + \frac{c_2}{K_1} \dot{z}_2 + \left(\frac{K_2}{K_1} + 1\right) z_2 - z_1 = -\eta \end{cases} \quad (5)$$

where $\omega_1 \sqrt{K_1/m_1}$, $\omega_2 \sqrt{K_1/m_2}$, and \ddot{z}_1 is the axle box acceleration.

Taking the single harmonic wave as an example, the Laplace transform of Eq. (1) is carried out:

$$\eta(s) = \frac{\frac{1}{2} a \omega^2}{s(s^2 + \omega^2)} \quad (6)$$

To obtain z_1 , **the response**, and \ddot{z}_1 which is the acceleration of unsprung mass, one can use the following formulas under the incentive of $\eta(s)$:

$$\begin{cases} z_1(s) = H_{Z_1}(s)\eta(s) \\ z_2(s) = H_{Z_2}(s)\eta(s) \end{cases} \quad (7)$$

where $z_1(s)$ and $z_2(s)$ are, respectively, the Laplace transforms of z_1 and z_2 , $H_{Z_1}(s)$ and $H_{Z_2}(s)$ are the transfer functions of Z_1 and Z_2 , respectively.

Then, transform the Eq. (5) by Laplace and bring it into the calculation with Eq. (7). Make all equations above simultaneous, then we can get:

$$H_{Z_1}(s) = \frac{\frac{s^2}{\omega_2^2} + \frac{c_2 s}{K_1} + \frac{K_2}{K_1}}{\frac{s^4}{\omega_1^2 \omega_2^2} + \frac{c_2 s^3}{K_1 \omega_1^2} + \left[\frac{1}{\omega_2^2} + \left(\frac{K_1}{K_2} + 1\right) \frac{1}{\omega_1^2}\right] s^2 + \frac{c_2 s}{K_1} + \left(\frac{K_2}{K_1} + 1\right)} \quad (8)$$

Finally, under zero initial conditions, the Laplace transform of \ddot{z}_1 can be described by Eq. (9):

$$\ddot{z}_1(s) = s^2 H_{z_1}(s) \eta(s) = \sum_{i=1}^3 \frac{A_i}{s + p_i} + \frac{Bs + C}{s^2 + \omega^2} \quad (9)$$

where A_i , B , C , and p_i are both constants, and the value is determined by $\omega_1 \omega_2$, K_1 , and K_2 ; then, we can get Eq. (9) inverse Laplace transformed:

$$\ddot{z}_1 = \sum_{i=1}^3 A_i e^{-p_i t} + D \cos(\omega t + \psi) \tag{10}$$

where ω is the frequency of rail corrugation. By solving the model, $\sum_{i=1}^3 A_i e^{-p_i t}$ is frequency independent, $D \cos(\omega t + \psi)$ is a steady-state value. Therefore, \ddot{z}_1 and ω are equal, so the rail corrugation can be detected by vibration signal of axle box. When the train is at the speed of v , there will exist corrugation at the length of λ with the frequency $\omega = 2\pi v/\lambda$. And the response frequency of axle box can be described as $f = v/\lambda$. Therefore, the corrugation wavelength can be obtained by knowing the speed and the response frequency of the axle box.

2.3 Establishment of Vehicle Coupling Dynamics Model Based on SIMPACK

Using SIMPACK, the first thing to build a model of a railway vehicle is to construct a topological graph of the model. Considering the repeatability and complexity of train components, a bottom-up structural modeling method is adopted to model the car body model in three steps: wheelset, bogie, and vehicle, and the rigid body model is used for the construction [8]. The specific topology graph is shown in Fig. 3.

As can be seen from Fig. 3, the whole car was divided into the body and the substructure, in which the dummy body, bogie frame, and wheelset were encapsulated.

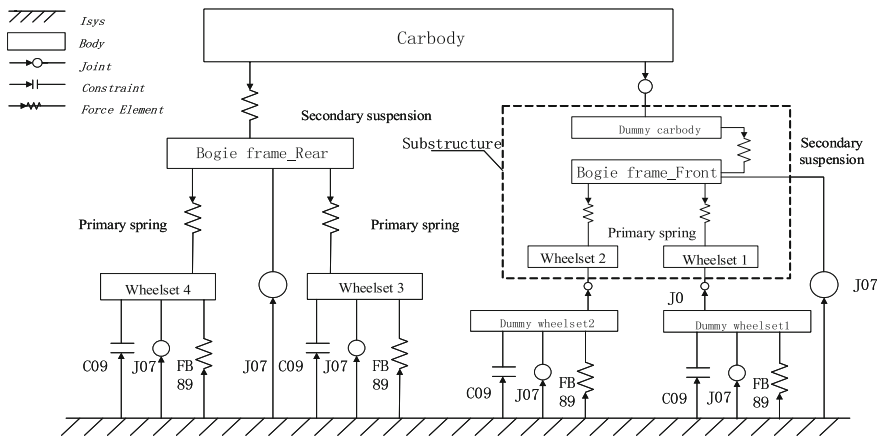


Fig. 3 The topological diagram of a railway vehicle

Between the body and the substructure adopts a secondary suspension force element connection; the inner of the substructure between bogie frame and wheelset adopts a primary spring force element, and the bogie frame is connected and hinged with Nos. 07 hinge. The wheel and rail are connected by Nos. 09, 07 and 89 hinge force elements. SIMPACK modeling process is parameterized, must be based on the actual track vehicle parameters for the corresponding component parameter settings. Select a certain type of vehicle dynamics parameters to establish a single vehicle model shown in Fig. 4.

3 Wavelet Packet Energy Entropy

If there is a signal to be analyzed, and the signal length is N , then the wavelet packet energy entropy of this signal can be obtained by the following steps:

Step 1: j -layer of wavelet packet decomposition is used to test signals, and a series of sub-signal sequences with different frequency bands are obtained after decomposition.

Step 2: The wavelet packet reconstruction coefficients of the obtained high- and low-frequency sequences are expressed as S_{jk} , ($k = 0, 1, \dots, 2^{j-1}$).

Step 3: Calculate the wavelet packet energy values of each subsequence: E_{jk} ($k = 0, 1, \dots, 2^{j-1}$).

$$E_{jk} = \int_{t_{i-1}}^{t_i} |A_i(t)|^2 dt \quad (11)$$

where $A_i(t)$ is the maximum amplitude of the node, and t_{i-1} and t_i are the start and stop time points of the signal, respectively.

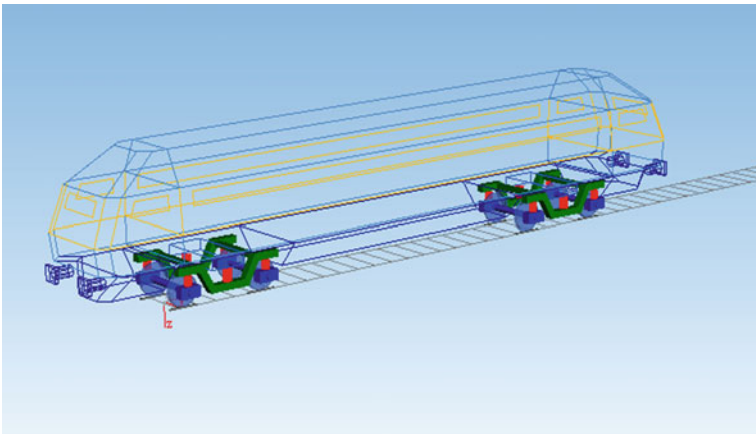


Fig. 4 The model of single vehicle

Step 4: Get the total energy value E .

$$E = \left(\sum_{k=0}^{2^{j-1}} |E_{jk}|^{1/2} \right) \quad (12)$$

Step 5: Calculate the probability of each node's energy relative to the total energy as p_{jk} .

$$p_{jk}(i) = \frac{E_{jk}}{E} \quad (13)$$

Step 6: Calculate the wavelet packet energy entropy:

$$H_{jk} = - \sum_{i=1}^N p_{jk}(i) \log |p_{jk}(i)| \quad (14)$$

In the formula, H_{jk} is the k th wavelet packet energy entropy estimate after the j -layer decomposition of the signal.

4 Simulation Verification

In this paper, the vibration signals of the axle box based on SIMPACK software are not disturbed by noise signals, so the filtering process can be ignored for a while.

4.1 Fault Diagnosis and the Setting of the Threshold for Rail Corrugation Fault Detection

In this paper, the setup of rail corrugation is carried out by superimposing the rail corrugation data on the American track spectrum of fifth grade, which the simulation test is based on. There is no processing on both sides of the track from 0 to 50 meters of the track, and 14 sets of different rail corrugation signals are set on the right track at the 50–95 m. And in the meantime, the left track only be added the vertical irregularity of the American track spectrum of fifth grade (AAR5).

Analyzing of the results in Table 1, it shows that the first group contains American track spectrum of fifth grade only, and its wavelet packet energy entropy is $1.8356e^{-5}$, while the other 13 groups which contain the corrugation section, the range is between 0.3 and 2.5. It can be seen, for the tracks containing the corrugation section, its wavelet packet energy entropy is far greater than the fault-free tracks.

Table 1 The total entropy calculation results of each grouping experiment

Track number	Speed (m/s)	Track: 50–95 m	Grinding wavelength (mm)	Corrugation depth (mm)	Adding position of corrugation (m)	Wavelet packet energy entropy
1	10	AAR5 only	none	None	None	$1.8356e^{-5}$
2	10	AAR5 + Corrugation	150	0.05	65–67	1.9964
3	10	AAR5 + Corrugation	150	0.10	65–67	2.3105
4	10	AAR5 + Corrugation	150	0.15	65–67	2.3450
5	10	AAR5 + Corrugation	200	0.05	65–67	1.3066
6	10	AAR5 + Corrugation	200	0.10	65–67	1.8334
7	10	AAR5 + Corrugation	200	0.15	65–67	1.9665
8	10	AAR5 + Corrugation	250	0.05	65–67	0.8135
9	10	AAR5 + Corrugation	250	0.10	65–67	1.4977
10	10	AAR5 + Corrugation	250	0.15	65–67	1.8185
11	15	AAR5 + Corrugation	250	0.10	65–67	1.7736
12	20	AAR5 + Corrugation	250	0.10	65–67	1.7948
13	10	AAR5 + Corrugation	150 200	0.10	65–67 75–77	0.4304
14	10	AAR5 + Corrugation	250 200	0.10	65–67 75–77	0.3933

Based on the above results of wavelet packet energy entropy in the rail corrugation fault detection, it is not difficult to find that the wavelet packet energy entropy has obvious measurability. And its value is affected by the grinding wavelength, the grinding depth, and the speed, and it has an obvious linear relationship with the above three. Therefore, the wavelet packet energy entropy of the vibration signal of the axle box can be taken as the reference value for the fault detection of the rail corrugation. In the simulation, according to the above data, it is suggested that the threshold is set to 0.3. If the wavelet packet energy entropy is greater than 0.3, the wave mill fault will be detected, otherwise, no wave mill fault will occur. Because the selection of the threshold is related to the vehicle running state and vehicle speed, it needs to be calibrated and adjusted according to the actual situation at the late stage.

4.2 Wavelength Detection of Rail Corrugation Fault

The data of Group 3, Group 6, Group 9, and Group 13 in Table 1 were selected for further analysis. The wavelet packet time–frequency analysis is performed on the sub-node signals obtained by the signal decomposition of the data, and then the time–frequency spectrums of these four groups are obtained as shown in Fig. 5.

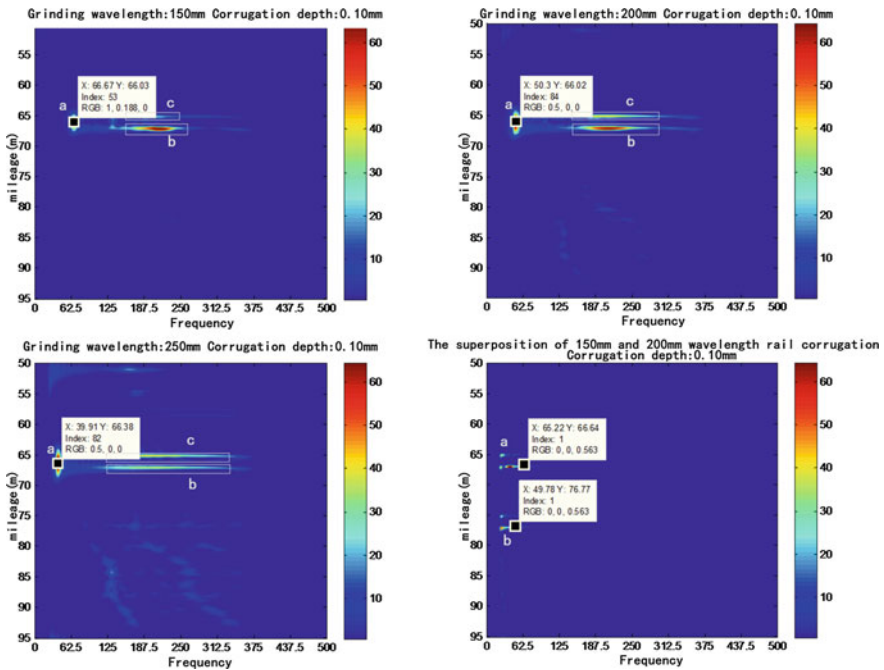


Fig. 5 The time–frequency spectrums obtained by the wavelet packet time–frequency analysis. (left, top) The third group in Table 1, (right, top) the sixth group in Table 1, (left, bottom) the ninth group in Table 1, and (right, bottom) the 13th group in Table 1

In Fig. 5, observe the time–frequency spectrums corresponding to the third, the sixth, and the ninth groups in Table 1. The frequency concentration areas of both the b and c are due to the impact of the train, through the transitional region of the corrugation section and non-grinding section, rather than the frequency concentration caused by the corrugation section. So, the frequency of vibration center caused by the rail corrugation fault of the third group, the sixth group, and the ninth group is 66.67, 50.3, and 39.91 Hz. The expression of grinding wavelength can be described as $\lambda = 1000 v/\omega$, where λ is the grinding wavelength, ω is the frequency of vibration center, and v is the driving speed. So the corresponding grinding wavelength of the third group, the sixth group, and the ninth group is 149.99, 198.81, and 250.56 mm. Observe the time–frequency spectrum in Fig. 5, corresponding to the 13th group in Table 1. The failure center frequency of a is 65.22 Hz, and the failure center frequency of b is 49.78 Hz. Then, calculate the corresponding wave wear wavelength were 153.32 mm, 200.88 mm. The results show that the diagnostic results are in good agreement with the actual wave setting parameters, the maximum error is 2.21%.

5 Conclusion

Based on wavelet packet decomposition and information entropy theory, a method of rail corrugation fault detection based on wavelet packet energy entropy is studied. First, the simplified model of track corrugation and the simplified model of train lumped parameter are established, and the feasibility of detecting rail corrugation based on vibration signal of axle box is verified. Then, SIMPACK multibody dynamics software is used to set up the model. The wavelet packet energy entropy is innovatively applied to track corrugation detection. Subsequently, the simulation experiments are carried out by setting 14 different kinds of rail corrugation faults. And the simulation results show that wavelet packet energy entropy can judge the condition of rail corrugation well.

Acknowledgements This work is supported by National key research and development plan project (2016YFB1200402).

References

1. Yufeng G, Futian W, Rengkui L (2006) Design and realization of analytical information system for railway track state. *China Saf Sci J* 16(7):128–132 (in Chinese)
2. Grassie SL (1996) Measurement of railhead longitudinal profiles: a comparison of different techniques. *Wear* 191(1):245–251
3. Tsunashima H, Mori H, Yanagisawa K et al (2014) Condition monitoring of railway tracks using compact size on-board monitoring device. In: 6th IET conference on railway condition monitoring (RCM 2014). IET, pp 1–5
4. Xining C, Xiaodong C, Shubin Z (2015) Analysis of acceleration of train axle box based on hilbert-Huang transformation. *Instrum Tech Sens* 03:92–95 (in Chinese)
5. Hongfeng L (2014) The research of rail wavy abrasion detection system. Southwest Jiaotong University (in Chinese)
6. Hongmei S (2013) Research on intelligent sensing algorithm of track irregularity based on vehicle dynamic responses. Beijing Jiaotong University (in Chinese)
7. Jian Z, Xuesong J, Xinbiao X et al (2011) Solution methods of rail model in vehicle-track coupling dynamics. *J Traffic Transp Eng* 11(2):32–38 (in Chinese)
8. Kurzeck B (2011) Combined friction induced oscillations of wheelset and track during the curving of metros and their influence on corrugation. *Wear* 271(1):299–310

Pantograph Slide Plate Abrasion Detection Based on Deep Learning Network

Yan Li and Xiukun Wei

Abstract As the key components of electrified railway power supply, pantograph has complex electrical and mechanical effects while working, resulting in a high fault ratio. It is important to detect the defects timely to guarantee the safety of the railway system. Manual detection is the most common detection at present, which is high in accuracy but low in efficiency. Another automatic detection system is limited in function or poor in accuracy. In this paper, deep learning method is used for defects recognition of pantograph slide plate to the identification and classification of different types of defects. Through a large number of experiments and parameters optimizing, the innovated proposed in this paper can reach an accuracy rate of 90.625% used to identify a variety of different defects. This provides an alternative for pantograph slide plate defect identification.

Keywords Pantograph slide plate defect · Image recognition · Deep learning Convolutional neural network

1 Introduction

Pantograph-OCS (overhead contact system) system, the key link in the electrified railway system, has been responsible for powering the electric locomotive. However, a series of complex electromechanical interactions cause a high fault rate, which badly damages the safe operation of the railway. Timely detection of the pantograph is significant important for ensuring the safety of railway systems.

Manual inspection of pantograph defect is the commonly used method at present. It is flexible to identify different types of defects accurately but poor in efficient and need maintenance window to operate [1]. Therefore, some noncontact detection methods are invented. In literature [2], the principle of laser ranging is

Y. Li · X. Wei (✉)

State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing, China
e-mail: xkwei@m.bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_22

215

used to measure the geometric parameters of the pantograph and the wear of the slide plate. In literature [3], ultrasonic is used to detect the crack and wear of pantograph. Noncontact detection is more efficient and less interference than contact way [4]; however, the function of laser detection is unitary, and the accuracy of ultrasonic detection is poor. Based on these, the abrasion recognition of the pantograph surface image has been developed. The images taken by the acquisition device are analyzed using the intelligent algorithm, to identify the different kinds of defects, which is efficient and non-interrupted. Therefore, the detection algorithm is the key to the pantograph defect inspection.

As one of the important deep learning algorithms, Convolutional Neural Network (CNN) is widely used in research and application of computer vision. CNN is first proposed by Yann LeCun in 1998; the LeNet model can reach the accuracy rate of 99.1% for handwritten numeral recognition [5] and has been applied to the postal sectors. After that, a model deeper in hierarchy called AlexNet is proposed by Alex Krizhevsky in 2012, which wins the first prize of ILSVRC that year for the classification error only 15.3% [6]. Since then, the research of CNN has been paid more attention and has been widely used in OCR, face recognition, and some other areas. In this paper, CNN is used for defect recognition of pantograph slide plate images with CAFFE (Convolutional Architecture for Fast Feature Embedding). Compared with the previous method, this model can detect various types of defects with high accuracy without the need for artificial identification of features.

The rest of this paper is organized as follows. In Sect. 2, the pantograph and its main defects are discussed; In Sect. 3.1, deep learning and CNN are brief introduced; In Sect. 3.2, we describe about image preprocessing and network training, and then a new CNN model is designed for this problem and optimized the parameters to reach a better effect. Finally, the conclusions are given in Sect. 4.

2 Pantograph and Its Common Defects

A pantograph is an apparatus mounted on the roof of an electric train, subway, or electric bus to collect power through contact with an overhead catenary wire [7]. The picture of pantograph in Beijing Metro 6 is shown in Fig. 1. There are complex and powerful electromechanical effects on the sliding plate, which has always been a common part of the defects. According to the manifestation of the defect, it can be divided into excessive wear (Fig. 2a), slide eccentric wear (Fig. 2b), groove wear (Fig. 2c), pantograph–catenary arcing (Fig. 2c), etc.

- Excessive wear: electromechanical effects can cause wear on the slide surface, making the slide thinner. If the abrasion exceeds the preset limit, it needs to be replaced in time to avoid further damage or other defects. In Beijing metro line 6 (Fig. 2a), the thickness of the collect head newly installed is 40 mm. It needs to be replaced immediately when the thickness reaches 24 mm.

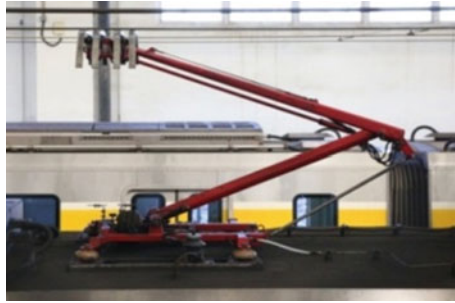


Fig. 1 Pantograph in Beijing Metro line 6

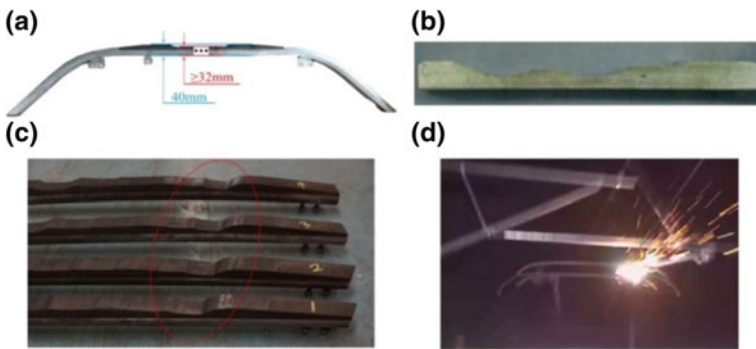


Fig. 2 Common defects of slide plate on pantograph

- Slide eccentric wear: The wear value on both sides of the slide inconsistently due to the improper stagger value is called slide eccentric wear (Fig. 2b). It causes the strength uneven on both sides. Hence, the slide plate needs to be replaced even without excessive wear.
- Groove-shaped wear: Compared with eccentric wear, the groove wear is narrower and deeper (Fig. 2c). Local strength would be seriously reduced, and the contact line possibly damaged by hindering the horizontal movement.
- Pantograph–catenary arcing: Arcing caused by the poor contact between pantograph and catenary is often because of the problem of catenary performance and installation, which will cause the surface slider burning and pitting damage.

In this paper, image recognition based on deep learning is used for pantograph surface defect detection research. The excessive wear, slide eccentric wear, and groove-shaped wear are the main defects to be detected. Besides, there are still some serious faults such as crack, collector head parts missed, etc., which harm traffic by preventing normal current. Therefore, detecting the early defect timely to reduce the damage occur is necessary.

3 Deep Learning Theory and the Application in the Detection of Pantograph Defect

3.1 Deep Learning and Convolution Neural Network

With the upgrade of hardware system and the arrival of big data era, artificial intelligence has been developed rapidly in recent years, which largely depends on the development of deep learning theory. CNN, a typical type of deep learning models, has been widely used in computer vision problems or image-related applications [8]. Compared with traditional artificial neural networks, the development of CNN mainly reflects the depth and complexity of the models. More complex learning tasks and feature information can be competent without the need for artificial identification of features.

In general, CNN consists of standard structure stacked convolutional layers, optionally followed by pooling layer or contrast normalization, and then followed by fully connected layers and have nonlinear function at the end or after each layer [9]. The schematic of LeNet, a typical CNN architecture, is shown in Fig. 3.

The input layer receives preprocessed images and transmits it to convolution layer for feature extraction; each unit receives inputs from a set of units located in a small neighborhood in the previous layer, which is called local receptive field. Then, the followed pooling layer subsamples the feature map from the output of the convolution layer, which greatly reduces the spatial size of the data while preserving the relative positional relationship between the features and also reducing the sensitivity to the displacement and illumination of the sample. Finally, after the convolution of the C3 layer and the subsampling of the P4 layer, the feature that maps output from C5 layer is converted into a required multidimensional vector by fully connecting to Fc layers, and then finishes the feature extraction and classification.

3.2 Image Identification of Pantograph Defect Based on CNN

3.2.1 Image Acquisition and Preprocessing

The image data of pantograph defects used in this paper are collected from Beijing Metro line 6 and Guangzhou subway. Handheld DSLR (Digital Single-Lens Reflex)



Fig. 3 Schematic diagram of CNN

camera and area-array camera installed in the tunnel are used as image acquisition equipment. The collected images contain color and grayscale in different sizes. According to the suggestions from the maintenance staff, the images of different defects are divided into five categories including normal, mild wear, excessive wear, groove wear, and slide eccentric wear; then, the image types are labeled from 0 to 4 for classification. Because the images are different in color mode, size, etc., it is necessary to normalize and preprocess the images before training.

- Normalization: The normalization of the images mainly includes the size and the color mode. The resize tool under Caffe was called and set the mode as “squash” to resize the images to 256*256 pixels in size. For this, research is not sensitive to color, and we unified images to grayscale by `rgb2gray()` in Matlab.
- Mean removal and thresholding: In order to reduce the influence of unrelated information and highlight the difference of defect characteristics, the mean value of images is calculated and subtracted from the dataset by configuring network description file. For the data used in Caffe, framework is defined in the range of [0,1], and it is necessary to thresholding the images to create binary images.
- Construction of the dataset: Caffe is a deep learning framework usually used for voice and image recognition, so a unified database file for stable and high-speed reads is necessary. LevelDB and LMDB are currently supported in Caffe, which load different types of files into unified key-value pairs. The dataset in LMDB can be generated by programming the shell file, and named pan5-sG256. 158 sets of data are contained and basically average in each category. 20% of which are used as validation data to evaluate the prediction accuracy of the CNN network.

3.2.2 Design and Optimization of the CNN Network Structure

The CNN under Caffe is designed mainly by coding two files: *.sh file, the shell file for network running, which sets the running rules, log file storage path and visualization; *.prototxt file, the description file of the network, including the configuration file and the network structure file. The configuration of the learning rate, max iterations, solver mode, and other hyper-parameters are defined in the former; the settings of the network structure, parameters, etc. are defined in the latter.

As the classical model of CNN, AlexNet has high accuracy rate in feature extraction and recognition. The structure is shown in Table 1. Pan5-sG256 dataset is used to train AlexNet and test the recognition effect of the pantograph surface defect. In addition, the output vector in last fully connected layer changed from 1000 to 5 dimensions to meet the classification needs. The training process of AlexNet is shown in Fig. 4.

As we can see, the accuracy is converged to 72.5% at about 1200 epochs, the final test loss is about 1.25, which means in this research, AlexNet is poor in

Table 1 Key parameters and hyper-parameters in AlexNet and optimized networks

Optimized items		AlexNet	Opt-AlexNet	PanNet
Structural parameters	Vision layer	C1-P1-C2-P2-C3-C4-C5-P5-Fc6-Fc7-Fc8		C1-P1-C2-P2-C3-P3-C4-C5-C6-P6-Fc7-Fc8-Fc9
	Network structure			
	Initialization of convolution layers	All-gaussian	1,3,5-gaussian 2,4-msra	1,3,5-gaussian-2,4,6-msra
	Type of pooling layers	All-MAX	1,2-AVE; 5-MAX	1,2-AVE-3,6-MAX
Others	Initialization of Fc_bias	6,7-constant0.1; 8-constant0	All-constant 0.1	All-constant 0.1
	dropout_ratio	0.5;0.5	0.3;0.7	0.3;0.7
	base_lr	0.005	0.015	0.008
	lr_policy	Step	Inv	Inv
Hyper-parameters	Learning rate			

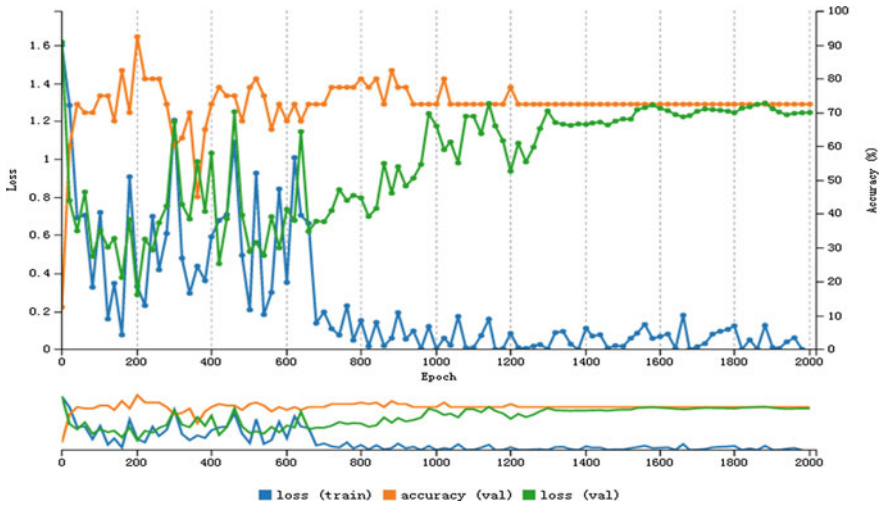


Fig. 4 Training and testing result of AlexNet by Pan-5-sG256 dataset

generalization and cannot be used directly to solve the problem. To achieve better recognition effect, the parameters of AlexNet need to be optimized according to the feature of surface defect images, which is shown in Table 1.

As shown in Table 1, the optimization of the network mainly concentrates on structure parameters and hyper-parameters. The optimization of the structural parameters through the choice of convolution kernel initialization mode uses the msra function as the initialization way in second and fourth convolution layers rather than Gaussian function. Besides, set the first and second pooling layers as the average pooling layer to reserve the feature of the image fully. Furthermore, modify the dropout ratio to improve the generalization ability. The optimization of the hyper-parameters mainly about the learning rate sets the base learning rate to 0.015 and the learning rate decay strategy to inverse decay. The optimized network is named as Opt-AlexNet, and the training process is shown in Fig. 5.

Compared with the original AlexNet, the accuracy rate of the Opt-AlexNet increased from 72.5% to 78.125%, while the loss rate reduced from 1.25 to 0.9, proving that the generalization ability has greatly improved. However, the accuracy of the identification is not good enough, so it is necessary to design a model with deeper hierarchies to achieve better feature extraction.

The depth of the network has a direct effect on the extraction of image features. Therefore, a CNN model based on the pattern of AlexNet which is deeper in hierarchy is built and named as PanNet. The structure of visual layer is shown in Table 1. The design of the visual layers mainly includes the parameters and the feature maps size. The kernel size of convolution layers has four sizes: 13, 11, 5,

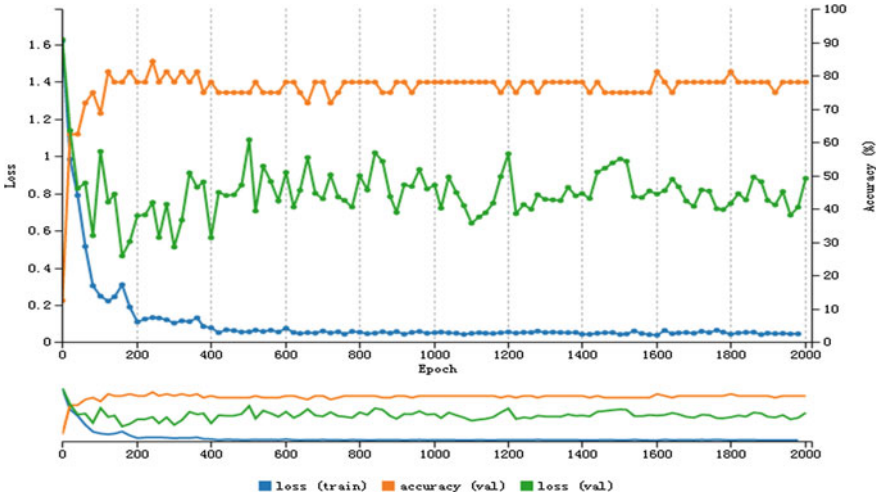


Fig. 5 Training and testing result of Opt-AlexNet by Pan-5-sG256 dataset

and 3. And the size of feature maps is related to the value of kernel size, stride, and pad in each layer, as shown in Eq. 1:

$$W_{out} = \frac{W_{in} + 2 \times pad_size - kernel_size}{stride} + 1, \tag{1}$$

where the W_{in} and W_{out} mean the size of input and output feature map. Key parameters and hyper-parameters in these three networks are shown in Table 1.

As can be seen in Table 1, the convolution layer is initialized in accordance with the experience of Opt-AlexNet, using Gaussian and msra alternately. Similarly, set the AVE and MAX pooling layer alternately. The hyper-parameters are basically consistent with Opt-AlexNet but fine-tuned based on a test to achieve better results. The result of the recognition effect is shown in Fig. 6.

As can be seen obviously, the accuracy rate of the PanNet increased from 78.125% to about 90.625%, and the loss rate reduced from 0.9 to 0.68, which means that both accuracy and generalization ability have been greatly improved. Compared with these three networks mentioned above, the features of the surface defect can be better identified by PanNet. We classified the typical defect images that are not trained in the network using the well-trained PanNet; the top-5 identified probability results shows that the model makes an exact judgment on the different types of defects. For the case of images containing multiple defects, some other defects are also identified based on judging the existence of the main defects, it is proved that the network is feasible and effective for the recognition of pantograph surface defect images.

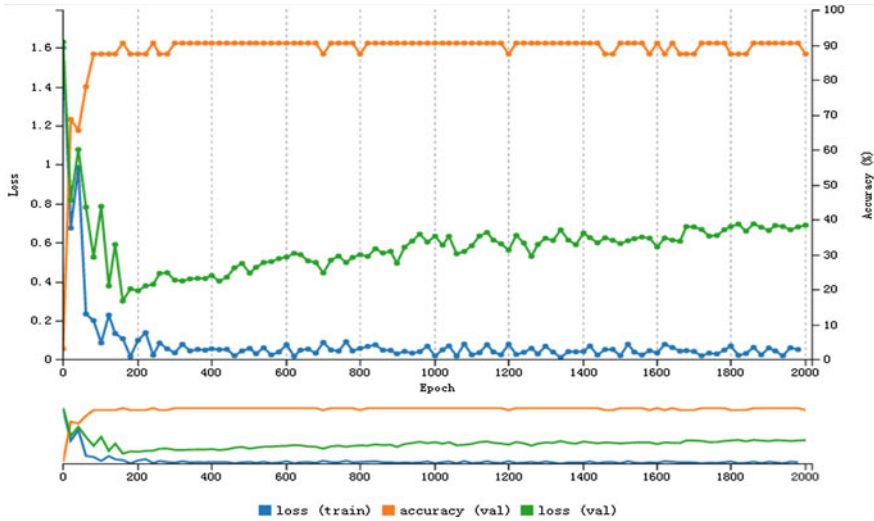


Fig. 6 Training and testing result of PanNet by Pan-5-sG256 dataset

4 Conclusion

In this paper, we present a deep CNN model for pantograph slide plate defect identification. The model is designed to detect different kinds of defects efficiently and finally achieved a better effect with many experiments and optimization. However, there are still some problems. First, the parameters of the network need to be trained with numerous of images with typical features. Second, there are no theoretical guidelines for parameters and network structures designing, which need to be set up based on experience and extensive experiments. How to design efficient network model and parameters, and achieve better recognition effect with limited training data, is the direction of further research.

Acknowledgements This work is partly supported by Chinese National Key Project of Research and Development (Contract No. 2016YFB1200400-2) and State Key Lab of Rail Traffic Control and Safety (Contract No. RCS2016ZT006), Beijing Jiaotong University, Beijing, China.

References

1. Han Z-W, Liu Z-G, Zhang G-N (2013) Overview of non-contact image detection technology for pantograph-catenary monitoring. *J China Railway Soc* 35(6):40-47 (in Chinese)
2. Kuen LK, Lee TKY, Ho SL, Cheng EW (2006) A novel intelligent train condition monitoring system coupling laser beam into image processing algorithm. *HKIE Trans* 13(1):27-34
3. NBao-lai YI, Bo-ming W (2008) Application of ultrasonic ranging principle in monitoring abrasion of pantograph slider. *Electr Drive Locomotives* 5:021

4. Boffi P, Cattaneo G, Amoriello L, Barberis A, Bucca G, Bocciolone MF, Collina A, Martinelli M (2009) Optical fiber sensors to measure collector performance in the pantograph-catenary interaction. *IEEE Sens J* 9(6):635–640
5. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
6. Krizhevsky A, Sutskever I, Hinton GE (2012) Image net classification with deep convolutional neural networks advances in neural information processing systems. 1097–1105
7. Pantograph (transport) [Internet]. En.wikipedia.org. (2017). Available from: [http://en.wikipedia.org/wiki/Pantograph_\(rail\)](http://en.wikipedia.org/wiki/Pantograph_(rail))
8. Zhu J, Liao S, Lei Z, Li SZ (2017) Multi-label convolutional neural network based pedestrian attribute classification. *Image Vis Comput* 58:224–229
9. Ma J, Wu F, Zhu J, Xu D, Kong D (2017) A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics* 73:221–230

Track Surface Defect Detection Based on Image Processing

Yuxin Liu and Xiukun Wei

Abstract In this paper, computer vision-based methods are presented to detect the rail track surface defects automatically. The detection is the key foundation to inspect and assess railways, and for the operation safety and rail maintenance, railways inspection is the critical task. To achieve this goal, the rail surface edge's likelihood is investigated, and the Canny edge detector for defects extraction is introduced to guarantee the detection of the rail surface damage accurately. The analysis performed on some image data captured on the field has demonstrated encouraging detection performance on rail track surface defect detection.

Keywords Image processing · Defect detection · Track surface
Field image analysis

1 Introduction

To ensure the safety and efficient operations, railway inspection is a very important issue in the urban rail transit. Usually, in order to achieve the above goal, this key issue is always operated by trained inspectors who periodically walk along the track lines to seek out any damaged rail components. However, with the extension of urban rail transit network, the manual inspection becomes more and more inappropriate for its cost and danger; meanwhile, the maintenance and the inspection face more challenges than ever before. So developing automatic inspection system has attracted much interest all over the world, especially applying the machine-vision technology.

Recently, many automatic systems based on machine-vision technology have been introduced to solve different tasks in railway inspection. The V-CUBE system for inspecting track surface defects with line-scan cameras, which is developed by MERMEC Group [1]; the TrackVue system developed by RailVision for analyzing

Y. Liu · X. Wei (✉)

State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing, China
e-mail: xkwei@bjtu.edu.cn

rail and wheel wear, track gauge with cameras and laser sensors [2]; however, technical details or performance report about both systems are not available. The VisiRail Joint Bar Inspection System, with laser equipment and scan line cameras, which is developed by ENSCO [3]. Yang et al present a real-time rail inspection system for detecting rail components such as ties and anchors performing at 16 km/h only [4]. In [5], the authors apply computer vision to finding the particularly worn and missing fasteners, but it does not involve in track surface defect. Some image processing approaches are used to achieve the rail surface detection in [6]. Similar efforts are also reported in [7] and [8] for inspecting the broken fasters and rail surface, experiments show that these methods are less robust, especially in strong light. Nevertheless, to improve the accuracy of detecting rail surface defection, rail component detection must be performed well. Hence, we propose new solutions to achieve the automation of inspecting rail component including rail surface and defects on image processing. The performance of the proposed method is also effective even in strong illumination.

For the rest of this paper, in Sect. 2, the adjustment of the collected image is presented. In Sect. 3, the detail detection of rail is described, and the defect detection is introduced in Sect. 4. Finally, we conclude this paper in Sect. 5.

2 Image Adjustment

2.1 *Compensate for Background Illumination*

Because of nonuniform illumination, edges of the rail often become blurred in the collected images. If preprocessing methods such as image denoising were used alone, the rail edge extraction is not satisfactory. Hence, a related morphological algorithm opening is presented to deal with the uneven illumination between rail and background highlighting the main rail edges in grayscale images.

The grayscale dilation and erosion are similar to spatial convolution in a grayscale image f , which are defined in terms of minima and maxima of pixel neighborhoods. The dilation by structuring element b is defined as [9]

$$(f \oplus n)(x, y) = \max(f(x - x', y - y') + b(x', y') | (x', y') \in D_b) \quad (1)$$

where D_b is the domain of b . As with dilation, the formula of grayscale erosion is

$$(f \odot n)(x, y) = \min(f(x + x', y + y') - b(x', y') | (x', y') \in D_b) \quad (2)$$

The opening of grayscale image f by structuring element b is defined as

$$f \circ n = (f \odot n) \oplus n \quad (3)$$

Applying the opening to a collected rail image, as a result, most of the grayscale information is preserved. But the difference between the original and the opened

image is calculated to emphasize the rail edge, resulted in Fig. 1. In this way, an ideal image of an obvious rail edge can be generated.

2.2 Image Rectification

It can be found that the above rail image has an angle trouble, creating problem in the next rail detection. In order to resolve this problem, one straight line that is parallel to the rail edge is needed by applying the Hough transform, and then the angel of this line is set to be vertical. However, binary processing of grayscale image is needed before the transform. In this section, the optimum global thresholding algorithm based on local maximal between-cluster variance $\sigma_B^2(k)$ is used to gain image threshold automatically [9].

$$\sigma_B^2(k) = P_1(k)[m_1(k) - m_G]^2 + P_2(k)[m_2(k) - m_G]^2 \quad (4)$$

where $P_1(k)$ is the probability of set C_1 occurring with level $[0, 1, 2, \dots, k]$, $P_2(k)$ is the probability of set C_2 occurring with level $[k + 1, \dots, L - 1]$, the terms $m_1(k)$ and $m_2(k)$ are the mean intensities of the pixels in sets C_1 and C_2 , and the term m_G is the global mean of the entire image.

As to Hough transform, it is an excellent way to search and find straight lines in the binary image [9], and the important idea about it is the transformation of parameter space. The line is a set of discrete points in the image; line detection is realized by the transform through a discrete polar coordinate formula. Applying the Hough transform to the previous binary track image, the needed line is found, the result is shown in Fig. 2.

It can be seen clearly that there are three straight lines in Fig. 2b. It seems like these lines are parallel to each other; actually, the slope of the lines is different. To achieve image rotation correction accurately, the longest one that is the closest to the rail edge is chosen from the three ones. The rotated result is shown in Fig. 2c.

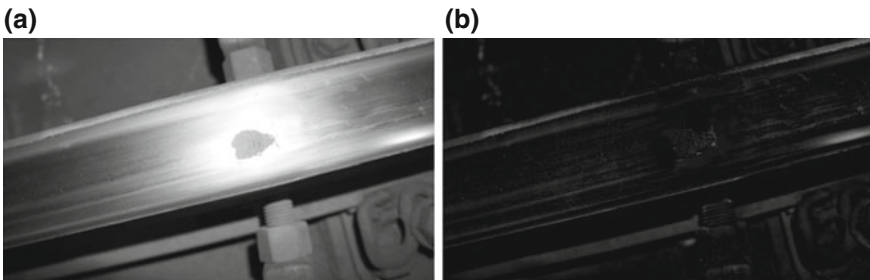


Fig. 1 Compensating for illumination: the left is original image, and right is the resulted image

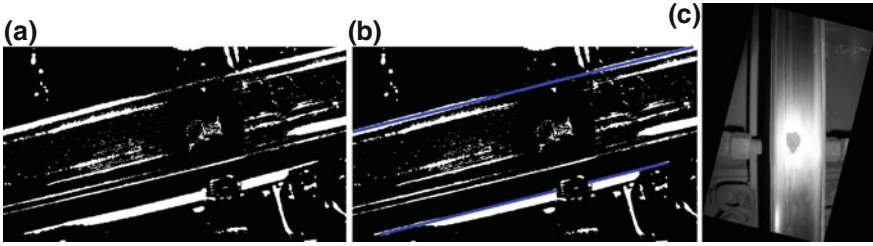


Fig. 2 The resulted picture of rectification

3 Rail Detection

Based on the results above, it is found that the gray value of the adjusted rail surface is basically continuous in the vertical and horizontal direction, while there is obvious difference in the horizontal. But in other regions, there is a significant difference in both directions. In view of this distinguish point, the vertical gradient operator is presented to enhance the vertical rail edge of the image. The operator templates p and q are shown below [10]:

$$p = \begin{bmatrix} -2 & 4 & -2 \\ -2 & 4 & -2 \\ -2 & 4 & -2 \end{bmatrix} \quad q = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (5)$$

$$G(i, j) = \sum_{i=2}^{h-1} \sum_{j=2}^{w-1} (|p * f(i, j)| + |q * f(i, j)|) \quad (6)$$

where h , w are the height and width, respectively, $G(i, j)$ is the resulted image, $f(i, j)$ is the original image, and $*$ is the convolution notation.

After the above template operation, the left and right borders of the rail surface are strengthened. Figure 3a shows the result.

Although there are still some disturbances, these gray values of the rail edge occupy the absolute advantage in the vertical direction, which provide a way for locating and cutting the rail surface. The vertical projection method is proposed to determine one rail edge in the image. First, its binary image is computed using the optimum global thresholding mentioned above and then the value is summed up for each column. One such magnitude map is plotted in Fig. 3b. By analyzing the vertical projection, the results appear many peaks inevitably, but the maximum value is corresponding to one rail boundary indicated by the blue line necessarily, it can be clearly shown in Fig. 3b.

According to the analysis and Fig. 3, one edge can be found accurately, while it does not confirm whether it belongs to the left or right, and the other cannot be determined only by the projection peak value. In the light of these observations, a

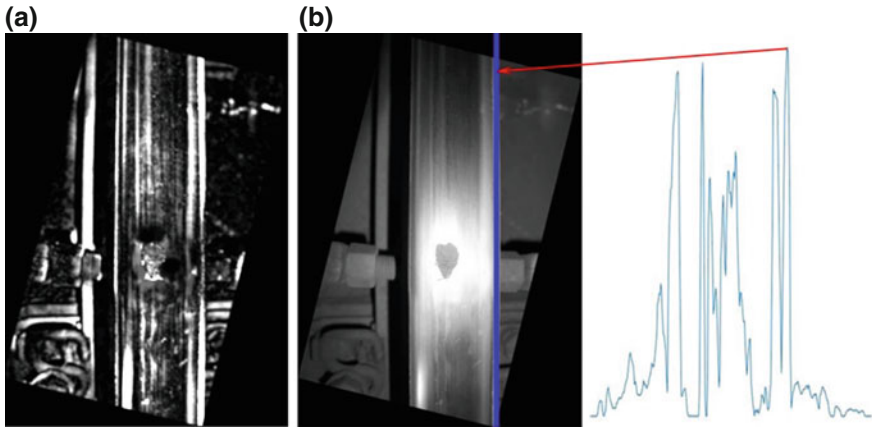


Fig. 3 Edge strengthen and the vertical projection

method of searching the midline of rail edges is proposed based on window statistics, offering a reference to decide the left or right side further.

Sum up all magnitudes within a window that is centered on each column [4], the window size is approximately equal to the width of the rail. Figure 4a shows the summation result, from which there is a distinct local maximum that corresponds to the midline of the rail surface. In the plot, the maximum is found, and the rail surface's left or right edge is determined and marked. Then, one candidate of the other edge is defined based on distance between the determined edge and midline. In consideration of peak value that corresponds to the edge between rail and non-rail region, another candidate is defined by maximum peak value of the vertical projection within 10 pixels around the first candidate (indicated in yellow).

Next, for an actual, its immediate left and right regions present diverse textures; as the rail surface is relatively smooth and other regions have irregular surface, the likelihood of being the other rail edge for each candidate line C_i is measured in terms of pixel intensity variation of binary image. 5 pixels to each candidate's left and right first are defined, respectively. Then sum up the intensity value of pairs of pixels over the whole length of the line C_i . Note that the parameter values such as 5, 10 used here are empirically determined. Finally, the absolute difference between the left 5 pixels summation and the right for each candidate C_i is calculated; moreover, the maximum difference value indicates which one candidate is the other edge (indicated in blue). Figure 4 shows an example of such edge candidates in yellow and accurate vertical edges of the rail in blue.

$$\text{edge} = \max \left\{ \left| \sum_{n=-5}^0 vp(C_i + n) - \sum_{n=0}^5 vp(C_i + n) \right| \right\} i = 1, 2 \quad (7)$$

where vp is the vertical projection, and C_i is the column of one edge candidate.

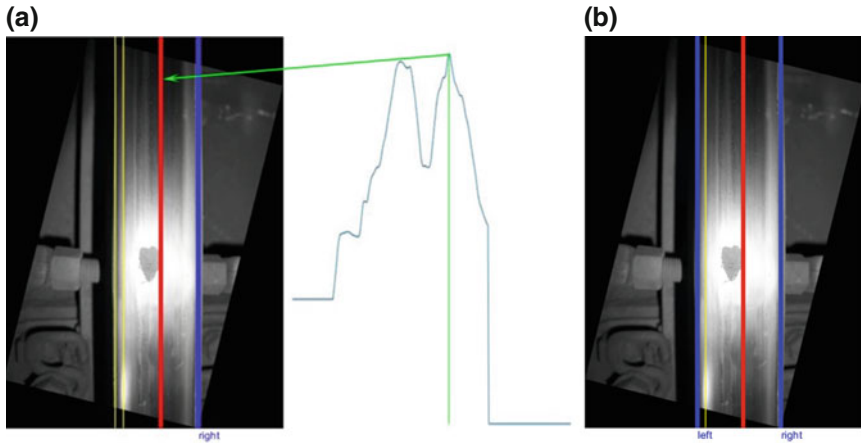
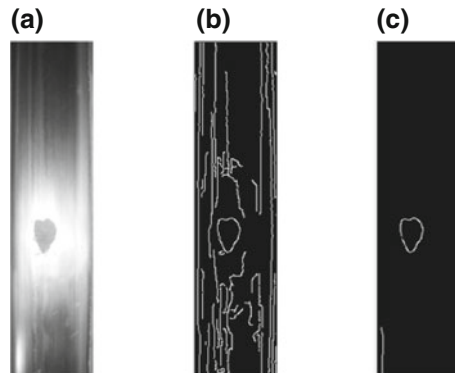


Fig. 4 The rail surface location; the edge of the rail surface is in blue

Fig. 5 The default result and improved result



4 Track Surface Defect Detection

In consideration of the location error of rail edge, a part of edge's pixel is shifted to non-rail region, and the deformity of the upper and lower is cut off simultaneously when track surface region is segmenting. The surface segmentation is shown in Fig. 5a. For the system, the only surface defect is interested defect is interesting for the system. The Canny edge detector with 2D recursive filters is used to detect the defect. There are two critical points that are non-maximum suppression and hysteresis thresholding about this detector [11].

The default Canny result is shown in Fig. 5b. As clearly seen from the figure, although computed edges in default parameter are rich, excessive interference is generated for edges of the interested defect in rail surface. In this paper, according to the grayscale value characteristic of the defects, the hysteresis thresholding is

adjusted to reduce unnecessary details. The result can be seen in Fig. 5c, which is superior by far to the previous result.

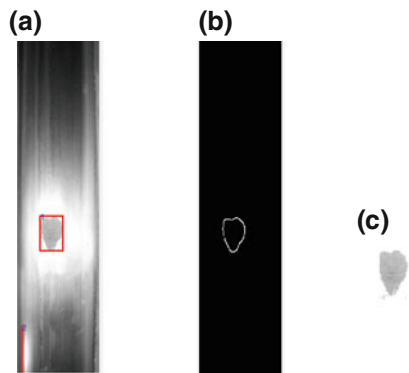
In the improved Canny resulted image, defect in the rail surface become obvious. To label and extract the object, an approach named connected-component labeling is used to search the foreground. This method is an algorithmic application of gray theory, where subsets of connected components are uniquely labeled based on a given heuristic.

However, some noise and rusty defect-free regions are divide as foreground by connected-component labelling, so some foreground components may not be true defects. In view of improving the detection’s accuracy to reduce false positive, the connected component is labeled and recognized as the following rules.

On the one hand, only the first four connected regions are considered, because the system should be more interested in the critical ones. And according to the arrangement of labeled sequence, the possible defects are further marked directly with the minimum boundary rectangle in the rail surface, and numbers are used to mark the importance of the labeled region. That is, the smaller number and the thicker line of the outer rectangle indicate the greater probability of real defect. The result is shown in Fig. 6a.

On the other hand, according to the practical railway maintenance in actual operation, the area of one labeled defect should be greater than a threshold. With the minimum area as the threshold, the defect is identified that is shown in Fig. 6b. Therefore, the defect image (seen in Fig. 6c) is acquired by completely removing unmarked background with the minimum boundary rectangle. It also lays a substantial foundation for further classification of track defect.

Fig. 6 The labeled connected component and defect extraction



5 Conclusion

This paper has shown the computer image processing methods can be successfully used to build an effective application for the detection of rail surface defect and the detection can perform with the collected field images. In the future, this work can be applied to study a complex system that monitors the defection of the railway components, which guarantees the safety of the railway operation.

Acknowledgements This work is partly supported by Chinese National Key Project of Research and Development (Contract No. 2016YFB1200400-2) and State Key Lab of Rail Traffic Control and Safety (Contract No. RCS2016ZT006), Beijing Jiaotong University, Beijing, China.

References

1. Track inspection system V-CUBE [Internet] (2017) Mermecgroup.com. Available from: <http://www.mermecgroup.com/en/01/inspect/track-inspection/524/track-inspection-v-cube.php>
2. TrackVue| Rail Vision Europe Ltd. [Internet] (2017) Rail-vision.co.uk. Available from: <http://www.rail-vision.co.uk/trackvue>
3. Berry A, Nejtkovsky B, Gilbert X, Tajaddini A (2008) A high speed video inspection of joint bars using advanced image collection and processing techniques. In: Proceedings of World Congress on railway research, vol 290
4. Li Y, Trinh H, Haas N, Otto C, Pankanti S (2014) Rail component detection, optimization, and assessment for automatic rail track inspection. *IEEE T Intell Transp* 15(2):760–770
5. Feng H, Jiang Z, Xie F, Yang P, Shi J, Chen L (2014) Automatic fastener classification and defect detection in vision-based railway inspection systems. *IEEE T Instrum Measur* 63(4):877–888
6. Yu H (2013) Research on defects inspection technology for rail surface based on machine vision. Master's thesis, Hunan University (in Chinese)
7. Yang D (2014) Railway fastener status detection under non-uniform illumination based on image processing. Master's thesis, Southwest Jiaotong University (in Chinese)
8. Nping C (2014) Detection of rail surface defects based on image. Master's thesis, Central South University (in Chinese)
9. Gonzalez RC, Woods RE, Eddins SL (2010) *Digital image processing using MATLAB®*. McGraw Hill Education
10. Jia M (2005) Design and research of recognition system on railway surface defects. Master's thesis, Southwest Jiaotong University (in Chinese)
11. Nixon MS, Aguado AS (2012) *Feature extraction & image processing for computer vision*. Academic Press

A Hybrid Temporal-Spatio Fusion Algorithm for Moving Pedestrian Detection in Traffic Scenes

Aili Wang and Yuanyun Sun

Abstract To accurately extract traffic information, the pedestrian detection algorithm of combining time domain, edge and internal–external marker information is proposed. First, a rough motion mask is extracted based on the temporal change of adjacent multi-frames and Canny edge detection, which effectively solves D -value localization and noise problem. Second, to improve the accuracy of airspace segmentation and eliminate over-segmentation, spatial mask image is received by introducing the technology of quadratic reconstruction, internal–external marker, and watershed transformation. Finally, an accurate pedestrian detection mask is obtained by integrating the rough mask and spatial mask image. The experiment results using Beijing South Station video show that this method can detect the complete target information and obtain the better pedestrian detection results in dynamic traffic scene.

Keywords Intelligent transportation · Temporal change of multi-frames
Quadratic reconstruction · Internal–external marker · Watershed
Temporal-spatio fusion

1 Introduction

With the development of the traffic video monitoring technology, the intelligent detection technology of traffic information has been widely concerned and plays an important part in Intelligent Transportation System. The research of moving pedestrian segmentation algorithm in video sequence has vital academic significance and practical value, many scholars devoted to the field and put forward many segmentation algorithm [1–3]. However, due to hardware limitations, the devel-

A. Wang (✉) · Y. Sun
China Railway Information Technology Center, Beijing 100038, China
e-mail: wangaili20050722@163.com

opment history of video monitoring technology is relatively short, a lot algorithm which is only for a specific scenario and the involved various subjects such as image processing, pattern recognition and artificial intelligence, etc. these factors caused incomplete theory. So many problems need to be solved and the effective methods and techniques need to be developed.

Target segmentation algorithm is mainly divided into three areas: temporal segmentation, spatial segmentation, and temporal-spatio segmentation. The temporal segmentation is based on the uniformity of sport information. For example, the history area detection method is proposed by Mech who obtains the detection templates according to threshold segmentation between two adjacent filtered frames [4–6], but it is sensitive to the noise for using only the difference information of two frames. In spatial segmentation, the image will be divided into multiple similar areas based on grayscale, color and texture features, etc. The typical spatial methods include region division, pixel clustering, and watershed method [7]. The watershed method which can get accurate and continuous sealing edge is the focus of attention, but the over-segmentation phenomenon is usually caused by the noise [8]. In temporal-spatio segmentation, first, the moving areas of image are positioned by temporal segmentation, and then the precise semantic information is obtained by spatial segmentation, finally, the precise segmentation objects are required by incorporating the temporal and spatial results [9]. It will be the future development direction, but the accuracy of temporal-spatio integration is influenced by the problems of the rough detection mask, over-segmentation, and threshold selection. To accurately extract traffic information, a hybrid temporal-spatio fusion algorithm for moving pedestrian detection in traffic scenes is proposed.

2 Algorithm Idea

The schematic diagram proposed algorithm is shown in Fig. 1. In initial segmentation, the time domain mask is extracted by adjacent three-frames difference, then edge detection mask is obtained by Canny edge detection method, finally the initial detection mask can be obtained by integrating time domain mask with edge detection mask and using adaptive morphological method to optimize image. In spatial segmentation, first, quadratic reconstruction technique is introduced for image processing, second, the internal–external region is marked according to the local maximum value of gray images and neighborhood information, third, the spatial mask image is extracted by modifying the original gradient and watershed transform. During the fusion process, the spatial segmentation results are projected into initial detection area and the final video detection image is extracted after post-proposing.

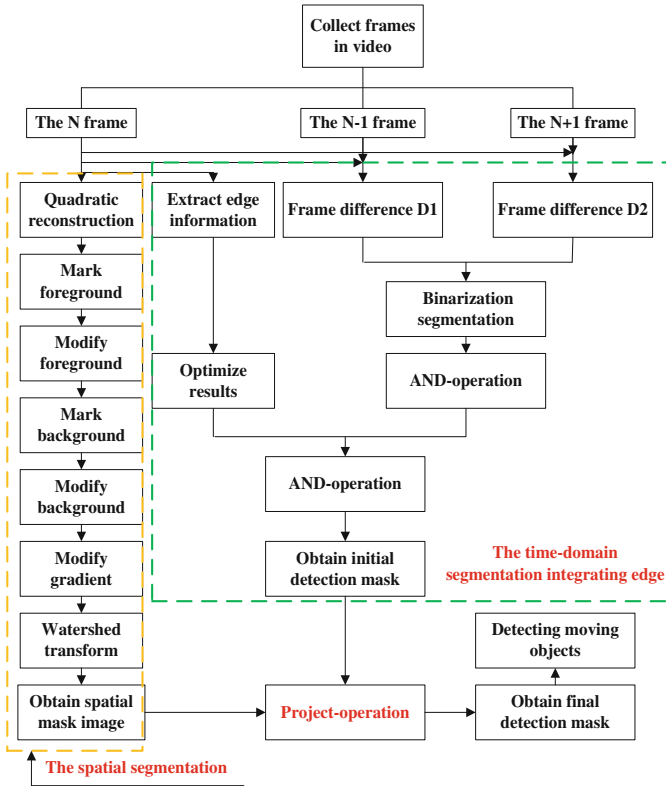


Fig. 1 Algorithm principle diagram

3 Algorithm Description

3.1 Time Domain Segmentation of Integrating Edge Information

3.1.1 The Adjacent Three-Frames Difference Method

Time segmentation is based on time domain, namely computing the difference value of former frame and after frame. Because of the different moving information between foreground and background, the approximate areas of moving objects are obtained by the inter-frame difference method. In order to extract the initial moving detection mask, this paper adopts the adjacent three-frame difference method (the previous, the current and the next frame) to detect moving targets.

Assuming $I(x, y, t)$ is the current frame, $I(x, y, t - 1)$ is the previous frame and $I(x, y, t + 1)$ is the next frame, difference formula is expressed as

$$D_{t,t-1}(x, y, t) = \begin{cases} 1 & |I(x, y, t) - I(x, y, t - 1)| \geq T \\ 0 & |I(x, y, t) - I(x, y, t - 1)| < T \end{cases} \quad (1)$$

$$D_{t,t+1}(x, y, t) = \begin{cases} 1 & |I(x, y, t) - I(x, y, t + 1)| \geq T \\ 0 & |I(x, y, t) - I(x, y, t + 1)| < T \end{cases} \quad (2)$$

In order to extract the precise target edge information, AND-operation between the different results $D_{t,t-1}(x, y, t)$ and $D_{t,t+1}(x, y, t)$ is used to determine the motion detection mask of the current frame, the formula is expressed as

$$\text{Time_mask}(x, y, t) = D_{t,t-1}(x, y, t) \cap D_{t,t+1}(x, y, t) \quad (3)$$

3.1.2 Edge Information Detection

In order to prevent the local tone changes from causing the interference of target extraction, the Canny edge detection method is adopted in the temporal computation to eliminate non-moving areas. Concrete implementation steps are as follows:

- ① Image smoothing: to get rid of noise by using two-dimensional Gaussian function, fixed filter ($\delta = 1$) is adopted in this paper to process video images.

$$I(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right] \quad (4)$$

- ② Calculate gradient amplitude and direction: the gradient amplitude and direction of each point in image are obtained by using first-order differential operator, the calculation formula is

$$\begin{cases} G_x = \frac{I(x,y+1) - I(x,y) + I(x+1,y+1) - I(x+1,y)}{2} \\ G_y = \frac{I(x,y) - I(x+1,y) + I(x,y+1) - I(x+1,y+1)}{2} \\ G(x, y) = \sqrt{G_x^2 + G_y^2} \\ \theta(x, y) = \arctan\left[\frac{G_x}{G_y}\right] \end{cases} \quad (5)$$

- ③ Non-maximum suppression: assuming the point (x, y) in the gradient image is the center of neighborhood ($S = 3 \times 3$), the interpolation calculation is conducted along the gradient direction. If the gradient

amplitude of point (x, y) is greater than the adjacent two interpolations in same direction that point (x, y) will be marked the candidate edge point, otherwise to be marked non-edge point.

- ④ Dual-threshold segmentation: set high threshold T_h and low threshold T_s , the strong edge image and the weak edge image will be connected according to the set threshold so as to extract the edge detection image $Edge_mask(x, y, t)$.
- ⑤ Optimization processing: using morphological thinning to optimize the extracted edge results and using *Endpoints* function to eliminate some irrelevant areas, thus the edge detection image of high signal-to-noise ratio and low mean square deviation is obtained.

3.1.3 Extract the Initial Detection Mask Image

In order to eliminate the influence of background on target extraction, AND-operation between the time domain result and candy edge detection result will remove noise effectively and get the initial detection mask image $Initial_mask(x, y, t)$.

$$Initial_mask(x, y, t) = Time_mask(x, y, t) \cap Edge_mask(x, y, t) \quad (6)$$

3.2 Spatial Segmentation Based on Watershed

3.2.1 Quadratic Reconstruction

To mark foreground targets clearly, the image processing is developed by adopting the morphological reconstruction technology in this paper. The common reconstruction technology is based on erosion and dilation operations, such as open and close reconstruction [10].

The specification of this method is:

For an image $f(x, y)$, reference image $g(x, y)$ and structural element b , the open reconstruction $O_b^{rec}(f, g)$ can be represented by

$$\begin{cases} D_b^1(f, g) = \min(f \oplus b, g) \\ D_b^i(f, g) = \min(D_b^{i-1}(f, g) \oplus b, g), \quad i = 1, 2, 3, \dots \\ D_b^{rec}(f, g) = D^n(f, g) \\ O_b^{rec}(f, g) = D_b^{rec}[(f \circ b), g] \end{cases} \quad (7)$$

Close reconstruction $C_b^{\text{rec}}(f, g)$ can be represented as follows:

$$\begin{cases} E_b^1(f, g) = \max(f \ominus b, g) \\ E_b^i(f, g) = \max(E_b^{i-1}(f, g) \ominus b, g), \quad i = 1, 2, 3, \dots \\ E_b^{\text{rec}}(f, g) = E_b^n(f, g) \\ C_b^{\text{rec}}(f, g) = E_b^{\text{rec}}[(f \bullet b), g] \end{cases} \quad (8)$$

Based on above knowledge, the current frame $I(x, y, t)$ is processed by adopting the mixed reconstruction operation of opening and closing. The “disk” with radius r ($r = 1, 3, 5, 7, 9$) is treated as the shape of structure element. This method is called the multiscale morphological quadratic reconstruction. Calculation formula is

$$H_{b,r}^{\text{rec}}(I(x, y, t), g(x, y, t)) = C_{b,r}^{\text{rec}}[O_{b,r}^{\text{rec}}(I(x, y, t), g(x, y, t)), g(x, y, t)] \quad (9)$$

3.2.2 Region Marking

After reconstruction transform, this paper adopts the internal–external region marking method to extract foreground and background, which can effectively suppress irrelevant targets and avoid over-segmentation.

(1) Internal region marking

In paper, the local maximum values [12] of reconstruction image $H_{b,r}^{\text{rec}}(I, g)$ are marked by analyzing the gray value of reconstruction image, to acquire the foreground marking image $\text{RegionalMax}(x, y)$, as shown in Fig. 2a.

From the figure, we can see that many irrelevant segmentation regions are produced and foreground cannot be marked well. To solve this problem, the previously extracting pixel value of the local maximum region and its neighborhood are analyzed and the set of “High-point” in images is recorded. Namely, the D -value of the gray value in local maximum regions and its neighborhood are calculated which is used to compare with the high threshold and the region whose D -value is higher than high threshold can be retained. This method eliminates the



Fig. 2 Internal marker

irrelevant regions and extracts foreground objects accurately, as shown in formula 10.

$$I_r^{\text{mark}}(x, y, t) = \begin{cases} 1 & \text{RegionalMax}(H_{b,r}^{\text{rec}}(I, g)) - N_L(H_{b,r}^{\text{rec}}(I, g)) \geq \text{Th} \\ 0 & \text{otherwise} \end{cases}, \quad (10)$$

where the variable $I_r^{\text{mark}}(x, y, t)$ is the foreground marking image, Th is high threshold, $N_L(H_{b,r}^{\text{rec}})$ is the neighborhood pixel value.

The images which are obtained through the operation of multiscale morphological quadratic reconstruction are marked, respectively, and then the final foreground marking image $I_{\text{mark}}(x, y, t)$ is acquired according to formula 11. It can be seen that the foreground objects got better marking from Fig. 2.

$$I_{\text{mark}}(x, y, t) = I_{r1}^{\text{mark}}(x, y, t) \cup I_{r2}^{\text{mark}}(x, y, t) \cup I_{r3}^{\text{mark}}(x, y, t) \cup I_{r4}^{\text{mark}}(x, y, t) \cup I_{r5}^{\text{mark}}(x, y, t) \quad (11)$$

(2) External region marking

The paper employs threshold transform, distance transform, and watershed method to mark external regions. Calculation process is as follows:

- ① Threshold transform: the reconstructed image is processed by employing adaptive binarization method, the calculation process is

$$\begin{cases} \text{BW}(x, y, t) = \begin{cases} 1 & H_{b,r}^{\text{rec}}(I, g) \geq T \\ 0 & \text{otherwise} \end{cases} \\ T = \text{graythresh}(H_{b,r}^{\text{rec}}(I, g)) \end{cases} \quad (12)$$

- ② Optimizing process: the obtained image is optimized, which mainly includes *imerode()*, *bwareaopen()* and *imfill()*, etc.
- ③ Distance transform: the ‘‘Quasi Euclidean distance’’ method is used to transform for thinning background, the computing formula of the smallest neighborhood distance between edge pixel (a, b) and region pixel (m, n) is

$$\text{Dist}(m, n) = \begin{cases} |m - a| + (\sqrt{2} - 1) \times |n - b| & |m - a| > |n - b| \\ (\sqrt{2} - 1) \times |m - a| + |n - b| & \text{otherwise} \end{cases} \quad (13)$$

- ④ Watershed segmentation: the watershed transformation is employed to seek the division ridge line graph after distance transform, the background marking result $I_{B\text{mark}}(x, y, t)$ is

$$I_{Bmark}(x, y, t) = \text{Watershed}(\text{Dist}(BW(x, y, t))) == 0 \quad (14)$$

3.2.3 Gradient Modification

The *Imimposemin* technology is adopted to modify the gradient image. The local minimum regions only appear on the position of internal–external marking, other pixels will be pushed as required, and the other local maximum regions are deleted for eliminating over-segmentation. The modification process is followed:

- ① Calculate initial gradient: transverse gradient image I_x , longitudinal gradient image I_y , and gradient image $I_{\text{gradmag}}(x, y, t)$ are required by using the horizontal–vertical *Sobel* operator.
- ② Modify gradient: the new gradient image $I_{\text{gradmag}}^{\text{end}}(x, y, t)$ is extracted through cover local minimum areas according to the marked position of internal–external marker.

$$\begin{cases} I_x = \text{imfilter}(\text{double}(I), h_x, 'replicate') \\ I_y = \text{imfilter}(\text{double}(I), h_y, 'replicate') \\ I_{\text{gradmag}}(x, y, t) = \sqrt{I_x^2 + I_y^2} \\ I_{\text{gradmag}}^{\text{end}}(x, y, t) = \text{ImposeMin}(I_{\text{gradmag}}(x, y, t), I_{\text{mark}}(x, y, t) | I_{Bmark}(x, y, t)) \end{cases} \quad (15)$$

3.2.4 Extract Spatial Detection Mask Image

The modified gradient image is transformed through watershed method and the final spatial division ridge line $\text{Space_mask}(x, y, t)$ is obtained as follows:

$$\text{Space_mask}(x, y, t) = \text{watershed}(I_{\text{gradmag}}^{\text{end}}(x, y, t)) \quad (16)$$

The results are shown in Fig. 3.

3.3 The Algorithm Steps of Extracting Moving Objects

For extracting the information of moving pedestrians, this article will merge the two methods together that make up their lost information [10]. The basic steps of whole fusion algorithm are as follows:

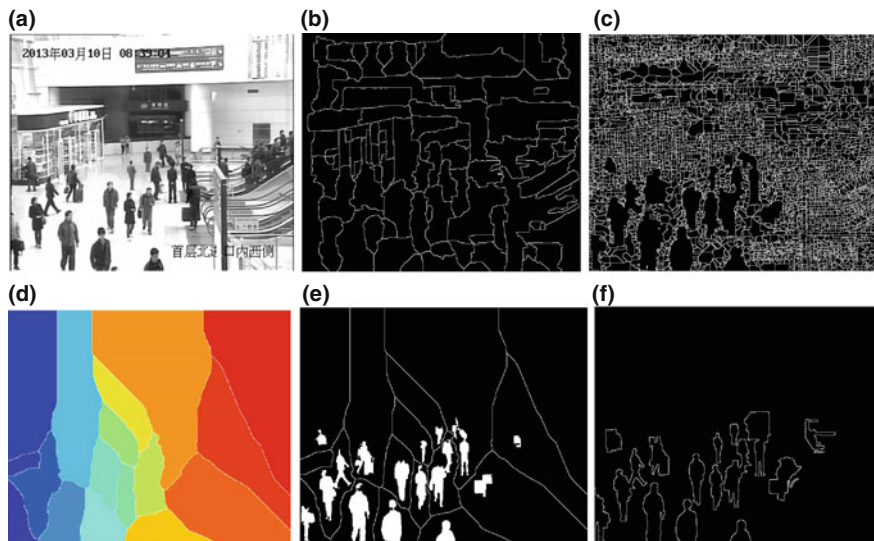


Fig. 3 Spatial segmentation process

- ① **Information fusion:** the spatial segmentation results $\text{Space_mask}(x, y, t)$ are projected into the initial detected mask image $\text{Initial_mask}(x, y, t)$ for obtaining the final detection mask image $\text{Final_mask}(x, y, t)$, as shown in formula 17:

$$F_i(x, y, t) = \begin{cases} 1 & (x, y) \in S_i \&\& \frac{\sum_{(x,y) \in S_i} \text{Initial_mask}(x,y,t)}{N_{S_i}} \geq \text{th} \\ 0 & \text{otherwise} \end{cases}, \quad (17)$$

$$\text{Final_mask}(x, y, t) = \cup \{F_i(x, y, t)\} \quad (i = 0, 1, \dots, n - 1)$$

where S_i is the any region segmented in spatial mask, N_{S_i} is the magnitude of S_i , th is threshold value which between 0.6 and 0.7 generally.

- ② **Information modification:** We took adaptive morphology to modify the final detected mask image $\text{Final_mask}(x, y, t)$. First, the dilation operation with elliptical structure element B is adopted to obtain the closed target edge image $\text{Final_mask1}(x, y, t)$; then the edge information of $\text{Final_mask1}(x, y, t)$ is thinned to relocate edge information; eventually, we did *Imfill* and *Bwareaopen* to obtain the final detected image $\text{End_mask}(x, y, t)$. The optimizing process is shown in formula 18.

$$\begin{cases} \text{Final_mask1}(x, y, t) = \text{Imdilate}(\text{Final_mask}(x, y, t), B) \\ \text{Final_mask2}(x, y, t) = \text{Final_mask1}(x, y, t) \cap \text{Edge_mask}(x, y, t) \\ \text{Final_mask3}(x, y, t) = \text{Bwmorph}(\text{Final_mask2}(x, y, t), 'Skel'Inf) \\ \text{Final_mask4}(x, y, t) = \text{Bwareaopen}(\text{Final_mask3}(x, y, t), TH) \\ \text{End_mask}(x, y, t) = \text{Imfill}(\text{Final_mask4}(x, y, t)) \end{cases} \quad (18)$$

4 Experiment Results

To verify the effectiveness of above algorithm, a target detecting platform by using the combination of C# and MATLAB is developed. The experiment took the video of west entrance of Beijing South Station as its input and aimed at detecting pedestrians in the plane domain.

(1) Comparing target segmentation algorithm

- ① The time domain segmentation based on frame difference: To compare the effectiveness of three different algorithms, including frame difference, Canny edge detecting and the frame difference of integrating edge information proposed in this paper, we collected three continuous frames as input and processed these three continuous frames by using these three different methods. The results are shown in Fig. 4. By comparing three results, we can see clearly that the proposed frame difference algorithm eliminates the “ghosting” phenomenon caused by frame difference and weaken background noises caused by Canny edge detecting. In addition, the edge information of targets have been obtained completely and the result image can be treated as high quality initial segmented mask image, but background noises still produced some interference.
- ② The spatial segmentation based on watershed: the 236th frame of the video is spatial segmented by using traditional watershed algorithm, smooth gradient watershed segmentation algorithm, and the improved watershed algorithm in this paper and the results have been shown in Fig. 5. The segmented independent regions and the edge of moving targets can be clearly seen in Fig. 5c, and the over-segmentation has been controlled effectively, but some edge information has been lost.
- ③ Fusion algorithm: We combined the improved time domain segmentation result and the improved watershed result, as shown Fig. 6. Compared with Figs. 4i and 5c, the edge of segmented image can be clearly seen and the target has been segmented completely without losing information and makes the extraction more accurate and more efficient.



Fig. 4 The contrast experiment of frame difference

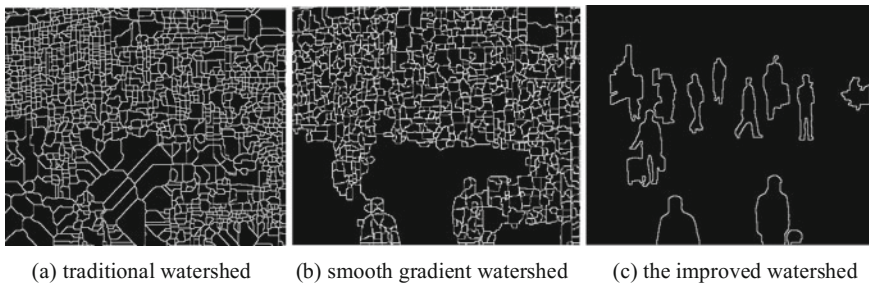


Fig. 5 The contrast experiment of watershed

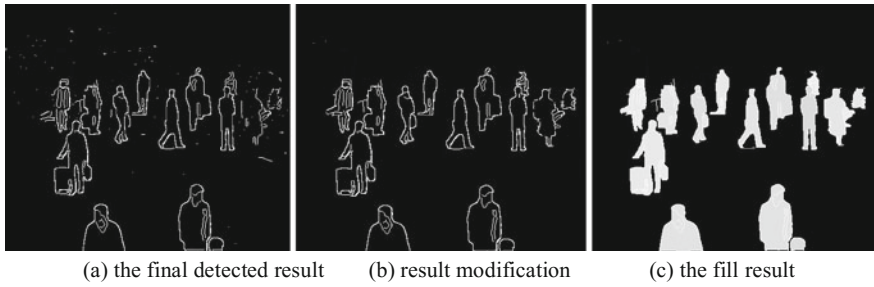


Fig. 6 The results of this paper

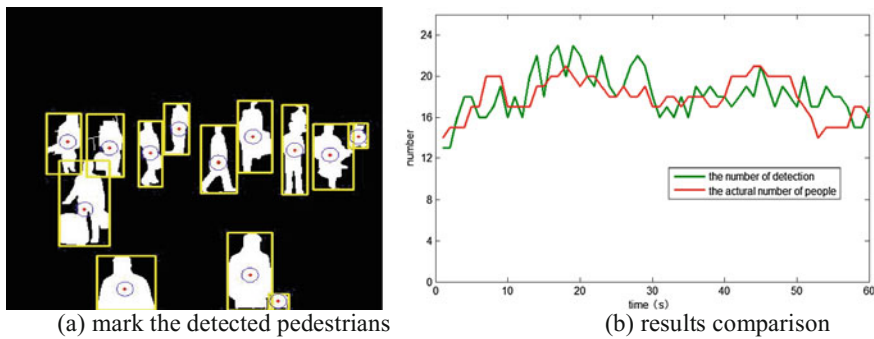


Fig. 7 The detection results

(2) Analyzing the detected results

Figure 7a shows the detecting results of 236th frame. In the detecting region, there are 15 people in total and the detecting result of fusion algorithm is 13 people. In Fig. 7b, the red line displays the real number of people in 60 s at the west entrance of Beijing South Station, the green line shows detection results by the fusion algorithm. Comparing these two line graphs, we conclude that the fusion algorithm works efficiently and has a high accuracy rate.

5 Conclusions

The new detecting algorithm in time-space domain which is developed on MATLAB platform can extract the moving pedestrian in traffic scene accurately. On one hand, the time domain segmentation algorithm based on adjacent multi-frames information and Canny edge detecting is able to eliminate the “ghosting” phenomenon and noises effectively. On the other hand, the improved watershed algorithm can not only extract the edge and outline of moving target but

also eliminate over-segmentation. Finally, an accurate motion pedestrian mask is obtained by integrating the rough mask and spatial mask image. The experiment of Beijing South Station demonstrates that the fusion algorithm is capable of extracting the information of moving targets and has a high accuracy and effectiveness in real-time detecting.

References

1. Meier T, Ngan KN (1998) Automatic segmentation of moving objects for video object plane generation. *IEEE Trans Circ Syst Video Technol* 8(5):525–538
2. Li J, Shao CF, Xu WT, Li J (2010) A real-time system for tracking and classification of pedestrian and bicycles. *Transp Res Rec* 2198:83–92
3. Weng M, Huang G, Da X (2010) A new interframe difference algorithm for moving target detection. In: 2010 3rd international congress on image and signal processing, 2010, pp 285–289
4. Ilea DE, Whelan PF (2011) Image segmentation based on the integration of colour-texture descriptors—a review. *Pattern Recogn* 44(10):2479–2501
5. Zhang Z, Yang G, Liu Z (2009) The principle and application of video object segmentation. Science press, Beijing, pp 52–60 (in Chinese)
6. Patras L, Hendriks EA, Lagendijk RL (2001) Video segmentation by MAP labeling of watershed segments. *IEEE Trans Pattern Anal Mach Intell* 23(3):326–332
7. Zhao L (2012) Research of video object segmentation based on spatial-temporal technology. Master Thesis of Jilin University, 6 (in Chinese)
8. Li J, Lan J, Li J (2013) A novel fast moving target detection method. *J Cent S Univ (Nat Sci)* 44(3):978–984 (in Chinese)
9. Zhu Z, Jiang G, Yu M (2003) New algorithm for extracting moving object based on spatial-temporal information. *J Image Gr*, 2013, 2003 8(4):422–426
10. Rafael CG, Richard EW, Steven LE (2005) *Digital image processing using MATLAB*. Publishing House of Electronics Industry, Beijing, pp 260–280

Research of Online Nondestructive Monitoring of Composite Components Using FBG

Xianfeng Wang, Zhongcheng Jiang, Xiaobo Liu and Jixiong Jiang

Abstract Fiber Bragg grating (FBG) is a new type of passive fiber optic device. It has been mainly used for the health monitoring of large structures such as bridges and buildings, while the application of the online nondestructive monitoring of composite components is less. This paper introduced the composite nondestructive testing method, fiber grating sensor application at home and abroad, as well as FBG sensing principle. Then simulation was performed for 3D model of carbon fiber levitation chassis frame of maglev vehicle under the normal operating condition and extraordinary operation condition. Moreover, the layout design of FBG sensors on the surface of carbon fiber levitation chassis frame was performed using the strain distribution, and then FBG sensing network and its monitoring system were designed. The preliminary design results showed that the online nondestructive monitoring system can meet the requirements of engineering measurement of carbon fiber levitation chassis frame.

Keywords FBG · Composite material · Online monitoring · Levitation chassis frame

1 Introduction

Compared to the metal material, the composite has advantages of anti-fatigue, higher strength to weight ratio, lighter weight, and longer service life, and is widely used in the manufacture of major engineering structures and equipment, such as aerospace, automotive industry, chemical industry, textile and machinery manufacturing [1–4]. Applications of composites in railway transportation are limited to

X. Wang (✉) · Z. Jiang · X. Liu · J. Jiang
The State Key Laboratory of Heavy Duty AC Drive Electric
Locomotive Systems Integration, Zhuzhou 412001, Hunan, China
e-mail: xfwang0812@qq.com

X. Wang · Z. Jiang · X. Liu · J. Jiang
CRRC Zhuzhou Locomotive Co. Ltd., Zhuzhou 412001, Hunan, China

no-load or partial load bearing components, such as interior materials, cab hood, and so on, and it is still in the research and development phase for the car body, bogie, levitation chassis, and so on. During the service under condition of tension compression, impact and fatigue load, the composite materials is prone to be damaged or destructed, which will lead to huge economic losses and casualties, and even disastrous consequences if not taking corresponding measures in time. Therefore, it is very important to use dynamic damage detection and identification for composite structures during service.

The detection methods of composite materials include ultrasonic testing, ray testing, computer tomography, microwave testing, acoustic ultrasonic testing, acoustic emission testing, optical fiber testing, and so on. Fiber optic sensing technology has many advantages, such as small size, high sensitivity, easy to compound, and so on. It is a popular method for testing composite materials in the future. In 1979, NASA organized an optical fiber sensor embedded in composite material to monitor its strain and temperature, the earliest application of fiber optic sensors. In 1989, Mendez et al. proposed the application of fiber optic sensors to concrete structure detection for the first time [5]. After that, researchers in Japan, the United Kingdom, the United States, Germany, and other countries have studied fiber optic sensing systems in civil engineering [6]. In 2004, Jiang Yi and Xu Ying proposed a new layered nondestructive testing method for composite materials. The layered optical fiber sensor was used to detect the delamination of composite materials, and good results were obtained [7].

As one of the fiber sensors, fiber Bragg grating sensor has its unique advantages as follows. More than one grating can be written in an optical fiber, which composed to a sensor array. The distributed sensing can be realized for more than one optical fiber by using wavelength division multiplexing and time-division multiplexing system. The measuring parameter of information carrier of FGB is wavelength, therefore FBG has strong anti-interference ability and not been influenced by the optical fiber grating sensor the fluctuation of the light source, optical fiber connection and coupling loss, and light the polarization state changes. In 2005, Hiroshi Tsuda measured carbon fiber composite materials using fiber grating sensor at first time, and concluded that the fiber grating sensor can distinguish the spread path of acoustic signal, linear or around the damage zone spread, compared with piezoelectric sensor. And Hiroshi Tsuda verified that the detection of acoustic wave using tunable laser system is more sensitive than using broadband optical system [8]. Hiroshi Tsuda proposed that FBG can play a better effect in resonance structure, and verified the feasibility for using fiber grating ultrasonic sensors to detect metal fatigue crack [9, 10]. The length of grating sensor for sensitively detecting the ultrasonic was assured, and the location of the crack tip was realized.

At present, it is still in the initial experimental stage for the research on fiber grating used for composite material detection in China, but has great potential for development. FBG sensors were used for structure damage detection of carbon fiber levitation chassis frame in the paper. Location, direction, and density of FBG sensor network on levitation chassis frame were researched using the strain calculation simulation of levitation chassis frame, and the distributed fiber grating sensing

network system was designed, which provides the preliminary technical support for the application of FBG sensor in carbon fiber levitation chassis structure.

2 Sensing Principle of FBG

FBG is a reflection type of optical fiber grating which was manufactured using a strong ultraviolet laser irradiation with property of spatial periodic on the doped fiber with the ultraviolet photosensitive property. The refractive rate distribution of the doped fiber core is periodic along the axial during production process [11]. When the broadband light is transmitted in the FBG, the mode coupling occurs. The incident light satisfying the condition of the fiber grating is coupled into the reflection mode. The fiber Bragg grating reflection condition is expressed by the equation

$$\lambda_B = 2n_{eff}\Lambda, \tag{1}$$

where λ_B is the reflection wavelength of the fiber grating, Λ is the grating period, and n_{eff} is the effective refractive index of the fiber core. The light satisfying the conditional wavelength of the Eq. (1) is reflected by the grating, and the reflectance is as high as 99%, but the light of other wavelengths does not degrade significantly after passing through the fiber grating. In fact, the grating acts like a narrowband filter.

When the temperature, strain, stress, or other measured physical quantity around the grating change, it will lead to changes in the grating period or core refractive index to produce Bragg wavelength drift. Through detecting the drift of the Bragg wavelength, you can get the physical quantity changes. The fiber Bragg grating is an optical waveguide device in which the refractive index of the core is periodically changed in the axial direction so that the forwardly transmitted energy is coupled to the reverse guide to form a sharp reflection peak with a reflection wavelength period. Its structure and principle are shown in Fig. 1, the broadband light is

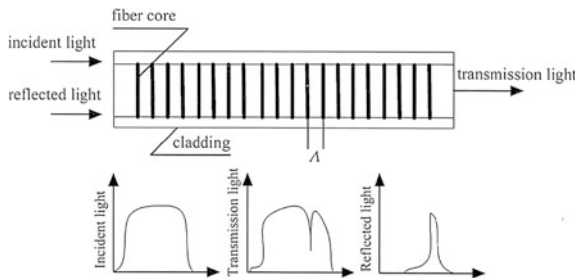


Fig. 1 The structure and principle of fiber grating [12]

emitted by the light transmission to the measured point, a narrowband light is selectively reflected back by the fiber grating, which is transmitted to the wavelength discriminator or wavelength demodulator by the optical splitter, and has photoelectric conversion through the photodetector. Under the action of temperature or stress, when the deformed body is deformed by external temperature or pressure, the wavelength of the narrowband light reflected by the fiber grating will change accordingly, which reflects the change of the corresponding temperature or pressure of the deformed body.

3 Strain Simulation of Carbon Fiber Levitation Chassis Frame

3.1 3D Model of Carbon Fiber Levitation Chassis Frame

Levitation chassis frame was originally made of aluminum alloy structure, which included the arm, stringer, anti-side beam, and traction seat. After using carbon fiber composite material, its weight loss effect reached 36.8%, which is shown in Fig. 2.

3.2 Strain Simulation Results Under Normal Operating Condition

The composite material levitation chassis frame was loaded using a working condition with the largest load in the normal operating conditions, and the finite element simulation was carried out. The results are shown in Fig. 3. The maximum strain is $4238\mu\epsilon$, which is located near the bolt connection hole of the arm. The strain cloud of each carbon fiber component is shown in Fig. 4, which can be used for the arrangement of the sensor network.

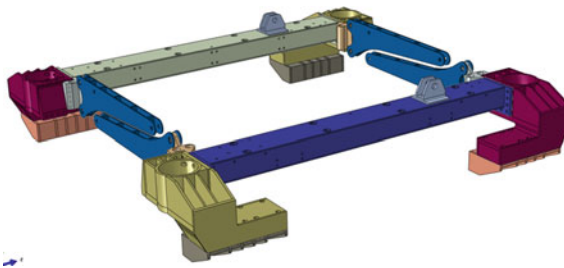


Fig. 2 Composite material levitation chassis frame assembly

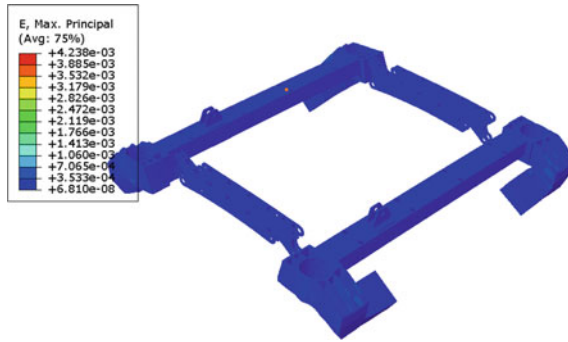


Fig. 3 Structural strain cloud diagram of carbon fiber levitation chassis frame under normal operating condition

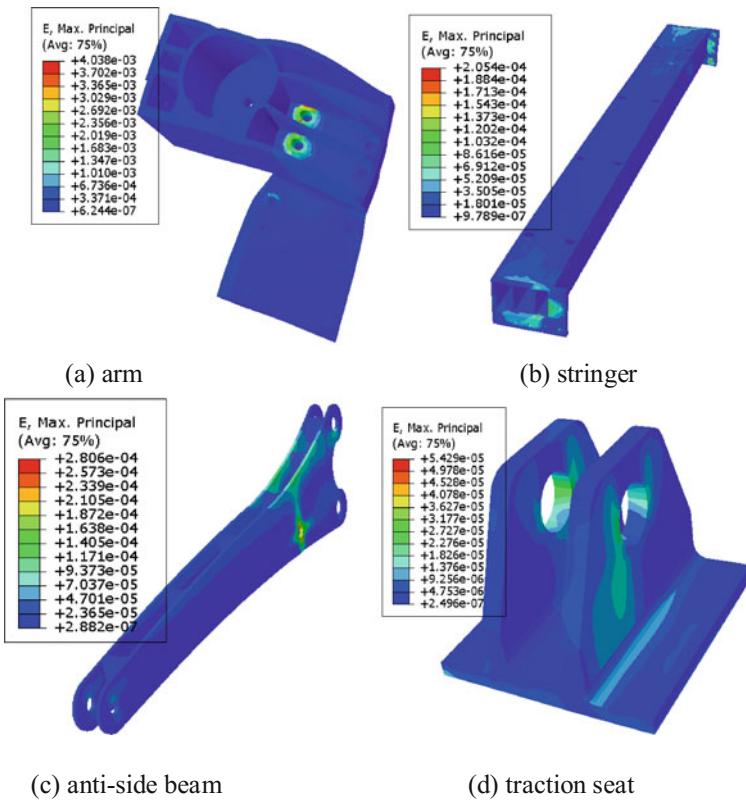


Fig. 4 Structural strain cloud of each component of carbon fiber levitation chassis frame under normal operating conditions

3.3 Strain Simulation Results for the First Extraordinary Operating Condition

The finite element simulation of the carbon fiber levitation chassis frame was performed under the first extraordinary operating condition with the largest vertical load, in which results are shown in Fig. 5. The maximum strain is $5086\mu\epsilon$, which is located near the bolt connection hole of the arm. The strain cloud of each carbon fiber component is shown in Fig. 6, which can be used for the arrangement of the sensor network.

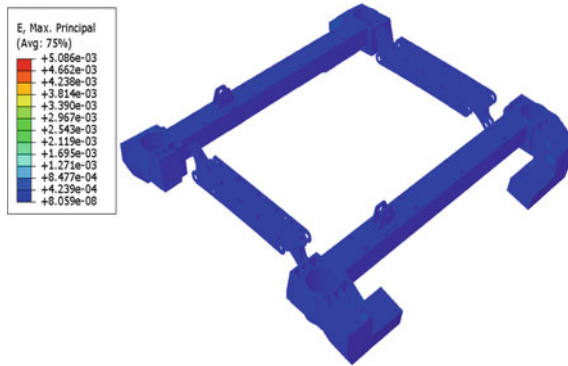


Fig. 5 Structural strain cloud diagram of carbon fiber levitation chassis frame under the first extraordinary operating condition

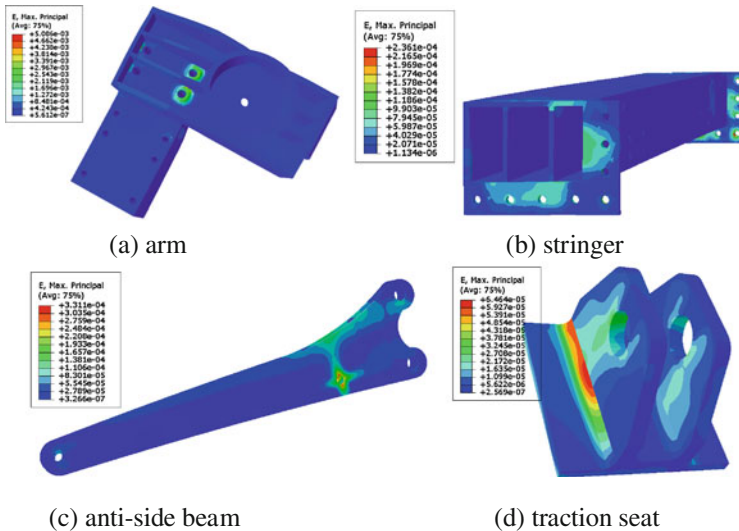
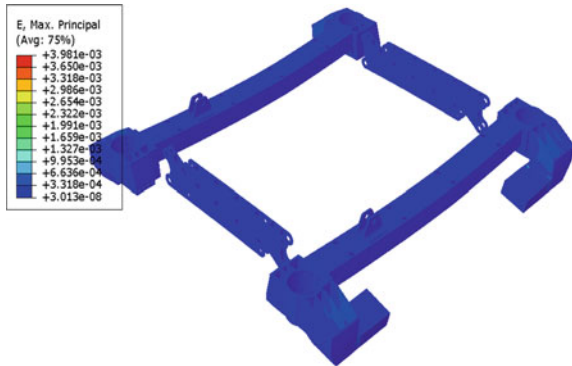


Fig. 6 Structural strain cloud of each component of carbon fiber levitation chassis frame under the first extraordinary operating condition

Fig. 7 Structural strain cloud diagram of carbon fiber levitation chassis frame under the second extraordinary operating condition



3.4 Strain Simulation Results for the Second Extraordinary Operating Condition

The finite element simulation of the carbon fiber levitation chassis frame was performed under the second extraordinary operating condition with the largest longitudinal load, in which results are shown in Fig. 7. The maximum strain is $3981\mu\epsilon$, which is located near the bolt connection hole of the arm. The strain cloud of each carbon fiber component is shown in Fig. 8, and the maximum strain of the traction seat is $505\mu\epsilon$, which can be used for the arrangement of the sensor network.

3.5 Comparison of Strain Simulation Results of Each Operating Condition

The maximum strain of the carbon fiber levitation chassis frame under each operating condition is shown in Table 1. Obviously, the strain of the arm is greater, followed by stringer, anti-side beam, and traction seat smaller.

4 FBG Sensor Network Layout

According to the strain analysis of each component of the carbon fiber levitation chassis frame, FBG sensors were arranged on the outer surface of each component, as shown in Fig. 9. 8 fiber gratings were arranged on each arm at position with maximum or larger strain, and 32 fiber gratings were arranged on four arms. 6 fiber gratings were arranged on each anti-side beam at position with maximum or larger strain, and 24 fiber gratings were arranged on four anti-side beams. 16 fiber gratings were arranged on each group of stringer and traction seat at position with maximum

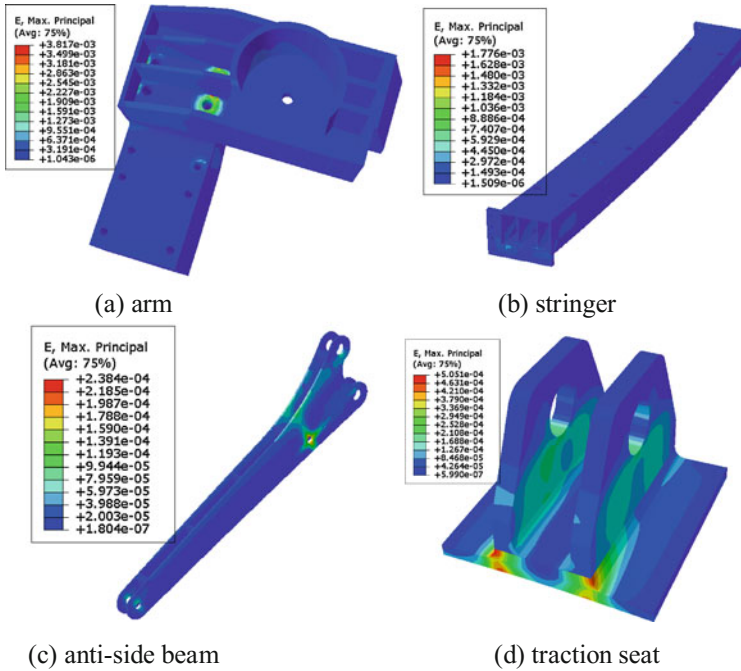


Fig. 8 Structural strain cloud of each component of carbon fiber levitation chassis frame under the second extraordinary operating condition

Table 1 Maximum strain of each component of carbon fiber levitation chassis frame ($\mu\epsilon$)

Parameters	Arm	Stringer	Anti-side beam	Traction seat
Normal operating conditions	4038	205	281	54
The first extraordinary operating condition	5086	236	331	65
The second extraordinary operating condition	3817	1176	238	505

or larger strain, and 32 fiber gratings were arranged on two group. Therefore, 88 fiber gratings in all were arranged on a carbon fiber levitation chassis frame.

There are four sensing lines designed for FBG sensor network of the carbon fiber levitation chassis frame, as shown in Fig. 10. The first sensing line has 18 fiber gratings in all, which consists of 6 fiber gratings on an anti-side beam, 8 fiber gratings on one arm, 4 fiber gratings at one end of the stringer, and the second sensing line is the same with the first. The third sensing line has 26 fiber gratings in all, which consists of 6 fiber gratings on an anti-side beam, 8 fiber gratings on one arm, 12 fiber gratings on one end of one stringer and one traction seat, and the fourth sensing line is the same with the third.

Fig. 9 The location of FBG sensors for each component of the carbon fiber levitation chassis frame

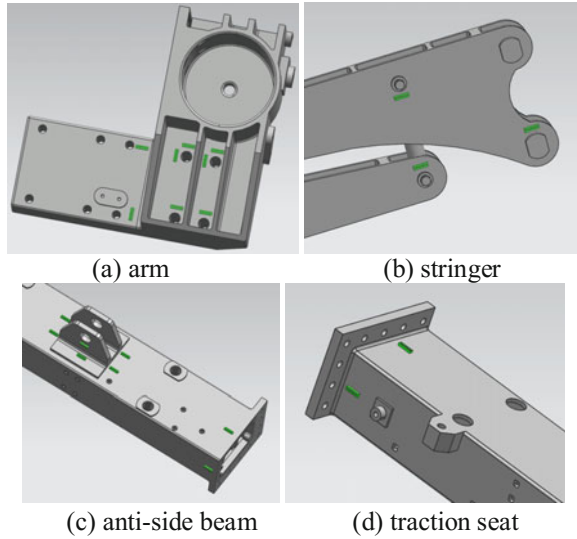
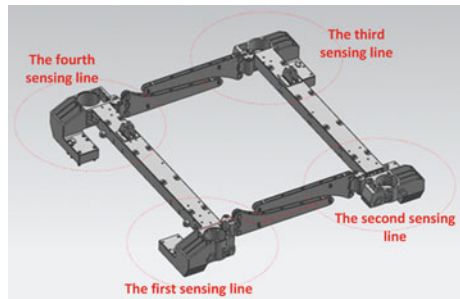


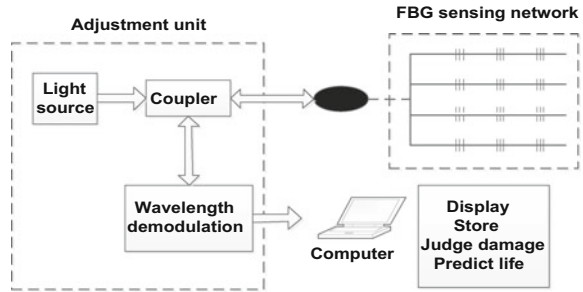
Fig. 10 Distribution of four sensing lines of FBG sensor network in carbon fiber levitation chassis frame



5 Structural Health Monitoring System for Carbon Fiber Levitation Chassis Frame

Carbon fiber suspension structure health monitoring system includes hardware system and software system. The hardware system is mainly composed of FBG sensor network, demodulator, and computer, as shown in Fig. 11. Sensor layout and acquisition system should be integrated to design. The stress concentration of the levitation chassis frame was monitored using fiber grating sensors, and the strain signals were collected and demodulated by the demodulator. Then, the strain signals of each component of the levitation chassis frame were transmitted to the computer and displayed online.

Fig. 11 Illustration of structural health monitoring system for carbon fiber levitation chassis frame



6 Conclusion

After describing the sensing principle and application of FBG sensor, the strain calculation simulation was performed for carbon fiber levitation chassis frame, and the maximum strain positions of each composite component were found, in which FBG sensors were arranged. According to the symmetrical structure of the carbon fiber levitation chassis frame and the decoupling connection structure of the anti-side beam, the fiber grating sensing network was designed as four sensing lines, and the hardware and software systems were designed for the structure health monitoring system. The design of online nondestructive monitoring system provided the pretechnical support for application of FBG sensor in carbon fiber levitation chassis frame.

References

1. Khashaba UA, Khair AI (2017) Open hole compressive elastic and strength analysis of CFRE composites for aerospace applications. *Aerosp Sci Technol* 60:96–107
2. Boland CS, Kleine RD, Keoleian GA et al (2016) Life cycle impacts of natural fiber composites for automotive applications: effects of renewable energy content and lightweighting. *J Ind Ecol* 20(1):179–186
3. Friedrich K (2016) Carbon fiber reinforced thermoplastic composites for future automotive applications. *AIP Conf Proc* 1736:020001
4. Fiorea V, Calabrese L, Proverbiob E et al (2017) Salt spray fog ageing of hybrid composite/metal rivet joints for automotive applications. *Compos B Eng* 108:65–74
5. Fuhr PL, Huston DR (1998) Corrosion detection in reinforced concrete roadways and bridges via embedded fiber optic sensors. *Smart Mater Struct* 7(2):217–228
6. Zonta D, Pozzi M, Wu H (2008) Experimental evaluation of a smart prestressed element instrumented with multiplexed interferometric fiber optic sensors. In: *Proceedings of the 4th European workshop on structural health monitoring*, pp 863–870
7. Leung CKY, Yang Z, Xu Y et al (2005) Delamination detection in laminate composites with an embedded fiber optical interferometric sensor. *Sens Actuators, A* 119(2):336–344
8. Tsuda H (2006) Ultrasound and damage detection in CFRP using fiber Bragg grating sensors. *Compos Sci Technol* 66(5):676–683

9. Tsuda H, Lee JR, Guan YS et al (2007) Investigation of fatigue crack in stainless steel using a mobile fiber Bragg grating ultrasonic sensor. *Opt Fiber Technol* 13(03):209–214
10. Yashiro S, Ono R, Ogi K (2014) Prediction of shear-cutting process of CFRP cross-ply laminates using smoothed particle hydrodynamics. In: *Proceedings of the 16th US-Japan conference on composite materials*, San Diego, CA, 8–10 Sept 2014, 747
11. Hao X (2013) *Technology and application of dynamic testing*. Electronic Industry Press, Beijing, 1–16 (in Chinese)
12. Zhang Y, Zhao CY, Chen JX (2013) Damage detection of composite materials based on fiber Bragg grating sensor technology. *Chin J Quantum Electron* 30(5):608–614 (in Chinese)

Study on the Adhesion Optimization Control Strategy and Its Application Based on Trend Forecasting

Yunxin Fan, Biao Ye, Zhongcheng Jiang and Jixiong Jiang

Abstract Wheel–rail adhesion is a complex question which is quite easily interrupted by external time-varying signal. Furthermore, the maximum adhesion point is time variant. It is difficult to find the maximum adhesion point for adhesion control procedure. This paper gives a deeply analysis on wheel–rail adhesion mechanism and then proposes an adhesion optimization strategy based on trend forecasting. This strategy is robust which is not sensitive to external environment, but mainly related to train’s operation condition. For validating its feasibility and reliability, the traction performance test in wet rail condition of HXD1 electric locomotive before and after adhesion program optimized is preceded. The test results indicate that the adhesion strategy proposed in this paper is useful to search the maximum adhesion point and improve the traction ability. It also indicates that this strategy is right and robust.

Keywords Control strategy · Adhesion optimization · Trend forecasting
Traction ability

1 Introduction

With the development of the economy of China economy, railway transportation as one of the most important modes of transportation has been inducted into a rapid development period. In order to satisfy the fast transportation requirement of energy, goods, and other lot cargo, more and more AC electric locomotives, which are much powerful, much fast, and more reliability are produced to replace DC

Y. Fan (✉) · B. Ye · Z. Jiang · J. Jiang
The State Key Laboratory of Heavy Duty AC Drive Electric
Locomotive Systems Integration, Zhuzhou 412001, Hunan, China
e-mail: stevenhust@163.com

Y. Fan · B. Ye · Z. Jiang · J. Jiang
CRRC Zhuzhou Locomotive Co., Ltd., Zhuzhou 412001, Hunan, China

electric locomotive. There are 10836 AC electric locomotives in China until 2013, 52.1% in all locomotives [1]. They are distributed among 54 locomotive depots in 18 railway bureaus. As to the large national area and its unbalance economy, the railway lines condition is quite different. There are not only lots of level I railway lines which are flat and firm, but also many level II, III railway lines which are ramp more and unsteady. One locomotive could be used in different level railway lines.

Due to the influence of external environment condition (such as rain, snow and oil et al.), the surface condition of iron rails is also quite different. That is to say, the adhesion coefficient between the iron rail and wheel is time varying. In order to reduce the number of electric locomotive in service in the same traction load, improving the traction performance of each single electric locomotive is one of the most effective methods. Searching optimum adhesion control point becomes a really important work in the control of modern AC electric locomotive, because the maximum adhesion force is affected by the choosing the optimum adhesion control point [2–4].

Han Changhu [5] did some of adhesion tests in simulating raining condition for HXD3 AC electric locomotive. This work provides some experience to optimize adhesion control system. Ge Xinglai [6] proposed to carry out online traction performance test for DJ1 AC electric locomotive used in Daqin railway. Zhang Wei-hua and Yang Ying [7, 8] studied the adhesion mechanism of high-speed train and electric locomotive. Li Jianghong [9] proposed some optimum methods for adhesion control based on the research of adhesion mechanism. Wang Hui [10] presented an adhesion control method based on wavelet analysis. Ren Qiang [11] proposed to study adhesion considering rail-wheel condition recognition. Chen Kai [12] did some research in optimization adhesion control based on the building of adhesion control simulation platform.

Based on the analysis of previous research, this paper proposes an adhesion control method based on trend forecasting. Then, its effectiveness and robust are verified through a traction performance test carried out in a typical heavy duty freight electric locomotive controller.

2 Adhesion Analysis

2.1 Adhesion Mechanism

Some research shows that adhesion traction force could be formed in the condition that some tangential micro-movement (called creep) taking place in the contact surface between wheel and iron rail. Also, adhesion traction force will increase with the adding of creep speed. However, adhesion traction force will decrease rapidly when creep speed is added too much. The relationship between adhesion traction force and creep speed is called adhesion property which can be represented as adhesion coefficient, as in.

$$\mu = F/P, \tag{1}$$

where P is weight of the locomotive, and F is the adhesion force. The relationship between P and F is shown in Fig. 1.

Also, we can find that the contact area forms a micro-elliptic region when locomotive stands on iron rail. The contact area in wheel is compressed in the forward direction. While in the opposite direction it is stretched. However, the contact area in iron rail is stretched in the forward direction and in the other direction it is compressed. The compressed contact area in wheel will recover when the wheel is rolling on the iron rail. This phenomenon indicates that wheel is not purely rolling but also along with some sliding. In other words, the circle line speed of a wheel is bigger than its traveling speed. The difference value between circle line speed and traveling speed of wheel is called “creep speed”. The adhesion characteristic between wheel and rail will be destroyed and wheel will slip when creep speed is added to a critical speed. We use the ratio of creep speed to traveling speed to represent the creep level, as in.

$$\gamma = \frac{\omega \cdot R - v_t}{v_t} = \frac{v_m - v_t}{v_t}, \tag{2}$$

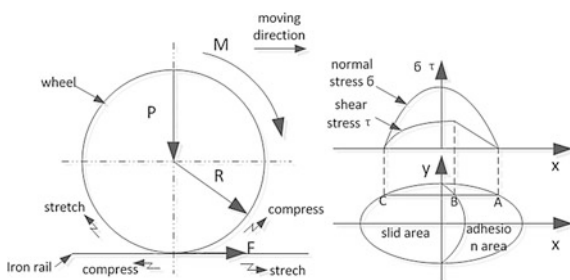
where ω is wheel angular speed, R is wheel diameter, v_m is wheel circle line speed, and v_t is traveling speed.

2.2 Adhesion Algorithm

From the above research, we can see that the bigger μ means the bigger available adhesion force between wheel and iron rail. To make locomotive in the biggest traction force, searching out the optimization adhesion coefficient point μ in real time is really an important job. As shown in Fig. 2, we can see that there is a biggest adhesion coefficient point at any time.

This paper proposes an optimization adhesion control point searching algorithm based on trend forecasting which is shown in Fig. 3. While $T_h, T_a, T_m, \omega_m, v_t,$

Fig. 1 Wheel rail adhesion diagram



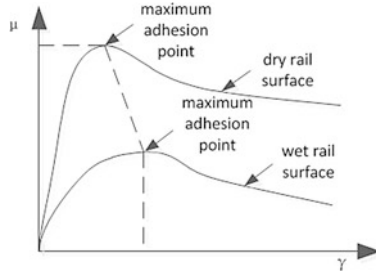


Fig. 2 The relationship of adhesion coefficient and creep ratio

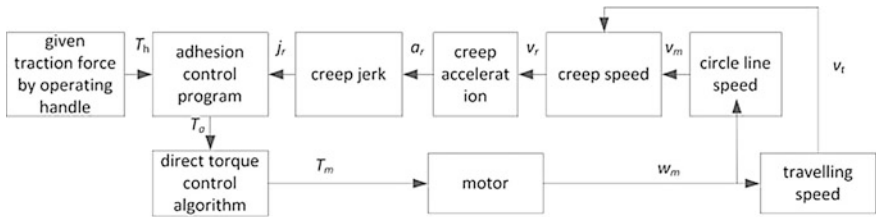


Fig. 3 Adhesion control point searching algorithm based on trend forecasting

v_m, v_r, a_r, j_r is the given traction force, traction force output from adhesion program, traction force output from control program, traction motor angular speed, locomotive traveling speed, wheel circle line speed, creep speed, creep acceleration, and jerk, respectively. v_m is obtained from the speed transducer. v_r is the average value of each axis' speed. v_r is the difference value of v_m and v_r . a_r is the differential of v_r while j_r is the differential of a_r .

The increase/decrease in amplitude ΔT of traction motor torque is constant in former adhesion programs which only take into account creep acceleration a_r . But in this paper, ΔT is time varying as we take into account creep jerk. When current adhesion control point is far away from the optimization adhesion control point, we can make ΔT much bigger. When current adhesion control point is close to the optimization adhesion control point, we can make ΔT smaller. That is to say, we can adjust the motor torque variation ΔT in real time. So, the accuracy will be much better. Also, because ΔT is much bigger when current adhesion control point is far away from the optimization adhesion control point, the search speed will be faster.

3 Adhesion Test

In order to validate the validity and efficiency of the adhesion control point searching strategy, this paper proposes to take some online traction performance test using HXD1 AC electric locomotive.

3.1 Test Condition

HXD1 locomotive was allocated to An-kang depot of Xi'an railway bureau. It is used to pull freight wagon on xi-kang railway line which starts north from xin-feng town and end south to an-kang city. It is across the Qin Mountain which is high and wide. The iron rail condition between each station is quite different because there are many tunnel, curve, ramp, and so on. There are 9.5 km rail line whose average ramp reach to 12% and 14.75 km rail line whose average ramp reaches to 11%.

Test was carried out in HXD1-1193 locomotive. The whole train is composed by an assisting locomotive and 4489 ton cargo. The test process is shown in Fig. 4. It was a sunny day when the test was carried out. In order to distinguish the traction force of each axis, first, we give a definition to locomotive axis which is shown in Fig. 5. In all test, we always use A-end as the operation end.

In order to compare the traction performance of HXD1 locomotive which starting from ramp or running across ramp in different adhesion program, this paper chooses some special section of xi-kang railway line to do the test, as in Table 1.

Fig. 4 HXD1-1193 locomotive



Fig. 5 Axis order of HXD1-1193 locomotive

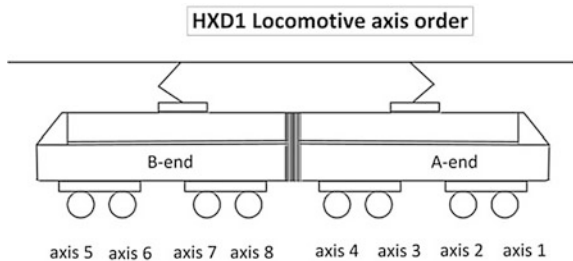


Table 1 Test scheme

Program	Section K44 ~ K52	Section K56 ~ K63	Ramp (‰)
Old adhesion program	Running across ramp	Starting from ramp	12
New adhesion program	Running across ramp	Starting from ramp	12

3.2 Test in Old Adhesion Program

3.2.1 Running Across Ramp

The running across ramp test in old adhesion program is carried out between Ma-xin station and Xi'an south station whose kilometer post is K44 and K52, respectively. Water was sprayed on the iron rail surface to simulate the rainy weather and sand was sprayed to enhance the adhesion effect. This paper chooses axis 1 of and axis 5 to analyze. The curves are shown in Figs. 6 and 7, respectively.

From Figs. 6 and 7, we can see that there are some different levels of slid phenomenon take place in axis 1 and 5. In order to reduce slid level, adhesion program plays a role to decrease the output torque of traction motor with a constant

Fig. 6 Curve comparison of old adhesion program in running across ramp (axis 1)

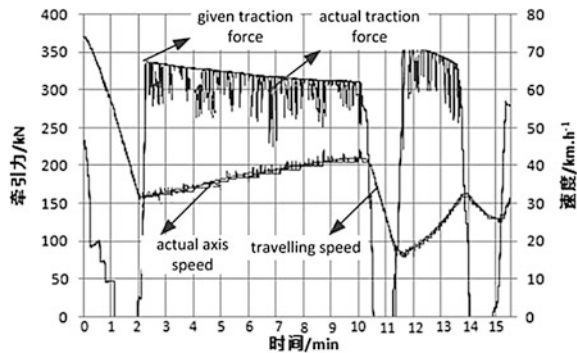
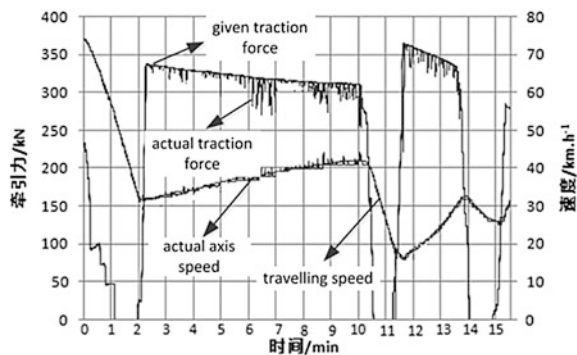


Fig. 7 Curve comparison of old adhesion program in running across ramp (axis 5)



value. The burr phenomenon of the curve in the figure is induced by actual output torque cannot track given output torque from operating handle.

Also, the traction performance of axis 5 is much better than the traction performance of axis 1. For axis 5, the traction force curve given by operating handle is quite consistent with the traction force curve produced by traction motor. The same thing also takes place in the speed curve.

3.2.2 Starting from Ramp

With the same method, starting from ramp test in old adhesion program is carried out between kilometer post K56 and K63. First, the train was stopped on the ramp whose value is 12‰. Then, water was sprayed on the iron rail surface to simulate the rainy weather and sand was sprayed to enhance the adhesion effect. The obtained curves are shown in Figs. 8 and 9, respectively.

From Figs. 8 and 9, we can see that there is obvious slid phenomenon taking place in axis 1 and axis 5 when HXD1 locomotive starts from ramp. The maximum slid takes place in axis 1.

Fig. 8 Curve comparison of old adhesion program in starting from ramp (axis 1)

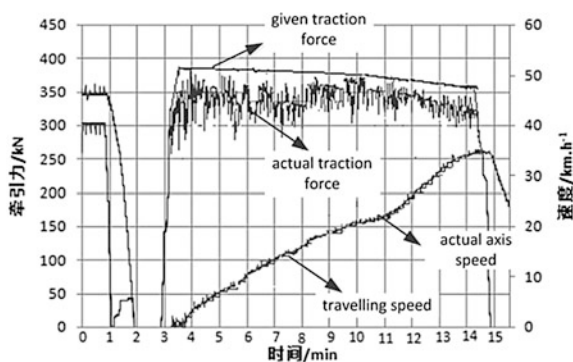
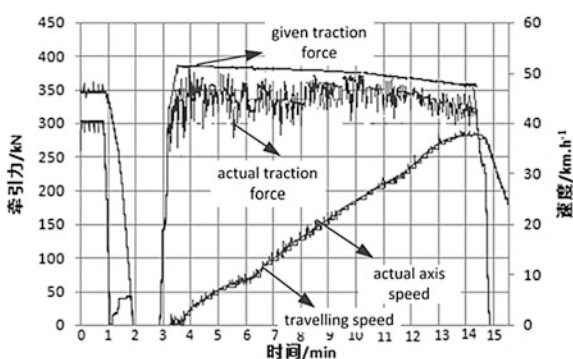


Fig. 9 Curve comparison of old adhesion program in starting from ramp (axis 5)



In order to reduce sliding, adhesion program plays a role to repeat a process which is decrease torque, increase torque, decrease torque. This is why there is some burr phenomenon of the curve in figures, especially when the traveling speed is among in 0–20 km/h. The same thing also takes place in the traveling speed curve. However, the traction force curve and traveling speed curve gradually tend to stabilize when the traveling speed is up to 20 km/h.

3.3 Test in New Adhesion Program

3.3.1 Running Across Ramp

With the same test method in part 3.2, the obtained curves for running across ramp in new adhesion program are shown in Figs. 10 and 11, respectively.

From Figs. 10 and 11, we can see that burr phenomenon in traction force curve and traveling speed curve is not quite obvious using the new adhesion program. That is to say, the new adhesion program plays a positive role in reducing slide phenomenon and in enhancing traction performance.

Fig. 10 Curve comparison of new adhesion program in running across ramp (axis 1)

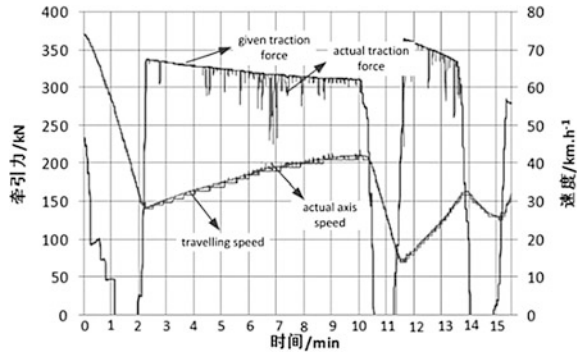
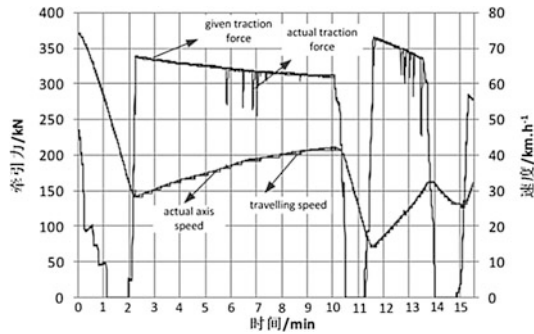


Fig. 11 Curve comparison of new adhesion program in running across ramp (axis 5)



3.3.2 Starting from Ramp

With the same test method in part 3.2, the obtained curves for starting from ramp in new adhesion program are shown in Figs. 12 and 13, respectively.

We can see that burr phenomenon in traction force curve and traveling speed curve is not quite obvious when the HXD1 locomotive is using the new adhesion program. Also, we can feel that the locomotive is much stabilized in the start process. That is because the increase/decrease in amplitude and speed of torque in the new adhesion program is much better.

3.4 Results Analysis

Taken together, the traction performance in the new adhesion program proposed in this paper is much better than the effect of old adhesion program. The traveling speed up to 20 km/h need 7 min and the actually used average traction force account for only 82% in given traction force by operating handle for the old adhesion program. But for the new adhesion program, the data is 5 min and 90%.

Fig. 12 Curve comparison of new adhesion program in starting from ramp (axis 1)

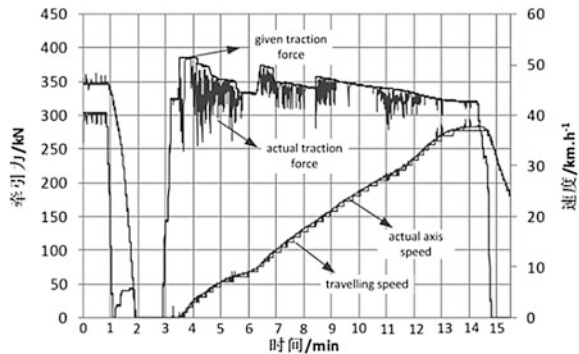
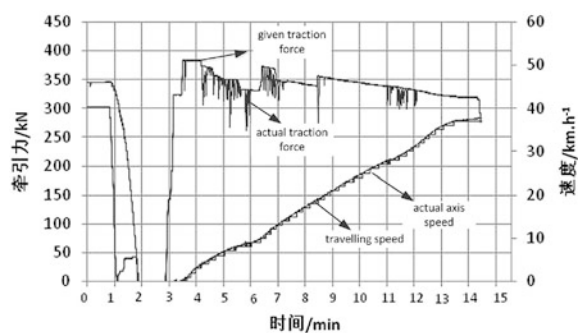


Fig. 13 Curve comparison of new adhesion program in starting from ramp (axis 5)



Further analysis showed that the old adhesion program is passive. When slid phenomenon takes place, adhesion program controls the traction torque of motor to reduce rapidly in a constant value which induces torque fluctuation. The new adhesion program is proposed in this paper considering creep jerk. So, we can know the variation trend of creep acceleration. We can use a time-varying value not constant value ΔT to adjust traction torque. The bigger the jerk, the bigger the ΔT . The smaller the jerk, the smaller the ΔT .

4 Conclusion and Outlook

This paper proposes a new adhesion coefficient searching strategy based on trend forecasting. By considering creep jerk, we can know the variation trend of creep acceleration. So we can give a time-varying value, not a constant value, to adjust output torque of motor. Also, the reliability and efficiency of new adhesion program are validated by a traction performance test in HXD1-1193 locomotive.

We can also see that the railway lines condition is really different. A same adhesion program may be not suitable for all railway lines. So, in order to play its best traction performance, it is necessary to carry out traction performance online test. Also, the next two questions are worth to study in the future. (1) Study the irregularity PSD of railway line. (2) Study the coupling dynamics between wheel and railway line.

References

1. Transport department integrated planning division (2013) Development statistical bulletin of transportation industry in 2013. China transportation report, 2014-05-13(02) (in Chinese)
2. Dabin C, Li L, Xuesong J (2010) Wheel-rail profiles matching design considering railway track parameters. *Chin J Mech Eng* 23(4):410–417 (in Chinese)
3. Qiyue L, Zhongrong Z (2003) Investigation of friction characteristic of steels under rolling-sliding contact. *Chin J Mech Eng* 39(7):56–58 (in Chinese)
4. Wang W, Guo J, Liu Q (2009) Effect of track structure parameters on rolling contact stresses of wheel-rail. *J Mech Eng* 45(5):39–44 (in Chinese)
5. Han C, Ning J, Chen Y (2014) HXD3 locomotive traction force and its three lines/three segments division analysis. *Chin Railways* 52(10):59–63 (in Chinese)
6. Ge X (2003) Traction test and its performance analysis of AC heavy train. Southwest Jiaotong University, Chengdu (in Chinese)
7. Zhang W, Zhou W, Chen L (2000) Experiment research on wheel/rail adhesion characteristics for high speed railway. *J China Railway Soc* 22(2):20–25 (in Chinese)
8. Yang Y (2003) Probe into adhesion control technology of AC drive locomotive. *Electr Drive for Locomot* 43(4):8–11 (in Chinese)
9. Li J, Ma J, Peng H (2002) Basic principle and methods of adhesion control of locomotive. *Electr Drive Locomot* 10(6):4–8 (in Chinese)

10. Ren Q, Huang J, Zhang S (2015) Adhesion control of electric motor coach based on fuzzy road identification. *Comput Simul* 32(3):173–177 (in Chinese)
11. Wang H, Xiao J (2003) Application of the wavelet analysis theory to optimized locomotive adhesion control. *J Chin Railway Soc* 25(5):32–38 (in Chinese)
12. Chen K, Lin P, Zhang S (2015) Research and implementation of the locomotive adhesion control multidisciplinary simulation platform. *Comput Simul* 32(1):211–214 (in Chinese)

Fault Identification for High-Speed Vehicle Suspension System Using Nonlinear Filtering

Xiaozhong Zhang, Xiukun Wei and Limin Jia

Abstract Suspension system is vital important to safe operation of the high-speed vehicle. Parameter estimation is a new way for fault identification of suspension system by monitoring parameter attenuation and sudden failure of key components. A lateral dynamic model is built in this paper, based on which a nonlinear filtering algorithm called Rao-Blackwellized Particle Filter (RBPF) is applied for parameter estimation. Furthermore, RBPF based on repeat-uniform-sampling strategy is proposed for avoiding impoverishment of parameter particles, and fault identification of sudden failure is realized. Simulation results show that algorithms proposed and studied in this paper are effective and reliable with high accuracy.

Keywords Fault identification · Parameter estimation · RBPF
Repeat-uniform-sampling strategy

1 Introduction

The suspension system is vital important to safe operation of high-speed vehicles. With development of high-speed railway in China, higher requirements for the safety and reliability of suspension system are raised because of the high-intensity running of high-speed vehicles. In order to evaluate the secure state of the suspension system, sequential fault identification of parameter attenuation and sudden failure is essential.

In recent years, parameter estimation research based on mathematical model and filtering algorithm develops rapidly. In the process of parameter estimation, high-speed vehicle suspension system dynamic model shows obvious nonlinear characteristics, which means that nonlinear filtering algorithm is required. Particle Filter (PF) is a typical nonlinear filter, and ideal parameter estimation results are

X. Zhang · X. Wei (✉) · L. Jia
State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University,
Haidian District, Beijing, People's Republic of China
e-mail: xkwei@bjtu.edu.cn

obtained when the dimension of the dynamic model is less than 5 [1]. However, if the dimension is much higher than 5, convergence speed and accuracy of PF drop rapidly. Based on the theory of state marginalization, RBPF is proposed in [2, 3], which separates state estimation and parameter estimation, and improves convergence speed and accuracy when the dimension of the dynamic model is high. Multiple parameter estimation based on RBPF is studied in [4], and the influence of rail noise to estimation result is studied in [5]. Joint estimation strategy based on multiple filters is studied in [6].

In this paper, a lateral dynamic model of the high-speed vehicle suspension system is built in Sect. 2, based on which parameter estimation of six key components using RBPF is introduced in Sect. 3. Fault identification under sudden failure condition using RBPF based on repeat-uniform-sampling strategy is introduced in Sect. 4, and some conclusions are summarized in Sect. 5.

2 System Description and Model Discretization

The lateral dynamic model is built by reference to the CRH2 trailer car, whose overview and rear view are sketched in Fig. 1. Parameters and meanings of the key components are shown in Table 1.

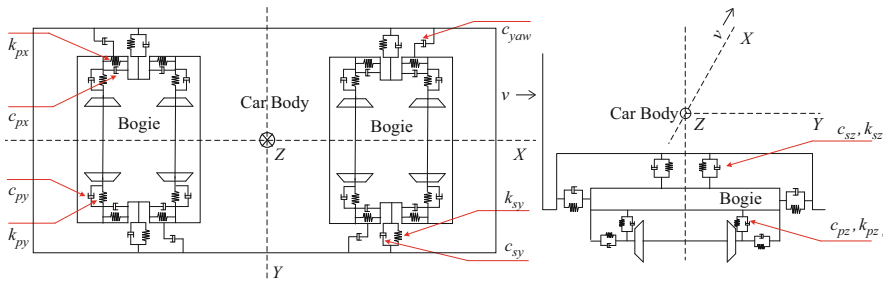


Fig. 1 Overview view and rear view of the lateral dynamic model

Table 1 Parameters and meanings of the key components

Meaning	Damping of secondary lateral damper	Stiffness of secondary lateral damper	Vertical damping of air spring	Vertical stiffness of air spring
Parameter	c_{sy}	k_{sy}	c_{sz}	k_{sz}
Meaning	Lateral damping of axle box guidance	Lateral stiffness of axle box guidance	Damping of primary damper	Vertical stiffness of axle box spring
Parameter	c_{py}	k_{py}	c_{pz}	k_{pz}
Meaning	Longitudinal damping of axle box guidance	Longitudinal stiffness of axle box guidance	Damping of anti-yaw damper	
Parameter	c_{px}	k_{px}	c_{yaw}	

Therefore, the continuous state space model can be written as

$$\dot{x} = \tilde{A}x + \tilde{B}u + \tilde{E}w \tag{1}$$

$$y = \tilde{C}x + \tilde{D}u + \tilde{F}v \tag{2}$$

where

$$x = [\dot{y}_c \ \dot{\phi}_c \ \dot{\rho}_c \ y_c \ \phi_c \ \rho_c \ \dot{y}_{b1} \ \dot{\phi}_{b1} \ \dot{\rho}_{b1} \ y_{b1} \ \phi_{b1} \ \rho_{b1} \ \dot{y}_{b2} \ \dot{\phi}_{b2} \ \dot{\rho}_{b2} \ y_{b2} \ \phi_{b2} \ \rho_{b2} \ \dot{y}_{w1} \ \dot{\phi}_{w1} \ \dot{y}_{w1} \ \phi_{w1} \ \dot{y}_{w2} \ \dot{\phi}_{w2} \ \dot{y}_{w2} \ \phi_{w2} \ \dot{y}_{w3} \ \dot{\phi}_{w3} \ \dot{y}_{w3} \ \phi_{w3} \ \dot{y}_{w4} \ \dot{\phi}_{w4} \ \dot{y}_{w4} \ \phi_{w4}]^T \tag{3}$$

$$y = [\ddot{y}_c \ \ddot{\phi}_c \ \ddot{\rho}_c \ \ddot{y}_{b1} \ \ddot{\phi}_{b1} \ \ddot{\rho}_{b1} \ \ddot{y}_{b2} \ \ddot{\phi}_{b2} \ \ddot{\rho}_{b2} \ \ddot{y}_{w1} \ \ddot{\phi}_{w1} \ \ddot{y}_{w2} \ \ddot{\phi}_{w2} \ \ddot{y}_{w3} \ \ddot{\phi}_{w3} \ \ddot{y}_{w4} \ \ddot{\phi}_{w4}]^T \tag{4}$$

x is the system state, y is the system observation, u is the system input, w is the process noise, and v is the observation noise. y_c, ϕ_c and ρ_c are lateral, yaw and roll movement of car body. $\dot{y}_c, \dot{\phi}_c$ and $\dot{\rho}_c$ are lateral, yaw and roll (angular) velocity of car body. $\ddot{y}_c, \ddot{\phi}_c$ and $\ddot{\rho}_c$ are lateral, yaw and roll (angular) acceleration of car body. The rest symbols belong to bogie and wheelset, which have similar meanings. The parameter estimation based on nonlinear filtering is carried out in discrete domain; therefore, the continuous state space model is rewritten as following discrete version:

$$x_{k+1} = Ax_k + Bu_k + Ew_k \tag{5}$$

$$y_k = Cx_k + Du_k + Fv_k \tag{6}$$

where $A = e^{\tilde{A}T}, B = \int_0^T A(t)\tilde{B} dt = \tilde{A}^{-1}(e^{\tilde{A}T} - I)\tilde{B}, E = \int_0^T A(t)\tilde{E} dt = \tilde{A}^{-1}(e^{\tilde{A}T} - I)\tilde{E}$, and $C = \tilde{C}, D = \tilde{D}, F = \tilde{F}$.

3 Parameter Estimation Using RBPf

3.1 State Marginalization

Based on Bayesian rules, posterior probability density function (pdf) of $\xi_k = [\theta_k \ x_k]^T$ given Z_k can be written as

$$p(\xi_k | Z_k) = p(x_k, \theta_k | Z_k) = p(x_k | \theta_k, Z_k)p(\theta_k | Z_k) \tag{7}$$

The system observation up to time k is Z_k . The system state and parameter at time k are x_k and θ_k . The estimation process of x_k and θ_k are partitioned and realized based on Kalman Filter (KF) and PF. Therefore, MMSE of ξ_k given Z_k can be written as

$$\hat{\xi}_k = \int \xi_k p(\xi_k | Z_k) d\xi_k = \int \left[\int (x_k, \theta_k) p(x_k | \theta_k, Z_k) dx_k \right] p(\theta_k | Z_k) d\theta_k \quad (8)$$

The idea of partitioning estimation process of x_k and θ_k is called state marginalization, which is also known as Rao-Blackwellization, and the combination of Rao-Blackwellization and PF is called RBPF.

3.2 Parameter Estimation Algorithm

The discrete state space model is partitioned into linear part and nonlinear part, which takes the following form:

$$\theta_k \sim p(\theta_k | \theta_{k-1}) \quad (9)$$

$$x_k = A(\theta_{k-1})x_{k-1} + B(\theta_{k-1})u_{k-1} + E(\theta_{k-1})w_{k-1} \quad (10)$$

$$y_k = C(\theta_k)x_k + D(\theta_k)u_k + F(\theta_k)w_k \quad (11)$$

The implementation of RBPF is sketched in Fig. 2.

(1) Initialization

Draw initial parameter samples $\theta_{1|0}(i)$ from initial pdf $p(\theta_0 | Z_0)$, and if there is no prior knowledge for $p(\theta_0 | Z_0)$, $\theta_{1|0}(i)$ can be uniformly sampled in $[\theta_{\min}, \theta_{\max}]$. Suppose that initial state is \hat{x}_0 , and the corresponding covariance matrix is P_0 . Draw initial state samples $x_{1|0}(i)$ from \hat{x}_0 and the corresponding covariance matrix $P_{1|0}(i)$ is calculated in the following way:

$$P_{1|0}(i) = E[(x_{1|0}(i) - \bar{x}_{1|0})(x_{1|0}(i) - \bar{x}_{1|0})^T], \bar{x}_{1|0} = \hat{x}_0 \quad (12)$$

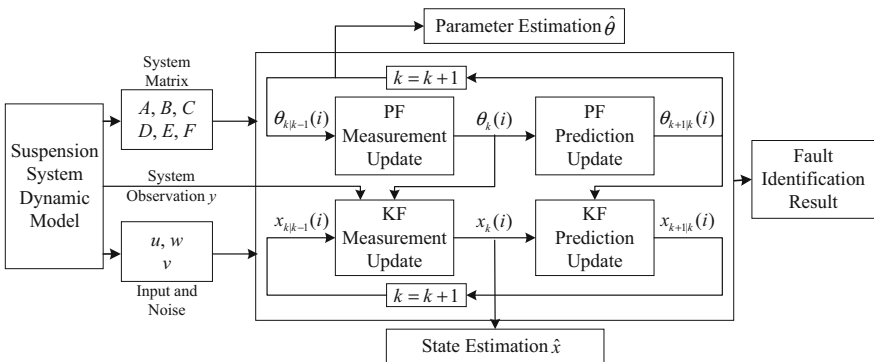


Fig. 2 Implementation of RBPF

(2) Repeat the following steps at time k ($k = 2, 3, \dots, T$)

(1) Importance weights update and normalization

$$y_{k|k-1}(i) = C(\theta_{k|k-1}(i))x_{k|k-1}(i) + D(\theta_{k|k-1}(i))u_k \quad (13)$$

$$R_k(i) = C(\theta_{k|k-1}(i))P_{k|k-1}(i)C^T(\theta_{k|k-1}(i)) + Q_v \quad (14)$$

$$\tilde{\alpha}_k(i) = p(y_k | Z_{k-1}, \theta_{k|k-1}(i)) \sim N(y_{k|k-1}(i), R_k(i)) \quad (15)$$

$$\alpha_k(i) = \frac{\tilde{\alpha}_k(i)}{\sum_{j=1}^N \tilde{\alpha}_k(j)} \quad (16)$$

where Q_v is covariance matrix of observation noise.

(2) Parameter estimation

$$\hat{\theta}_k = \sum_{i=1}^N \alpha_k(i) \theta_{k|k-1}(i) \quad (17)$$

(3) Resampling (measurement update of PF)

Resample the particles $\{\theta_{k|k-1}(i), x_{k|k-1}(i), P_{k|k-1}(i), \alpha_k(i) : i = 1, 2, \dots, N\}$ based on importance weight $\alpha_k(i)$, and new particles $\{\theta_k(j), x_{k|k-1}(j), P_{k|k-1}(j), N^{-1} : j = 1, 2, \dots, N\}$ are obtained, which satisfies [7]

$$Pr\{\theta_k(j) = \theta_{k|k-1}(i)\} = \alpha_k(i) \quad (18)$$

(4) Measurement update of KF

Calculate gain matrix $K_k(i)$, and estimation results of system state $x_k(i)$ and covariance matrix $P_k(i)$ are obtained.

$$K_k(i) = P_{k|k-1}(i)C^T(\theta_k(i))R_k^{-1}(i) \quad (19)$$

$$R_k(i) = C(\theta_k(i))P_{k|k-1}(i)C^T(\theta_k(i)) + Q_v \quad (20)$$

$$x_k(i) = x_{k|k-1}(i) + K_k(i)(y_k - C(\theta_k(i))x_{k|k-1}(i)) \quad (21)$$

$$P_k(i) = P_{k|k-1}(i) - K_k(i)C(\theta_k(i))P_{k|k-1}(i) \quad (22)$$

(5) Prediction update of PF

Prediction update of PF is realized based on the following evolution equation, proposed by Liu and West [8]:

$$\theta_{k|k+1}(i) = a\theta_k(i) + (1 - a)\bar{\theta}_k + W_k^\theta \tag{23}$$

where $a = (3\delta - 1)/2\delta$, and δ is the discount factor with recommended value range between 0.95 and 0.99. $\bar{\theta}_k$ is the Monte Carlo mean of parameter particles, and $W_k^\theta \sim N(0, h^2V_k)$ is the noise. V_k is the variance matrix of parameter particles and $h^2 = 1 - a^2$.

(6) Prediction update of KF

Prediction update results of system state $x_k(i)$ and covariance matrix $P_k(i)$ at time $k + 1$ are obtained in the following way:

$$x_{k+1|k}(i) = A(\theta_{k+1|k}(i))x_k(i) + B(\theta_{k+1|k}(i))u_k \tag{24}$$

$$P_{k+1|k}(i) = A(\theta_{k+1|k}(i))P_k^T(i)A^T(\theta_{k+1|k}(i)) + E(\theta_{k+1|k}(i))Q_wE^T(\theta_{k+1|k}(i)) \tag{25}$$

where Q_w is covariance matrix of process noise.

3.3 Simulation and Analysis

The parameter setting for simulation is shown in Table 2. The parameters estimated in this paper is $\theta = [C_{yaw} C_{sy} C_{sz} k_{sz} C_{pz} k_{pz}]$, which is shown in Table 1. Verification of algorithm effectiveness requires that the estimation result converges in $[0.9\theta_{real}, 1.1\theta_{real}]$, and θ_{real} is the real value of the parameter.

Each parameter is estimated separately using RBPF, and the result is sketched in Fig. 3. Meanwhile, estimation speed and accuracy of each parameter is shown in Table 3.

As shown in Fig. 3 and Table 3, estimation speed and accuracy using RBPF are satisfactory, with all estimated values converging in $[0.9\theta_{real}, 1.1\theta_{real}]$. Further, Error Ratio (ER) is introduced for evaluating the estimation process, which satisfies

Table 2 The parameter setting for simulation

Velocity	Sampling interval	Number of particles	Iteration	Initial particles sampling
$v = 80$ m/s	$T_s = 0.5$ ms	$N_s = 1000$	$T = 300$	$[0.5\theta_{real}, 2.5\theta_{real}]$

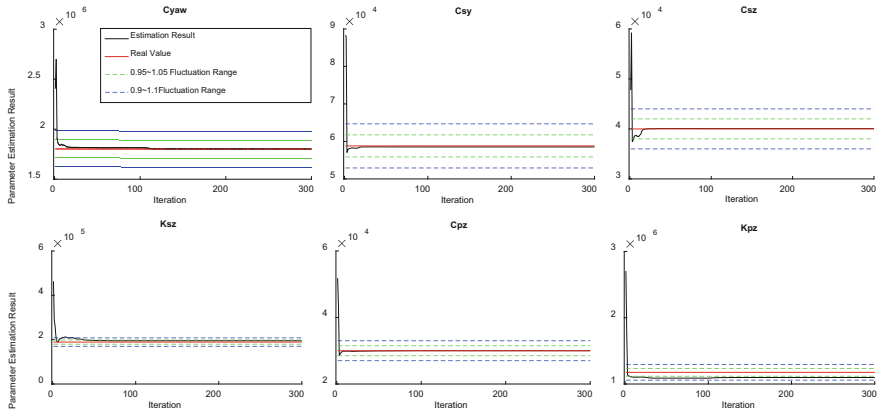


Fig. 3 Parameter estimation result using RBPF

Table 3 Estimation speed and accuracy of each parameter using RBPF

Parameter	Speed (convergence iteration)	Accuracy (convergence value)	Real value
C_{yaw}	115	1.802e + 6	1.800e + 6
C_{sy}	21	5.854e + 4	5.880e + 4
C_{sz}	34	4.009e + 4	4.000e + 4
K_{sz}	91	1.947e + 5	1.891e + 5
C_{pz}	53	2.998e + 4	3.000e + 4
K_{pz}	118	1.098e + 6	1.176e + 6

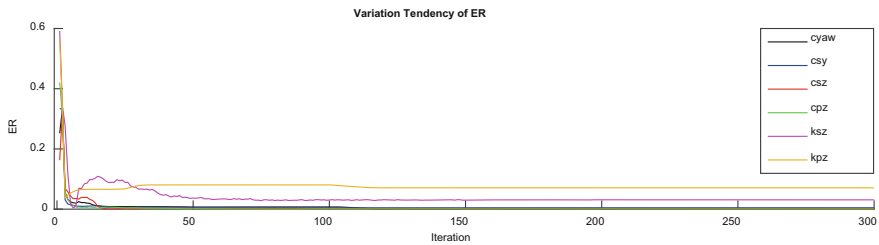


Fig. 4 Variation tendency of ER

$ER = |\hat{\theta}_k - \theta_{real}| / \hat{\theta}_k$, and $\hat{\theta}_k$ is the parameter estimation result at time k . The variation tendency of ER is sketched in Fig. 4.

As shown in Fig. 4, ER drops rapidly and converges under 0.1, which indicates that parameter estimation using RBPF converges quickly and accurately.

4 Fault Identification Using RBPf

RBPf based on repeat-uniform-sampling strategy [9] is proposed, based on which the impoverishment of parameter particles is solved, and sudden failure of key components is under continuous monitoring. The principle of the algorithm talked above is sketched in Fig. 5, and parameter setting for simulation is shown in Table 4.

The simulation is partitioned into four steps, specific setting of which is shown in Table 5. Sudden failure occurs at the 126th sample point of step 2, and parameter changes from θ_{normal} to $\theta_{fault} = 0.5\theta_{normal}$.

Suppose that sudden failure occurs at components which located on front side of vehicle, and fault identification results of normal and fault components are sketched in Fig. 6.

As shown in Fig. 6, fault identification result of each parameter converges in $[0.9\theta_{real}, 1.1\theta_{real}]$, which indicates that algorithm studied here is effective. Furthermore, the reaction time of the algorithm to sudden failure can be shortened by cutting down the iteration of each step in case that the convergence speed is fast enough.

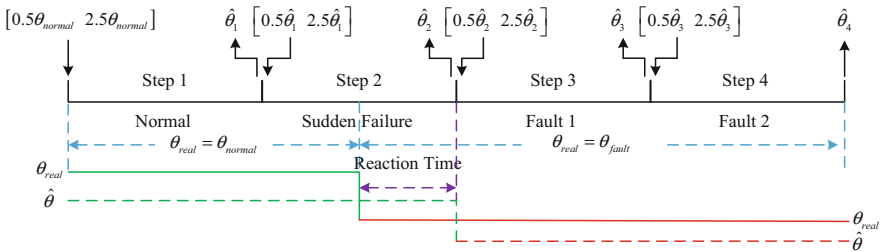


Fig. 5 Principle of RBPf based on repeat-uniform-sampling strategy

Table 4 The parameter setting for simulation

Velocity	Sampling interval	Number of particles	Iteration	Normal value	Fault value
$v = 80 \text{ m/s}$	$T_s = 0.5 \text{ ms}$	$N_s = 1000$	$T = 250$	θ_{normal}	$0.5\theta_{normal}$

Table 5 Specific setting for each step of the simulation

Step	Step 1	Step 2	Step 3	Step 4
Condition	Normal	Sudden failure	Fault 1	Fault 2
Real value θ_{real}	θ_{normal}	$\theta_{normal} \rightarrow \theta_{fault}$	θ_{fault}	θ_{fault}
Sample reference $\tilde{\theta}$	θ_{normal}	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$
Identification result $\hat{\theta}$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$

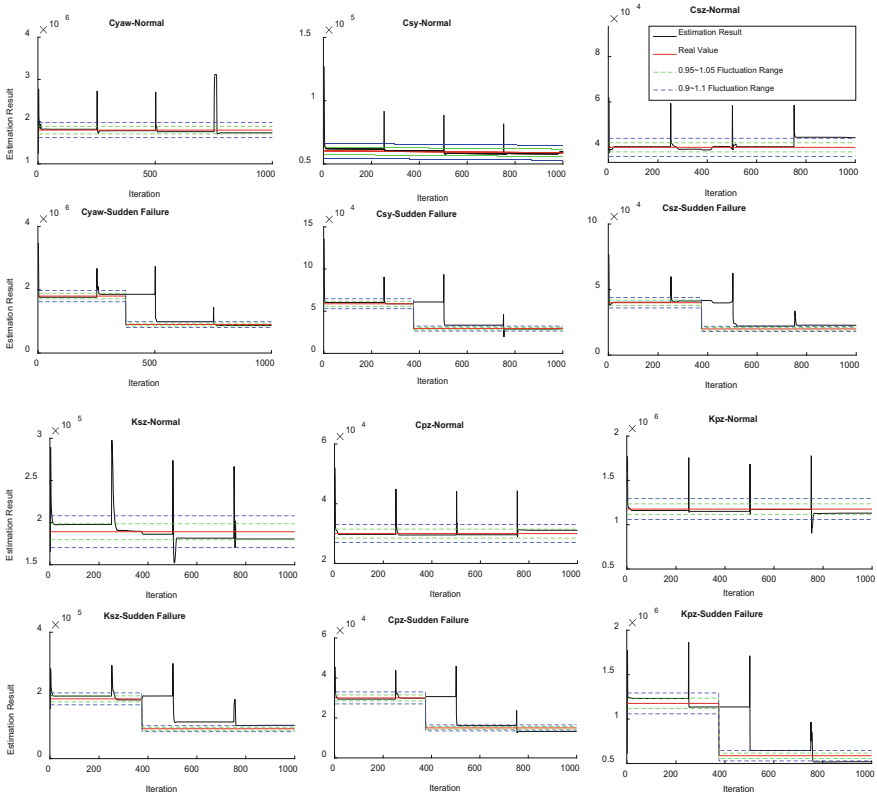


Fig. 6 Fault identification results using RBPF based on repeat-uniform-sampling strategy

5 Conclusions

In this paper, lateral dynamic model of the high-speed vehicle suspension system is built in SIMULINK, and nonlinear filtering algorithm called RBPF is applied for parameter estimation of key components. Moreover, RBPF based on repeat-uniform-sampling strategy is proposed for parameter estimation under sudden failure condition, which implements fault identification with high convergence speed and accuracy. Further extension of this research lies in modeling method improvement and real data verification, which will be studied in the future.

References

1. Gustafsson F, Hriljac P (2003) Particle filters for system identification with application to chaos prediction. In: Proceedings of SYSID03, 2003

2. Li P, Goodall R, Kadiramanathan V (2004) Estimation of parameters in a linear state space model using a Rao-Blackwellised particle filter. *IEEE Proc—Control Theor and Appl* 151 (6):727–738
3. Li P, Goodall R, Weston P et al (2007) Estimation of railway vehicle suspension parameters for condition monitoring. *Control Eng Pract* 15(1):43–55
4. Ding JM, Lin JH, Wang H et al (2014) Performance parameter estimation for high speed train using RBPF. *J Transp Eng* 3:52–57 (in Chinese)
5. Xv BW (2015) Parameter estimation for high speed vehicle suspension system using RBPF. Southwest Jiaotong University (in Chinese)
6. Geng SY (2015) Parameter estimation for high speed vehicle suspension system based on multiple filters. Southwest Jiaotong University (in Chinese)
7. Kitagawa G (1996) Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J Comput Gr Stat* 5(1):1–25
8. Liu J, West M (2001) Combined parameter and state estimation in simulation-based filtering. In: *Sequential Monte Carlo methods in practice*. Springer, New York, pp 197–223
9. Fang Y, Chen L, Li XB et al (2013) Improved parameter estimation method for secondary suspension system based on repeat-uniform-sampling strategy. *Chin Railway Sci* 34(3):72–78 (in Chinese)

Research on Gearbox Fault Isolation Based on VPMCD

Xiukun Wei and Dong Yan

Abstract As a power transmission device, gearbox is an indispensable component of the urban rail vehicle, and it is significant to determine the fault type of the gearbox and to provide the basis for the maintenance and maintenance plan of gearbox. In this paper, traditional VPMCD method is improved in the selection of the prediction variables and the parameter estimation, and based on the error between the estimated value and the actual value, the fault type of the gearbox can be accurately judged. The applicability of VPMCD is analyzed by an example, and compared with other methods, the superiority of the optimized VPMCD is verified.

Keywords Gearbox · Feature extraction · Fault isolation · VPMCD

1 Introduction

Gearbox is used to change the running speed of the railway vehicle; it includes shaft, gear, rolling bearing, and box. Because of the complex structure, severe working conditions, gears, and bearings are susceptible to damage. Therefore, it is significant to determine the fault type of the gearbox, making the management measures and to protect the safety of people.

Fault isolation includes traditional pattern recognition and modern pattern recognition [1]. Neural network and support vector machine have been widely used in fault identification. There is a direct relationship between the precision and the parameters of the neural network algorithm, and it is easy to fall into local optimum [2]; the isolation accuracy of support vector machine (SVM) is related to kernel function [3], and there is no unified parameters and kernel function selection basis. The above method ignores the internal relationship between the fault eigenvalues. Literature [4] presents a new isolation method—Variable predictive model based

X. Wei (✉) · D. Yan
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing, China
e-mail: xkwei@bjtu.edu.cn

class discriminate (VPMCD); it makes full use of the internal relationship between the characteristic values, and mathematical models are established for the different characteristics of different faults in an objective way [5]. However, traditional VPMCD has some defects in variable selection, parameter estimation, and model type determination. In this paper, select the optimal variable according to the Laplace value, and robust regression method is used instead of the least square method to reduce the effects of outliers in parameter estimation. At last, the optimal model is determined according to the minimum error value by selecting the type and order of the model.

The rest of this paper is organized as follows. Section 2 introduces the mechanism of VPMCD, the calculation steps of LS score, and the advantages of robust regression. In Section 3, the applicability of VPMCD is verified by a case, and the accuracy of VPMCD algorithm is compared with other methods. The conclusions are drawn in Sect. 4.

2 Variable Prediction Model

2.1 Mechanism Analysis of Variable Prediction Model

VPMCD is a method based on the relationship between the characteristics in different states, related research point out that the number of features between 2 and 6 is better [6].

The principle of this method is shown in Fig. 1. In this figure, training starts with a given training data set $N[n \times p:c]$, of which n notes the samples of the group, p notes the number of features, and c is the number of fault classes. Matrix N is separated into submatrices for each class of observations, $G_g[n_g \times p : c]$, where n_g is the number of samples in the training set belonging to class g , $n_1 + n_2 + \dots + n_g + \dots + n_c = n$ ($g = 1, 2, \dots, c$). After establishing the equation, the parameters such as b_j , b_{jj} , and b_{jk} are estimated by least square method for every training samples. The training process is over, and the function VPM of the features for each class is obtained. The number of training function is $c \times p$. As for testing progress, the same form testing data set should be prepared like training data set, and put them into the model which has been trained. Then, the prediction data for each feature in each class is obtained by calculating the function. The prediction value is x'_i , and the real value is x_i , ($i = 1, 2, \dots, p$). At last, calculate the sum errors S_g for each class, and the class with the smallest error is the fault type of testing data.

There are four VPM model types [4]:

Linear:

$$X_i = b_0 + \sum_{j=1}^r b_j X_j \quad (1)$$

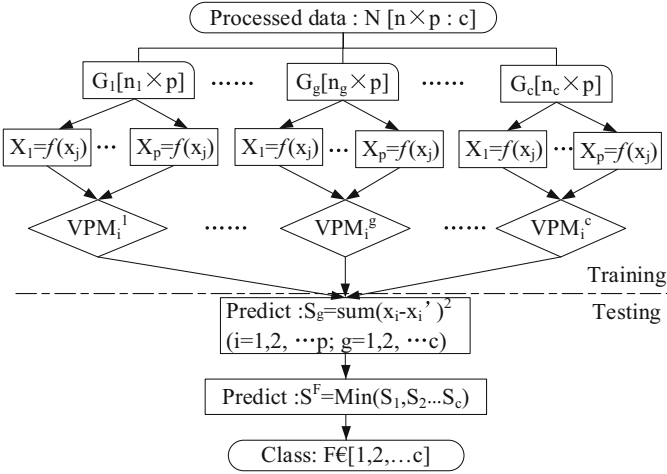


Fig. 1 VPMCD processing

Linear interaction:

$$A \tag{2}$$

Quadratic:

$$X_i = b_0 + \sum_{j=1}^r b_j X_j + \sum_{j=1}^r b_{jj} X_j^2 + \sum_{j=1}^r \sum_{k=j+1}^r b_{jk} X_j X_k \tag{3}$$

Quadratic interaction:

$$X_i = b_0 + \sum_{j=1}^r b_j X_j + \sum_{j=1}^r b_{jj} X_j^2 \tag{4}$$

2.2 Laplace Score

Select the better characteristics according to Laplace score (LS). The Laplace score transforms the complex high-dimensional feature space into a low-dimensional feature space according to Laplace’s eigenvalue mapping and local preserving projection, and determines the importance of eigenvalues by local retention [7].

Let L_r be the Laplace score of the r th characteristics, f_{ri} is the i th eigenvalue of the r th sample, $r = 1, 2, \dots, m$. The algorithm can be represented as follows: Let L_r be the adjacent map of r th sample points. The i th sample corresponds to x_i . When x_i

and x_j are close enough, they are the neighbors of each other. Then, there is a side connection; otherwise, there is no edge to connect. When the label of the node is known, an edge can be connected between two nodes of the same label:

$$S_{ij} = \begin{cases} e^{-\frac{\|x_i - x_j\|^2}{t}} & i \text{ is connected with } j \\ 0 & i \text{ is not connected with } j \end{cases} \quad (5)$$

The r th eigenvalue is defined as

$$f_r = [f_{r1}, f_{r2}, \dots, f_{rm}]^T \quad D = \text{diag}(SI) \quad (6)$$

$$I = [1, 1, \dots, 1]^T \quad L = D - S \quad (7)$$

The matrix L is called the Laplacian matrix of graph G . The deaveraging of each feature is obtained:

$$\tilde{f}_r = \tilde{f}_r - \frac{f_r^T DI}{I^T DI} I \quad (8)$$

The Laplace score for the r th eigenvalue is

$$L_r = \frac{\sum_{ij} (f_{ri} - f_{rj})^2 S_{ij}}{\text{Var}(f_r)} = \frac{\tilde{f}_r^T L \tilde{f}_r}{\tilde{f}_r^T D \tilde{f}_r} \quad (9)$$

2.3 Robust Regression

The robust regression is an iterative weighted least squares estimation regression, the weight w_{ij} of the points according to the regression error values. The larger weight is given to the smaller error point, and the smaller weight is given to the point where the error is large. The iterative correction weight is repeated, and the influence of the ‘‘abnormal value’’ is reduced [8].

For the general regression model,

$$Y_j = \sum_{j=1}^p x_{ij} \beta_j + e_i \quad i = 1, 2, \dots, n \quad (10)$$

The β_j is regression coefficients, and e_i is the error value, where the mean is zero independent distribution. Actual and estimated errors and the optimization goals of robust estimates are shown in (11) and (12):

$$e = \left[Y_j - \sum_{j=1}^p x_{ij}\beta_j \right]^2 \tag{11}$$

$$\sum_{i=1}^n w_i \left[Y_j - \sum_{j=1}^p x_{ij}\beta_j \right]^2 = \min \tag{12}$$

3 Case Analysis

3.1 Data Sources

To validate the efficiency of this method, data from the bearing experimental platform of the NSF/UCR intelligent maintenance center in the United States are used. The parameters of the bearing are shown in Table 1, and data sampling rate is 20 kHz. And in gear data from the QPZZ-II rotary mechanical vibration analysis and fault diagnosis test platform system, large gear teeth number is 75, pinion teeth number is 55, and the gear modulus is 2. There are six kinds of fault types: normal, large gear broken, large gear pitting, pinion wear, large gear pitting and pinion wear, and large gear cutter and pinion wear.

3.2 Bearing Fault Isolation

40 samples of each state were extracted, which include 30 samples of training data and 10 samples of the test data. The LS scores of each characteristic are calculated and arranged them from small to large; the result is shown in Fig. 2. The numbers of characteristic from 1 to 10 are shown in Table 2.

The prediction accuracy of the model for calculating 2–6 variables is shown in Fig. 3, when the number of variables is 5; the prediction accuracy is the highest. Choose the least five characteristics of LS score as the variables of the model. The model type and model order are shown in Table 3. Taking the outer race fault data as example, the results are shown in Table 4. The error value of the outer race fault is smaller than others, indicating that the input data meet the outer race fault model, consistent with the actual situation.

Table 1 Parameters of the bearing test platform

Name	Parameter	Name	Parameter
Number of rollers	16	Roller diameter	0.331 in.
Pitch diameter	2.815 in.	Contact angle	15.17°

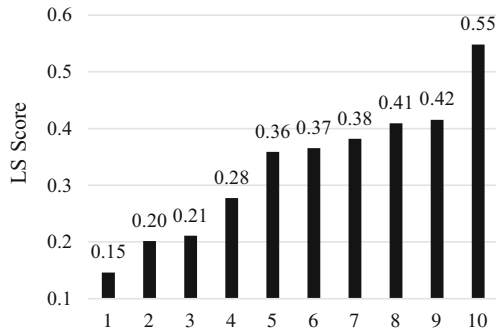


Fig. 2 Laplace score of characteristic

Table 2 Mathematical model and order number of model

No.	1	2	3	4	5
Characteristic	Wave factor	Kurtosis value	Power spectrum centroid index	Mean frequency	Mean square spectrum
No.	6	7	8	9	10
Characteristic	Shannon entropy	Margin index	Pulse factor	Frequency variance	Peak factor

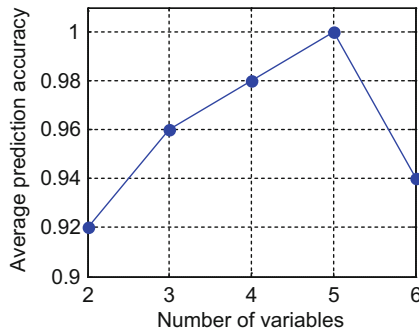


Fig. 3 Prediction accuracy

VPMCD prediction accuracy is compared with the SVM and backpropagation (BP) neural network algorithm in the condition of enough training samples and insufficient training samples. The results are shown in Figs. 4 and 5. It can be seen from Fig. 4 that the accuracy of the VPMCD algorithm is 100% in each state. It is demonstrated that VPMCD is effective and more accurate. In view of the situation of lack of fault training data, the four pattern recognition methods are also compared and analyzed, and the results are the same.

Table 3 Mathematical model and order number of model

Predicted variable	Bearing condition							
	Normal		Outer race fault		Inner race fault		Mixed fault	
	Model type	Model order	Model type	Model order	Model type	Model order	Model type	Model order
X_1	LI	2	LI	2	LI	4	LI	4
X_2	QI	4	LI	4	QI	4	QI	4
X_3	LI	4	LI	4	LI	4	LI	4
X_4	L	4	LI	4	L	4	LI	4
X_5	LI	4	QI	4	LI	4	QI	4

Table 4 Fault isolation of outer race fault

Bearing condition	Normal	Outer race fault	Inner race fault	Mixed fault
Error value	1531.38	44.37	0.027	65.71
Identification results	X	X	✓	X

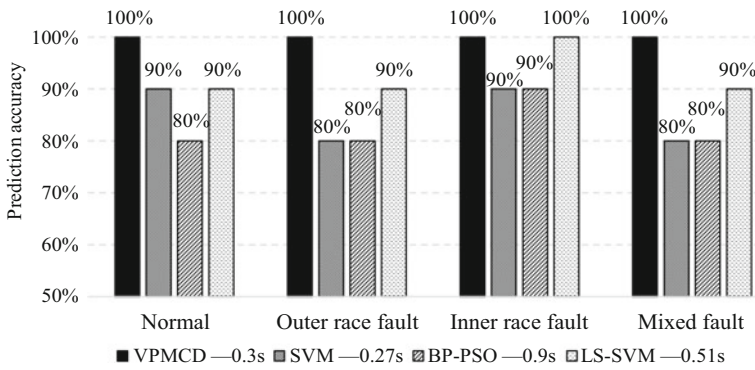


Fig. 4 Comparison between VPMCD and other methods (30 samples for each type)

3.3 Gear Fault Isolation

In this section, take the gear as example to verify the effectiveness of VPMCD. The number of training samples is 100, and there are 38 test samples. As with the training method of the bearing, the LS score of the characteristics is calculated and the optimal model is selected. VPMCD prediction accuracy is compared with the SVM and BP neural network algorithm in the condition of enough training samples and insufficient training samples. The results are shown in Figs. 6 and 7, respectively. The accuracy of VPMCD algorithm is 95 and 97% in normal and wear state, respectively, in Fig. 6, and the isolation accuracy of other states is 100%. The prediction accuracies of BP and LS-SVM are not satisfactory.

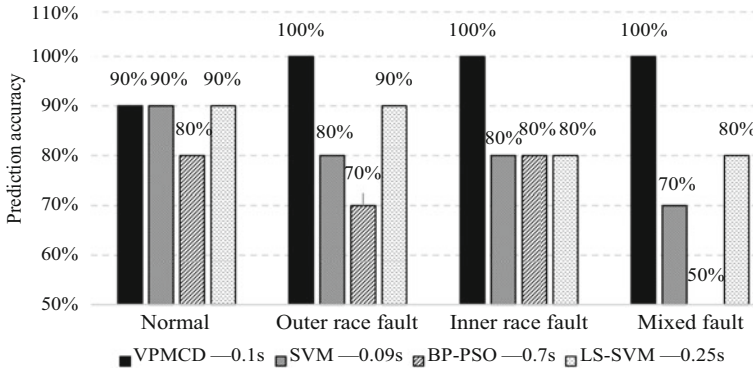


Fig. 5 Comparison between VPMCD and other methods (12 samples for each type)

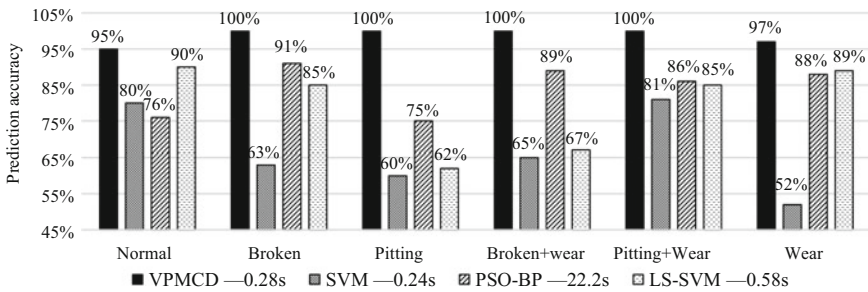


Fig. 6 Comparison between VPMCD and other methods (100 samples for each fault)

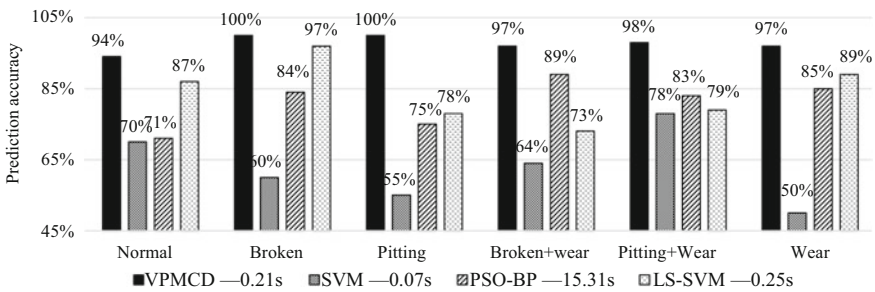


Fig. 7 Comparison between VPMCD and other methods (50 samples for each fault)

Reduce the number of training data of each state to 50, and compare the fault isolation effect of the four kinds of methods. It can be seen from Fig. 7 that all states of VPMCD are more than 94%. While the prediction accuracy of other algorithms has greatly decreased, the result is not ideal.

According to the above research, it can be seen that the training accuracy of VPMCD is higher than that of BP, SVM, and LS-SVM in the case of sufficient or missing training samples. Thus, VPMCD can distinguish single fault and mixed fault accurately.

4 Conclusions

This paper introduces the mechanism of VPMCD and the realization process, and the optimal predictor is selected by using the Laplace score method, which reduces the influence of the characteristics on the accuracy. The robust regression method is used to estimate the parameters instead of the least squares method, which reduces the influence of the “outliers”. The VPMCD algorithm is applied to the actual fault to verify the validity of the VPMCD. The accuracy, efficiency, and applicability of the VPMCD algorithm are proved by comparing with other algorithms.

Acknowledgements This work is also partly supported by State Key Lab of Rail Traffic Control & Safety (Contract No. RCS2016ZT006). This work is also partly supported by National Key R&D Program of China (Contract No. 2017YFB1201201).

References

1. Sun J (2008) Modern pattern recognition. Higher Education Press, Beijing
2. Saimurugan M, Ramachandran KI et al (2011) Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine. *Expert Syst Appl* 38(4):3819–3826
3. Wang CC (2010) Applications of fault diagnosis in rotating machinery by using time series analysis with neural network. *Expert Syst Appl* 37(2):1696–1702
4. Raghuraj R, Lakshminarayanan S (2009) Variable predictive models—a new multivariate classification approach for pattern recognition applications. *Pattern Recogn* 42(1):7–16
5. Raghuraj R, Lakshminarayanan S (2008) Variable predictive model based classification algorithm for effective separation of protein structural classes. *Comput Biol Chem* 32(4):302–306
6. Xu Z, Xuan J (2009) Application of a modified fuzzy ARTMAP with feature-weight learning for the fault diagnosis of bearing. *Expert Syst Appl* 36(6):9961–9968
7. He X, Cai D (2005) Laplacian score for feature selection. *Adv Neural Inf Process Syst* (18):507–514 (Cambridge)
8. Ma JJ, Li SW (2009) Retrieving method of differential optical absorption spectroscopy based on M-estimator robust regression. *Acta Photonica Sin* 38(8):2035–2039

A Feature Extraction Method of Rolling Bearing Fault Signal Based on the Singular Spectrum Analysis and Linear Autoregressive Model

Kun Xu, Gui Wang and Zongyi Xing

Abstract A feature extraction method of rolling bearing fault signal based on the singular spectrum analysis (SSA) and linear autoregressive (AR) model is proposed. The SSA is used to achieve the noise reduction, which has three steps: decompose original signal into multiple components, remove the components which have smaller contribution, and reconstruct the signal. Then, the reconstructed signal is modeled by the linear AR model, and the coefficients of the model are extracted as the characteristics of the signal. Finally, the proposed method is verified by using the experimental data of Case Western Reserve Lab.

Keywords Rolling bearing · SSA · AR model · Feature extraction

1 Introduction

The vibration signals generated by the rolling bearings of the rail trains are always nonstationary, usually due to the clearance between the bearing outer race and the housing or the sliding of the rolling elements [1]. And the non-stationarity of the bearing may also be due to the impact between the fault and the non-faulty component [2]. For these reasons, the traditional nonparametric technique has some limitations. For example, the Fourier transform (FFT) is an efficient algorithm for transforming a signal from time domain to frequency domain, but it is not suitable for an unstable signal, and it takes a longer time interval to form a good resolution spectrum. In this study, we proposed a feature extraction method based on the singular spectrum analysis and linear AR model, in which the SSA is used to remove the noise of the vibration signal, and then the unstable sub-signals are differentiated to realize the smoothness of the signal. Finally, the feature extraction of vibration signal is realized according to linear AR model.

K. Xu · G. Wang · Z. Xing (✉)

School of Automation, Nanjing University of Science and Technology,
Nanjing 210094, China
e-mail: xingzongyi@njust.cn

2 The Basic Theory

2.1 Singular Spectrum Analysis

SSA is a nonparametric model analysis method based on multivariate statistical principle. It performs the principal component analysis on the lagged component of the time series, which decomposes the original signal into multiple independent components (PCs), and reconstructs the original time series through multiple PCs. The main purpose of SSA is to reveal the trend of the signal, and it is also used as a noise removal by removing the low singular value components [3]. Only the independent principal component portion is selected for the construction phase, which will result in a significant improvement in the suitability of the linear autoregressive model (LTIVAR). In this study, SSA was used for decomposing the segmented signals, after which they were reconstructed using a number of the principal components. Thus, the procedure has two stages: decomposition and reconstruction [4].

2.1.1 Decomposition

In the decomposition stage, a sub-signal $\mathbf{X} = \{x_1, \dots, x_N\}$ of length N is mapped onto a window of length L , which can construct a trajectory matrix \mathbf{T}_X ($L \times K$), where $K = N - L + 1$. And the trajectory matrix is shown in Eq. (1).

$$T_{ij} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_N \end{pmatrix} \quad (1)$$

After the above step, the trajectory matrix \mathbf{T}_X is decomposed by singular values (SVD), and the trajectory matrix T_i after the decomposition is obtained is shown in Eq. (2):

$$T_i = U_i \times \sqrt{\lambda_i} \times V_i^T \quad (2)$$

where U_i ($K \times L$) is the orthogonal matrix of the; V_i ($L \times L$) is the diagonal orthogonal matrix; $\sqrt{\lambda_i}$ is the singular value; and $1 < i < L$. And the trajectory matrix \mathbf{T}_X has L singular values and satisfies the following conditions:

$$\sqrt{\lambda_1} > \sqrt{\lambda_2}, \dots, \sqrt{\lambda_L} \quad (3)$$

So the trajectory matrix T_X can be expressed in the lower form:

$$T_X = T_1 + T_2 + \dots + T_d = \sum_{i=1}^d U_i \sqrt{\lambda_i} V_i^T \tag{4}$$

where $d = \max(i: \sqrt{\lambda_i} > 0)$. In vibration analysis, one can impose on the width of the window a lower limit according to the desired minimum frequency resolution. Thus, if one wants to observe phenomena, which occur at frequencies at least equal to f , the minimum width of the window must be

$$L \geq \frac{f_s}{f} \tag{5}$$

where f_s is the sampling frequency of the vibration signal.

2.1.2 Reconstruction

The second stage of SSA involves reconstruction of the trajectory matrix T_X , which may be performed either by means of individual analysis or grouping analysis. The grouping is to index the $\{1, \dots, d\}$ is divided into m disjoint subsets $I = \{I_1, \dots, I_m\}$, so the trajectory matrix T_X can be expressed as follows:

$$T_X = \overbrace{T_1 + \dots + T_L}^{\text{SVD}} = \overbrace{T_{I_1} + \dots + T_{I_m}}^{\text{Grouping}} \tag{6}$$

After the grouping analysis, we proposed the diagonal mean strategy transformed the grouped trajectory matrix T_X into a new time series $\tilde{T}_{i,j}$:

$$\begin{aligned} \tilde{T}_{i,j} &= \frac{1}{s-1} \sum_{j=1}^{s-1} x_{j,s,j} & 2 \leq s \leq L \\ \tilde{T}_{i,j} &= \frac{1}{s-1} \sum_{j=1}^L x_{j,s,j} & L+1 \leq s \leq K+1 \\ \tilde{T}_{i,j} &= \frac{1}{N-s+2} \sum_{j=s}^{N-s+2} x_{j,s,j} & K+2 \leq s \leq N+1 \end{aligned} \tag{7}$$

The new reconstructed time series $\tilde{T}_{i,j}$ will be used for the LTIVAR.

2.2 Linear Autoregressive Model

Suppose that a time series $\{x_r\}$ satisfies the conditions of stationary, normality, and zero mean, then the time series $\{x_r\}$ can be represented by a linear autoregressive model as follows:

$$xr(m) = a_1 \bullet xr(m-1) + \dots + a_p \bullet xr(m-p) + \varepsilon(m) \quad (8)$$

where $xr(m)$ is the predicted signal value at time m which is linearly related to p previous values; p is the order of the model; a_i ($i = 1, 2, \dots, p$) are model weighting coefficients; and $\varepsilon(m)$ is the error term, which represents the difference between the actual and linearly predicted values.

The performance of the model in representing the experimental signal data is measured by calculating normalized loss function based on the normalized mean square error (NMSE), as given by the equation below:

$$\text{Gof}_{\text{NMSE}} = \left(1 - \frac{\sqrt{\sum_{i=1}^n (xr_p(i) - xr_m(i))^2}}{\sum_{i=1}^n xr_p \quad \bar{xr}_m} \right) \times 100\% \quad (9)$$

where xr_p is the forecasting sub-signal, xr_m is the real-time measured sub-signal and is the mean value of real-time sub-signal, and n is the number of data points.

2.2.1 Data Preprocessing

When the time series is nonstationary, differential can be used to make the time series stationary [5]. In this study, the stationarity of each reconstructed sub-signal is determined by the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) [6] test. After passed the KPSS test, the differential technique is only used to stabilize the signals which are not stable, and the difference method can be expressed as follows:

$$xr^{\text{diff}}(m) = xr(m) - xr(m - 1) \quad m = 2, 3, \dots, n \quad (10)$$

where $xr^{\text{diff}}(m)$ is the new stationary sub-signal. In this study, the primary difference can make the sub-signal stable.

2.2.2 Model Order Selection

The determination of the optimal model order is very important, and the incorrect model order may lead to over-fitting and under-fitting, which results in an inaccurate model prediction and leads to inaccurate or erroneous fault diagnosis [7, 8]. In order to find the order of reasonable AR model, the simplest way is to find the model order when the variance of the model prediction error is minimum. In this paper, the BIC criterion is used to determine the optimal model order. The BIC criterion can be defined by the following formula:

$$BIC(n) = \ln(\sigma^2) + n \times \ln(N)/N \tag{11}$$

where σ^2 is the variance of the residual, n is the order of the model, and N is the number of data points. According to this criterion, the model order corresponding to the BIC minimum is selected as the optimal order.

2.2.3 Model Parameter Estimation

In this study, the recursive least squares estimation is used to estimate the parameters of AR model. The least squares estimate of the parameter matrix is

$$\hat{a}_{N+1} = \hat{a}_N + P_N a_{p+N+1} (I + a_{p+N+1}^T P_N a_{p+N+1})^{-1} (x_{N+1} - a_{p+N+1}^T a_N) \tag{12}$$

where $P_N = (X_N^T \bullet X_N)^{-1}$, $\hat{a}_0 = 0$, $P_0 = \mu^2 I$, take $10^5 \leq \mu^2 \leq 10^{10}$ (as large as possible). Besides, the error term $\varepsilon(m)$ can be obtained as follows:

$$\varepsilon(m) = \frac{1}{N} \sum_{t=p+1}^N (xr(t) - \sum_{i=1}^p a_i xr(t-i))^2 \tag{13}$$

2.2.4 Dealing with Model Orders

There might be some cases where the model orders for different signals categories are not the same as some fault types may require a different model order. As these model orders will represent the lengths of the feature vectors formed for the diagnosis stage, it means the lengths of the feature vectors will not be the same. In this study, we proposed an equalization method called zero padding as given by the equation below:

$$\begin{aligned} & [a_0 \quad a_1 \quad a_2 \quad \cdots \quad a_l \quad \overbrace{0 \ 0 \ \cdots \ 0}^{p_{\max}-l}] \\ & [a_0 \quad a_1 \quad a_2 \quad \cdots \quad a_{p_{\max}}] \end{aligned} \tag{14}$$

where p_{\max} is the maximum order of the optimal order, and l is one of the optimal orders ($l < p_{\max}$).

3 Experimental Analysis

3.1 Algorithm Flow

The feature extraction process based on SSA and AR is divided into the following six steps:

1. Read the vibration signal in the time zone of the rolling bearing.
2. Vibration signal segmentation.
3. For each segment of data, SSA is used to eliminate the noise.
4. Stationarity test.
5. If the stationarity test fails, difference of the signal.
6. Build the AR model, select model order, and estimate model coefficients.

3.2 Model Simulation

In this study, the bearing vibration data were obtained from the test rig of Case Western Reserve University, the motor speed is 1797 r/min, the signal sampling frequency is 12 kHz. The original data are segmented at a length of 2048, where Fig. 1 is the vibration data for a segment of the rolling bearing. According to the rolling bearing failure frequency calculation formula, we can get the fault frequency, see Table 1. So we choose $f = 103.4$ Hz, and $L \geq \frac{f_s}{f} = \frac{12000}{103.4} \approx 116$. Finally,

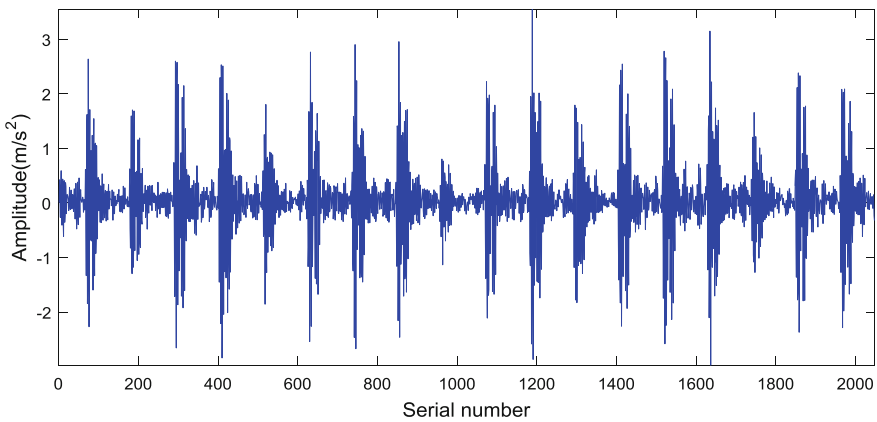


Fig. 1 The original vibration data

Table 1 The roller bearing failure frequency

Failure type	Rotational speed (Hz)			
	Inside circle fault	1797	1772	1750
162.2		159.9	157.9	156.1
Outside circle fault	107.4	105.9	104.6	103.4
Rolling body fault	141.2	139.2	137.5	135.9

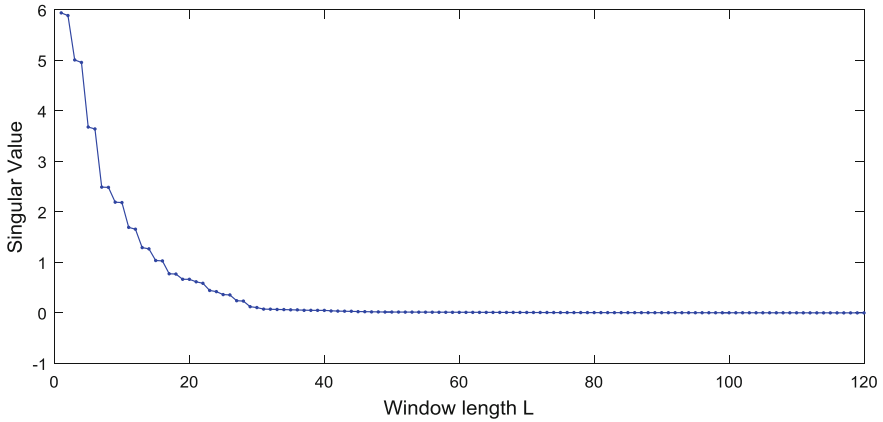


Fig. 2 The singular spectrum of different window lengths

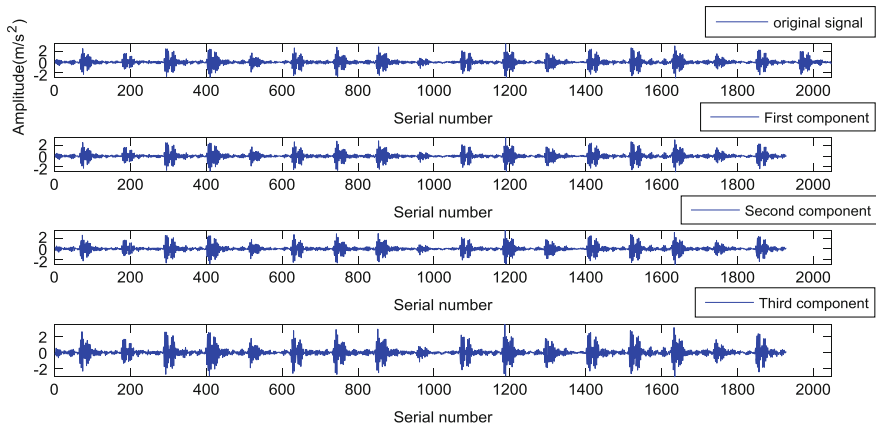


Fig. 3 The original signal and the first three components by SSA decomposition

we determinate $L = 120$, and the corresponding singular spectrum is shown in Fig. 2. Figure 3 shows the original sub-signals and the first three components obtained by SSA decomposition. Obviously, the first three components are good to show the trend of the original signal. The effect after the SSA de-noising is shown in Fig. 4, the signal becomes smoother. After the de-noising and one difference, the selected data are passed the KPSS test, which meets the requirement of stationarity. The de-noising signal is dealing with a difference of order 1, as shown in Fig. 5.

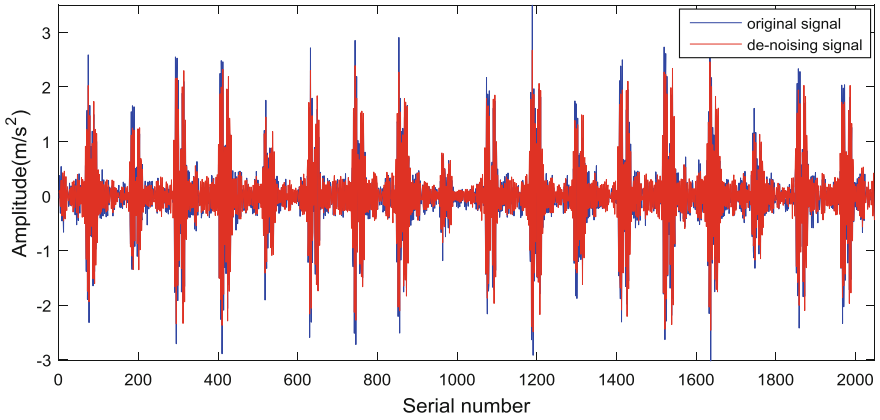


Fig. 4 The original signal and the de-noised signal

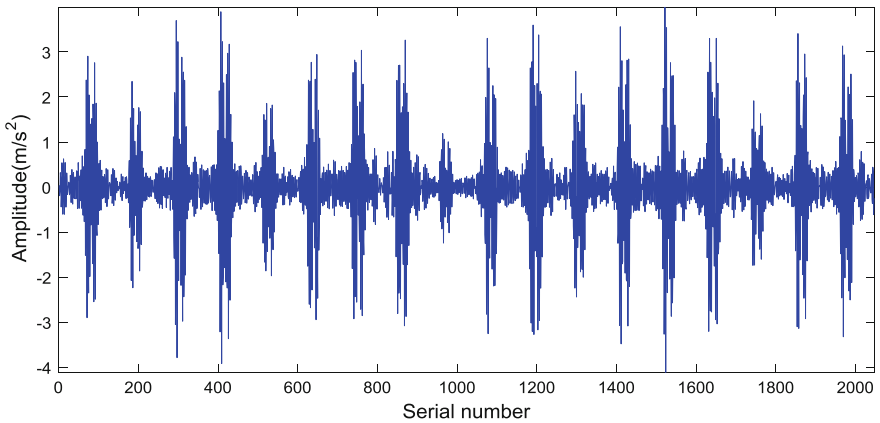


Fig. 5 The signal after the difference of order 1

The BIC value changes with the order of the AR model as shown in Fig. 6. It can be seen that when the order is 69, the BIC value is the smallest, that is, the AR optimal model order is 69. The recursive least squares method is used to calculate the coefficients of the AR model, and the coefficients of the AR model are the eigenvectors of this signal. $NMSE = 99.45\%$ can be obtained by evaluating the NMSE for the evaluation of the autoregressive model. It is shown that the fitting of the original signal by fitting the linear autoregressive model works well. The original signal and the fit signal are shown in Fig. 7.

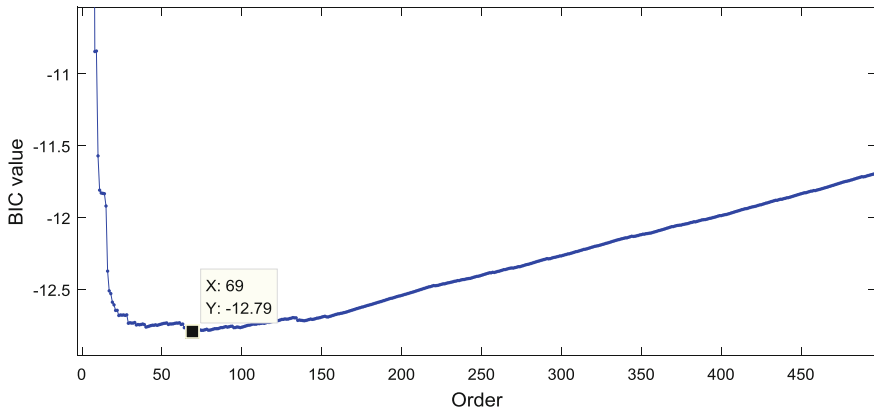


Fig. 6 The BIC values vary with order

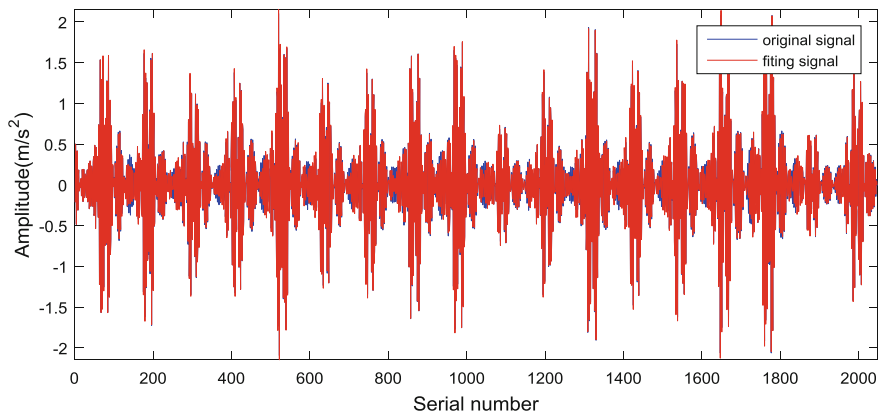


Fig. 7 The original signal and fitting signal

4 Conclusions

In this study, a fault feature extraction method based on SSA and linear autoregressive model is proposed. First, the basic theory of SSA is introduced, including two aspects: decomposition and reconstruction. Then, the related contents of linear autoregressive model are introduced, including data preprocessing, model order, and model parameter estimation. Finally, the proposed method is validated by the data from the test rig of Case Western Reserve University.

Acknowledgements This work is supported by National Key R&D Plan of China under Grant (2017YFB1201201).

References

1. Zheng J, Cheng J, Yang Y (2014) Multiscale permutation entropy based rolling bearing fault diagnosis. *Shock Vib*
2. Howard I (1994) A review of rolling element bearing vibration detection, diagnosis and prognosis'. Defence Science and Technology Organization Canberra (Australia)
3. Elsner JB, Tsonis AA (2013) Singular spectrum analysis: a new tool in time series analysis. Springer Science & Business Media
4. Myung NK (2009) Singular spectrum analysis. University of California, Los Angeles
5. Al-Bugharbee H, Trendafilova I (2016) A fault diagnosis methodology for rolling element bearings based on advanced signal pretreatment and autoregressive modelling. *J Sound Vib* 369:246–265
6. Kwiatkowski D, Phillips PCB, Schmidt P et al (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *J Econometrics* 54(1–3):159–178
7. Wang CC, Kang Y, Shen PC et al (2010) Applications of fault diagnosis in rotating machinery by using time series analysis with neural network. *Expert Syst Appl* 37(2):1696–1702
8. Pang B (2015) Rotational machinery fault feature extraction method. North China Electric Power University (in Chinese)

Passenger Flow Assignment of Evacuation Path in the Station Based on Time Reliability

Chunwei Sun, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan
and Jian Zhang

Abstract In order to realize the reasonable assignment of passenger flow on the evacuation route. Only considering the evacuation time cannot effectively assign the passengers. The path with the shortest theoretical time does not have the highest time reliability. Put forward the measurement of evacuation time reliability. Then, the path impedance function is constructed under the condition of the comprehensive evacuation time reliability and the evacuation time under the congested condition. Combining with the Wardrop equilibrium theory and establish the user equilibrium model based on the evacuation time and time reliability. A method of successive averages (MSA) is used for solving the route assignment result. Finally, the Fuxingmen station is used to verify the validity of the model.

Keywords Time reliability · Impedance function · Congestion degree
User equilibrium assignment model · MSA

1 Introduction

Subway stations are usually enclosed spaces with concentrated passenger flow. Efficient evacuation during emergencies would greatly reduce the loss. The reasonable assignment of the passenger flow can improve the evacuation efficiency. At present, there have been a great deal of researches on passenger flow assignment in station. Through constructing the framework structure of station network and using the restriction of the relationship between passenger flow and evacuation managers, then the game distribution method is used to distribute the evacuation passenger flow in the station [1]. However, the influence of congestion degree on evacuation

C. Sun (✉) · J. Xu · L. Jia · Y. Qin
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: 16125769@bjtu.edu.cn

K. Zhan · J. Zhang
Operations Division, Guangzhou Metro Group Co., Ltd, Guangzhou 510330, China

time is not considered in this method. In the study on evacuation route assignment in subway station, through establishing evacuation route assignment model to assign the passenger flow [2]. However, this study only takes time as the main parameter to describe the path impedance and does not consider the influence of the fluctuation of the evacuation time on the path choice. There also have been a great deal of researches on passenger assignment in rail transit network. In the research of passenger flow assignment in rail transit network, the variational inequality model is further constructed on the basis of the equilibrium principle [3]. The equilibrium demand assignment model based on elastic demand under congested condition [4] and the Frank Wolfe algorithm for solving the model are proposed [5]. Yet, most studies only consider the influence of the evacuation time and do not take into account the influence of other factors on the path choice.

This research proposes the measurement of evacuation time reliability and considers the evacuation time and time reliability in crowded condition to construct the impedance function. Based on the Wardrop equilibrium principle, a user equilibrium traffic assignment model considering the reliability of evacuation time is proposed. MSA algorithm is used to solve the model.

2 Impedance Function of Time Reliability

The impedance function is also called the generalized cost function, on the basis of evacuation time, combining with congestion to construct an impedance function based on time reliability.

2.1 Evacuation Time Function Based on Congestion Degree

The degree of congestion in subway station reflects the time characteristics of passenger aggregation during evacuation [6]. It is determined by the number of evacuees N_i and the evacuation facility capacity C_i . Assume that the number of evacuees was N_i . The congestion degree is calculated as follows:

$$c_i = N_i/C_i \quad (1)$$

The facility capacity is C_i :

$$C_i = \frac{L \times W}{\pi r^2} \quad (2)$$

where the L and W represent the effective length and width of the facility, respectively. Regard the human body as a circle with a radius of R and usually taken 0.25 m.

According to the literature [7], the exponential function is used to describe the congestion efficiency and speed of the station emergency evacuation. That is,

$$v_i = \begin{cases} V_{\max}, & c_i \leq 0.5 \\ V_{\max} \cdot \exp(-0.5 \times c_i), & c_i > 0.5 \end{cases} \tag{3}$$

$$t_i = l_i / v_i(c_i) \tag{4}$$

where l_i is the length of the road section i , and $v_i(c_i)$ is the evacuation speed of the road section i .

The evacuation path is mainly composed of passages and staircases. According to the formula (4), the evacuation time in evacuation path as follows:

$$t_k^{rs} = f(N_k) = \sum_i \sigma_{pi} t_{pi}(N_k) + \sum_i \sigma_{si} t_{si}(N_k), \quad \forall k \in W_{rs} \tag{5}$$

$$\sigma_{pi}, \quad \sigma_{si} = \begin{cases} 0, & i \text{ does not belong to path } k \\ 1, & i \text{ belongs to path } k \end{cases}$$

where t_{pi} represents the travel time of the passages, and t_{si} represents the travel time of the staircases.

2.2 Evacuation Time Reliability

When choosing the actual path, the passenger is more inclined to choose the path with less change in travel time rather than the path with shortest time. The degree of fluctuation of the evacuation time is measured by the reliability of the evacuation time.

Evacuation time reliability refers to the magnitude of the reliability from the starting point to the safe position within the maximum allowable evacuation time. The derivative of the evacuation time under congested conditions and take the reciprocal can be obtained:

$$R_k = \begin{cases} 1, & c_k \leq 0.5 \\ \frac{-0.5V_{\max}}{l_k} \exp(-0.5 \times c_k), & c_k > 0.5 \end{cases} \tag{6}$$

where R_k is evacuation time reliability of path k , l_k is the length of path k , and c_k is congestion of path k .

2.3 Construction of Impedance Function

Taking the maximum time reliability and the minimum evacuation time in congested condition as the criterion of evacuation path assignment. The impedance function is as follows:

$$T_k^{rs} = (1 - \rho)t_k^{rs} + \lambda\rho(1 - R_k) \quad (7)$$

where T_k^{rs} is from the starting point r to the target node s , the total impedance on the evacuation path k . R_k is the evacuation time reliability of path k . ρ is criterion coefficient, that is the degree of preference that passengers with evacuation time reliability maximum as path selection criteria, $0 \leq \rho \leq 1$. When $\rho = 0$, passengers only take the shortest time as a criterion for path selection. When $\rho = 1$, passengers only take the greatest time reliability as a criterion for path selection. λ is positive proportion conversion factor, and it is the parameter which converts the reliability of the evacuation time into the passage time of the road section. In general, its value is based on free flow time and is 10^n . Among it, n is the number of bits of passenger evacuation time.

2.4 User Equilibrium Assignment Model Based on Evacuation Reliability

According to Wardrop's first equilibrium principle [8], taking the system optimal assignment as the goal to make the impedance of all paths is minimized. Combining with the impedance function presented above. The user equilibrium assignment model considering the evacuation time reliability is established as follows:

$$\begin{aligned} \min Z &= \sum_k T_k^{rs}(N_k) \\ \text{s.t.} &\begin{cases} t_k^{rs} = f(N_k) \\ \sum_k N_k = N \\ N_k \geq 0 \end{cases} \end{aligned} \quad (8)$$

2.5 Algorithm Design

In the traditional traffic assignment method, the assignment result of MSA is closest to the equilibrium assignment. Combining with the traditional MSA allocation algorithm. Considering the evacuation time reliability and path time impedance. The specific assignment steps of evacuation path assignment based on the reliability of evacuation time are as follows:

Step 1. Conducting all-or-nothing assignment for the first time according to t_0 . Calculate the traffic volume of each road section; $X_a^0 = X_a(0)$, $\forall a$, and $n = 0$.

Step 2. Let $n = n + 1$, calculate the time impedance on the road section. $t_a^n = t_a(X_a^{n-1})$, $\forall a$.

Step 3. Calculate the time impedance T_k of all evacuation paths according to the road time impedance. Find the shortest path k_n by contrast.

Step 4. Calculate the evacuation time reliability R_a of each path according to the formula (6), and calculate the total impedance $T_{k_n}^Z$ of each path by the formula (7). Contrast the impedance of each route to find the shortest one.

Step 5. According to the new path impedance of the OD between the traffic on the k_n on an all-or-nothing assignment. Calculate the additional F_a^n of each road section.

Step 6. The original traffic volume and additional traffic weighted average. Calculate the new traffic volume X_a^n for each road section, $X_a^n = (1 - \frac{1}{n})X_a^{n-1} + \frac{1}{n}F_a^n, \forall a$.

Step 7. Calculate the shortest path impedance and other effective path impedance. When $|T_{k_n}^z - T_{k_{n-1}}^z| \leq \varepsilon$, the iteration stops. The result of the last round of assignment is the final assignment scheme. Otherwise, return to step (2).

3 Case Study

Verify the effectiveness of the model by the route assignment of the Beijing metro Fuxingmen Station.

Fuxingmen Station is the transfer station of Beijing Metro Line 1 and Line 2 with T-type structure. It is north and south sides symmetrical, and there are four exits, which are A (northwest), B (northeast), C (southeast), and D (southwest), respectively. The platform of Line 2 was north-south layout. There are two staircases in the middle of the platform. It is connected with the west side of the staircase of Line 1 for passenger flow transferring to the platform of Line 1. There is a staircase on each side of the platform, which is connected with the north-south transfer station hall. The platform of Line 1 was east-west layout. The stairs on the east side of the platform are connected to the east transfer hall. North and south sides have a transfer channel which can be transferred to the platform of Line 2. Dimensional structure of Fuxingmen station is shown in Fig. 1.

3.1 Construction of Evacuation Path Network

Using the method of graph theory to describe the structure of Fuxingmen station. Network $G = \{O, V, D, E\}$, where O, V, and D are node sets, and E is edge set. O is the starting node set which includes the platform. D is the exit node set which is the collection of safe outlets in the station, $D = \{d_j | j = i, i + 1, \dots, m\}$. V is the node set which is the connection of facilities and equipment, $V = \{v_t | t = n, n + 1, \dots, p\}$. E is edge set which represents the connection of

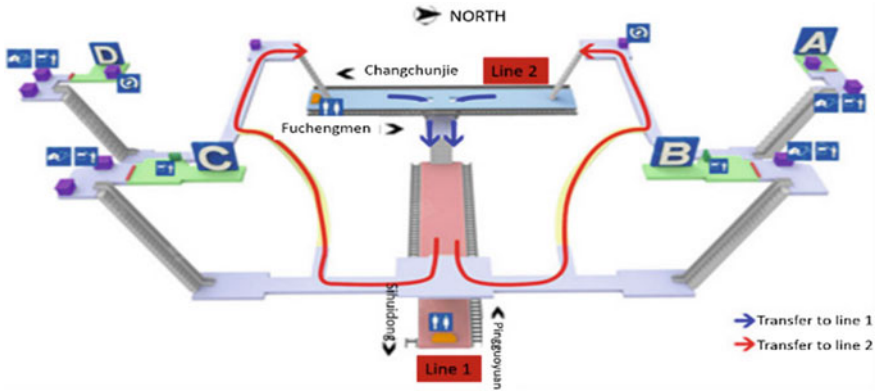


Fig. 1 Dimensional structure of Fuxingmen station

passage, stairs, escalators, and other facilities, $E = \{e_{ij} | e_{ij} = (v_i, v_j); v_i, v_j \in V\}$. Construct the evacuation path network model for station. As shown in Fig. 2.

3.2 Initialization Condition of Network

Calculating the capacity of each road section according to formula (2). Gaining the free flow evacuation speed according to the result of Ref. [9]. Velocity of passage take 2 m/s and stairs take 1.26 m/s. Calculate the free flow travel time (Table 1).

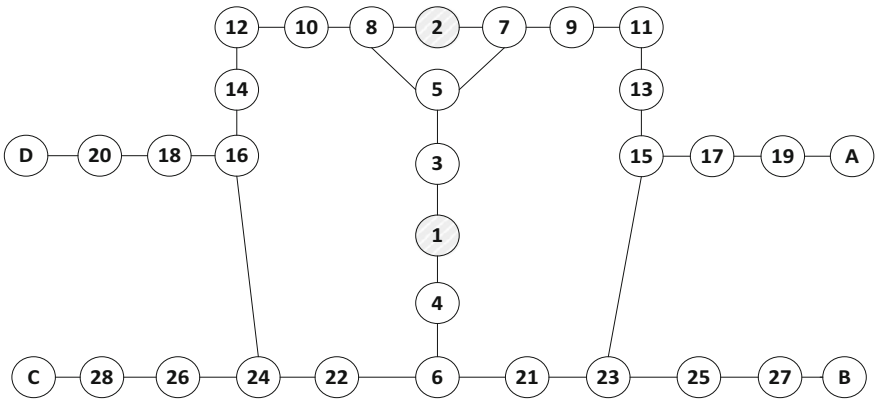


Fig. 2 The network structure of Fuxingmen station

Table 1 Initial conditions of road network

Road section	Length/m	Capacity (p/m ²)	Free flow travel time/s	Road section	Length/m	Capacity (p/m ²)	Free flow travel time/s
1-3	57	2903	28.50	14-16	26.8	546	13.40
1-4	57	2903	28.50	15-17	40	815	20.00
2-7	10	560	5.00	15-23	67.2	1506	33.60
2-8	10	560	5.00	16-18	40	815	20.00
3-5	4.8	112	3.81	16-24	67.2	1506	33.60
4-6	11	314	8.73	17-19	21	428	16.67
5-7	9.5	116	7.54	18-20	20.3	414	16.11
5-8	9.5	116	7.54	19-A	10	275	5.00
6-21	5	300	2.5	20-D	14	385	7.00
6-22	5	300	2.5	21-23	26.4	538	13.20
7-9	51	2857	25.50	22-24	26.4	538	13.20
8-10	51	2857	25.50	23-25	13.6	277	6.80
9-11	9.5	232	7.54	24-26	13.6	277	6.80
10-12	9.5	232	7.54	25-27	36	733	28.57
11-13	5.5	247	2.75	26-28	36	733	28.57
2-14	5.5	247	2.75	27-B	14	399	7.00
13-15	26.8	546	13.40	28-C	15	405	7.5

3.3 Passenger Flow Evacuation Path Distribution

Through the basic situation of the station and the road network initialization conditions, we can get all the evacuation path set and path evacuation free flow travel time (Table 2).

(1) Evacuation path assignment based on travel time

Combining the impedance function defined by 2.3 and the equilibrium allocation model defined in 3.1 and proposing the MSA algorithm, make evacuation

Table 2 Evacuation path set and free flow travel time

Number of path	Evacuation path composition	Free flow time/s
Path 1	1-3-5-7-9-11-13-15-17-A	127.956
Path 2	1-3-5-8-10-12-14-16-18-D	129.400
Path 3	1-4-6-19-13-15-17-A	125.697
Path 4	1-4-6-19-21-23-B	92.802
Path 5	1-4-6-20-14-16-18-D	127.141
Path 6	1-4-6-20-22-24-C	93.302
Path 7	2-7-9-11-13-15-17-A	93.106
Path 8	2-8-10-12-14-16-18-D	94.551

assignment for the constructed evacuation network of the Fuxingmen station. The criterion coefficient ρ is taken as 0, which indicates that the passengers take the shortest travel time as the evacuation path selection criteria. Approximate equalization interval $\varepsilon = [-0.5, 0.5]$. The results of the distribution and the evacuation time are shown in Tables 3 and 4.

Assigning based on the path of the travel time. The number of evacuees at the subway stations is, respectively: 379 people in exit A, 769 people in exit B, 756 people in exit C, and 536 people in exit D. As pedestrians take the shortest path travel time as the criteria for evacuation path selection. The path with the shortest travel time is preferred. So it is expected that the path with too long travel time is less likely to be selected and even not being chosen, such as path 3 and path 5. The shorter the travel time is, the greater the probability that the path will be selected. So the more number of pedestrian assigned.

(2) Assignment results considering evacuation time reliability

The criterion coefficient ρ is taken as 0.5, which indicates that the passengers at the moment take into account the minimum evacuation time and the maximum reliability of the evacuation time as the evacuation path selection criteria. Positive proportion conversion factor $\lambda = 1000$ (As the free flow of each path travel time is more than three digits and take $n = 3$). Approximate equalization interval $\varepsilon = [-0.5, 0.5]$. The results of the assignment and the evacuation time are shown in Tables 5 and 6.

Considering both the path travel time and the reliability of the time. The number of evacuees at the subway stations is, respectively: 453 people in exit A, 679 people in exit B, 664 people in exit C, and 644 people in exit D. The result shows the pedestrians take into account path travel time and time reliability when selecting path. The path with short travel time is not necessary choice for more people. In order to evacuate to a safe location safely. Taking into account the degree of

Table 3 The evacuation path assignment results of Line 1

Number of path	Flow distribution/ p	Distribution ratio/ %
Path 1	298	13.91
Path 2	317	14.83
Path 3	0	0
Path 4	769	35.92
Path 5	0	0
Path 6	756	35.34

Table 4 The evacuation path assignment results of Line 2

Number of path	Flow distribution/ p	Distribution ratio/ %
Path 7	81	27.00
Path 8	219	73.00

Table 5 The evacuation path assignment results of Line 1

Number of path	Flow distribution/p	Distribution ratio/%
Path 1	301	14.08
Path 2	342	15.98
Path 3	74	3.45
Path 4	679	31.72
Path 5	80	3.74
Path 6	664	31.03

Table 6 The evacuation path assignment results of Line 2

Number of path	Flow distribution/p	Distribution ratio/%
Path 7	78	26.00
Path 8	222	74.00

congestion in the evacuation process and pedestrians will choose the path with long travel time, such as path 3 and path 5. The number of evacuees at each exit is smaller than in the case where take the shortest time as criterion. At this time, the evacuation is more efficient.

4 Conclusion

At present, the problem of passenger flow assignment on evacuation path often only considers the evacuation time and other unilateral factors, which does not take into account the other factors on the passenger assignment of the composite effect. In this paper, considering the time reliability and the evacuation time in the congested condition, the following results are obtained:

- (1) Put forward the measurement of evacuation time reliability. This measurement takes into account the two factors of evacuation time and evacuation time reliability. Minimizes time and time reliability as the evacuation path assignment criteria. An impedance function considering time reliability is proposed to determine the path selection.
- (2) According to Wardrop equilibrium principle and establish a user equilibrium assignment model based on evacuation time reliability. The MSA algorithm is used to solve the user equilibrium assignment model and get the assignment results.
- (3) Using the user equilibrium assignment model to assign passenger flow in evacuation path under the emergencies of Line 1 of Fuxingmen station. Through the analysis, it is found that the assignment model which considering the evacuation time and the evacuation time reliability is more effective than the

model which only considering the unilateral factors. It is of great significance to the future evacuation strategy.

This paper only takes the platform passengers as the evacuation objective. The next study can research the assignment of evacuation path for the entire station passengers based on the user equilibrium assignment model.

Acknowledgements The authors gratefully acknowledge the support provided by China National “13th Five-Year” key research project “Safety assurance technology of urban rail system” (Grant No.2016YFB1200402) and “National Engineering Laboratory for System Safety and Operation Assurance of Urban Rail Transit.”

References

1. Beckman MJ, Mcguire CB (1956) Studies in the economics of transportation. Yale University Press, New Haven
2. Smith MJ (1979) The existence uniqueness and stability of traffic equilibria. *Transp Res B* 13 (4):295–304
3. Kong F, Li X (2008) Research on passenger flow assignment model and algorithm of rail transit under elastic demand. *W J Transp Sci Technol* 1:83–87 (in Chinese)
4. Wang Y (2014) A study on passenger flow assignment and design of emergency evacuation schedule in urban rail transit stations. Master thesis, Beijing Jiaotong University (in Chinese)
5. Zhao K, Zhao G (2004) The uncertainty urban planning ideology. *Urban Plan J*:33–36 (in Chinese)
6. Gao R, Li A, Hao X, Lei W, Deng B (2012) Prediction of the spread of smoke in a huge transit terminal subway station under six different fire scenarios. *Tunn Undergr Space Technol incorporating Trenchless Technology Research*
7. Zhou W (2010) Study on the model in the passenger flow distribution forecast of inner-city rail transit with its application. Master thesis, Southwest Jiaotong University (in Chinese)
8. Wardrop JG (1952) Some theoretical of road traffic research. In: *Proceedings of the Institute of Civil Engineers VI*, part 2, pp 325–378
9. Zuo D (2010) The user equilibrium traffic assignment method under interval uncertain demand. Master thesis, Changsha University of Science & Technology (in Chinese)

Simulation of Aging Characteristics of Power Switch Device Based on Saber

Yujia Guo, Mengzhu Wang, Lei Wang, Ruichang Qiu
and Guofu Chen

Abstract IGBT is a power switch device that is very important in modern power electronics. The reliability of IGBT concerns the security of circuits and even the whole system, while IGBT module's aging is inevitable, so it is necessary to study the aging of IGBT module. At home and abroad, the measurement of the electrical parameters when the bonding wires fall off is used to identify the aging failure state of IGBT, but the consideration is far from comprehensive, the evaluation is also inaccurate. This paper uses the saturation voltage drop which is easy to measure as a state characteristic parameter and considers the influence of junction temperature and collector current to the saturation voltage and IGBT module's aging failure to evaluate the aging failure of IGBT module.

Keywords IGBT · Saber · Aging characteristic simulation

1 Introduction

IGBT (Insulated Gate Bipolar Transistor) is composed of MOSFET (Metal–Oxide–Semiconductor Field-Effect Transistor) and BJT (Bipolar Junction Transistor). IGBT fusions excellent characteristics of the two devices, it is driven by voltage as well as MOSFET and also has the advantages of low saturation voltage, large capacity of BJT. The conductivity modulation effect of IGBT decreases the size of the chip and greatly reduces the manufacturing cost.

IGBT power module is widely used in the harsh condition of the wind power generation, electric traction, and aerospace power systems. The electric heat level is

Y. Guo (✉) · M. Wang · L. Wang · R. Qiu
Beijing Engineering Research Center of Electric Rail Transportation,
School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China
e-mail: 16121445@bjtu.edu.cn

G. Chen
State Key Laboratory of Advanced Power Transmission Technology,
Global Energy Interconnection Research Institute, Beijing, China

higher and higher which requires a high degree of safety and reliability of IGBT. According to the relevant survey, faults in power electronic circuits mostly come from the aging failure of power semiconductor devices, especially for new energy power system, the failure of IGBT module is the main reason for system shutdown and causes huge economic loss or safety accident [1]. Therefore, it is important to study the aging status of IGBT.

At present, the research on IGBT aging is done by monitoring the change of the state characteristics of IGBT module [2–5]. But for most detection of these state parameters, it was studied by accelerated aging test and does not realize the online monitoring of the IGBT aging and the dynamic prediction of service life.

In fact, the breakage of bonding wires is the main failure mode of IGBT. Saturation voltage drop is also affected by the collector current and junction temperature. After the breakage of bonding wires, less bonding wires bear the same current, lead to the rise of IGBT junction temperature. Junction temperature is not only accelerating aging of IGBT but also the influence factors of other aging characteristic parameters. This paper uses the saturation voltage drop which is easy to measure as a state characteristic parameter and considers the influence of junction temperature and collector current to the saturation voltage and IGBT module's aging failure to evaluate the aging failure of IGBT module.

2 Materials and Methods

2.1 Mechanism of IGBT Aging

The internal structural profile of IGBT module is as shown in Fig. 1 [6].

The main reason for the failure of IGBT is heat stress in the process of running. The internal temperature gradient is formed under different quantity of heat between different layers in Fig. 1. It is a hot-mechanical process. The coefficient of thermal expansion (CTE) of different material is as shown in Fig. 2. Under the certain thermal shock effect, different material bear different levels of heat stress due to different material CTE. After long-term accumulation, thermal stress is greater than

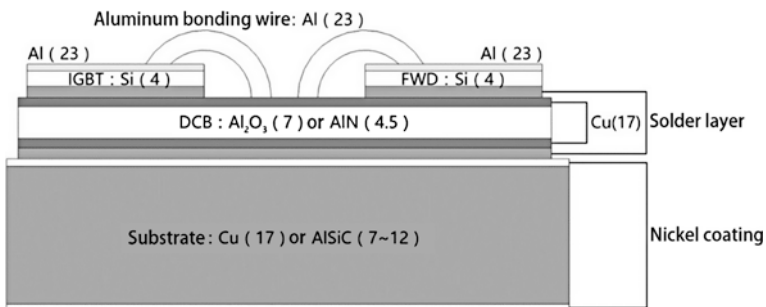
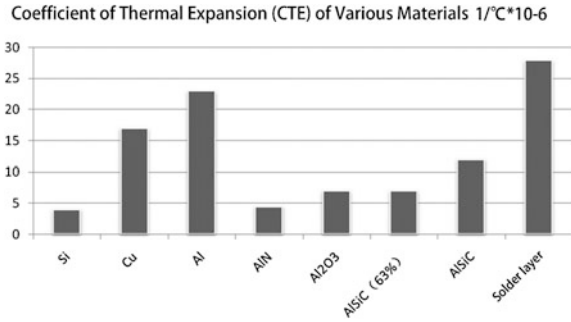


Fig. 1 IGBT module internal structure profile

Fig. 2 Coefficient of thermal expansion (CTE) of different materials



the strength of device due to external thermal shock effect, and the device is likely to fail.

This paper discusses the rule of IGBT module in the working area of the bathtub curve aging failure zone in Fig. 3 [7]. If we can find the aging failure characteristic parameter which is easy to get and changing obviously, we can evaluate and even predict the state of IGBT.

2.2 Saturation Voltage Drop of IGBT

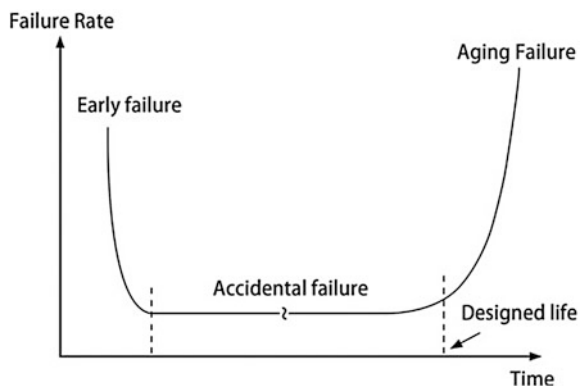
The structure and equivalent circuit of IGBT are given in Fig. 4.

The conduction voltage drop of IGBT is as shown in Eq. (1) according to the structure and equivalent circuit of IGBT in Fig. 4.

$$V_{cesat} = V_{J1} + V_{drift} + I_c \times R_{channel} \tag{1}$$

where V_{cesat} is saturation voltage drop, V_{J1} is conduction voltage drop of PN junction J_3 , and V_{drift} is the voltage drop on the drift resistance R_{drift} . $R_{channel}$ is the equivalent conduction resistance of p-zone, and I_c is collector current.

Fig. 3 The bathtub curve of IGBT module failure



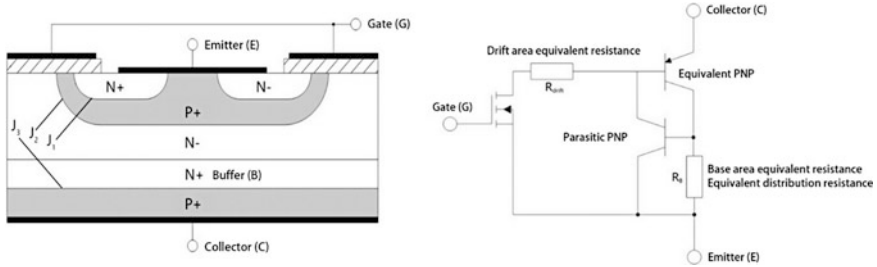


Fig. 4 The structure and equivalent circuit of IGBT

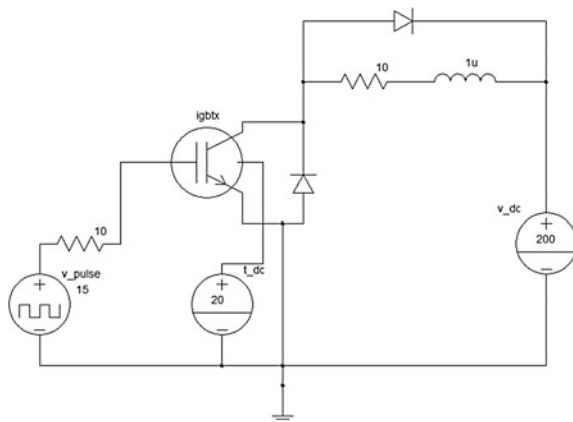
The temperature can affect the migration efficiency of the carrier, so the conductivity modulation effect and $R_{channel}$ and I_c will change along with the change of temperature. V_{cesat} has relation with I_c and $R_{channel}$, so T_j is closely related to V_{cesat} and also has a great influence on other aging characteristic parameters. Let us look at the relationships between V_{cesat} , I_c , and T_j .

3 Results

3.1 The Relationship Between Saturation Voltage Drop, Collector Current, and Junction Temperature

Use Saber (a system simulation software) to verify the relationship between saturation voltage drop, collector current, and junction temperature. Simulation circuit diagram is as shown in Fig. 5. IGBT module adopts the Hefner IGBT model of Self-Heat IGBT module which has a better dynamic thermal characteristic [8].

Fig. 5 IGBT aging simulation circuit diagram



First, I_c value is fixed by V_{cc} , and T_j junction temperature varies from 20 to 100 °C step by 20 °C. The waveform and enlarged figure are as shown in Fig. 6.

And then the temperature T_j is fixed, the current varies from 20 to 40 A step by 5 A which is the same as the voltage varies from 200 to 400 V step by 50 V. The waveform is as shown in Fig. 7a. The enlarged figure is as shown in Fig. 7b.

The measurement data of saturation voltage drop of IGBT at each current and junction temperature are given in Table 1.

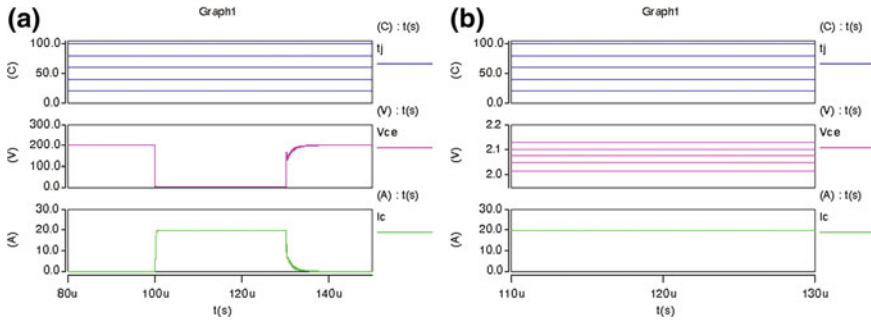


Fig. 6 a The saturation voltage drop when I_c is fixed and T_j is changed. b Saturation voltage drop under magnification

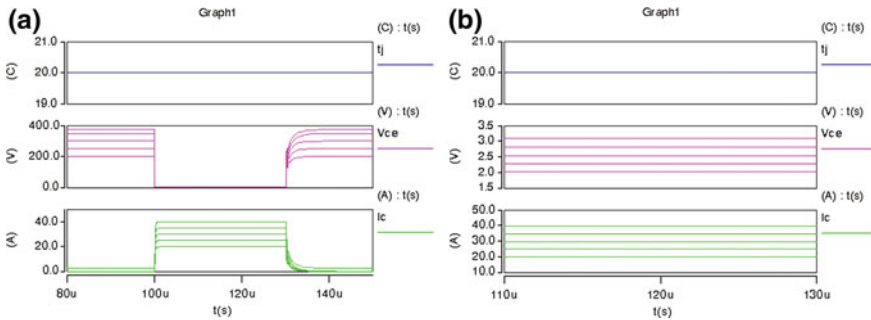


Fig. 7 a The saturation voltage drop when T_j is fixed and I_c is changed. b Saturation voltage drop under magnification

Table 1 The measurement data

T_j/C	I_c/A				
	20	40	60	80	100
20	2.0175	2.0463	2.0743	2.1015	2.1281
25	2.2768	2.3152	2.3524	2.3885	2.4237
30	2.5416	2.5899	2.6367	2.6821	2.7261
35	2.8144	2.8733	2.9302	2.9853	3.0387
40	3.0978	3.1681	3.2361	3.3017	3.3653

According to the data in table, three-dimensional surface of V_{cesat} , I_c , and T_j is as shown in Fig. 8. Saturation voltage drop, collector current, and junction temperature have a good linear relationship between each other. I_c has a large influence on V_{cesat} , and the influence of T_j is also not small, so it is necessary to consider junction temperature. So we can describe the aging failure of IGBT by saturation voltage drop under the influence of collector current and junction temperature.

3.2 Aging Rule Simulation of IGBT

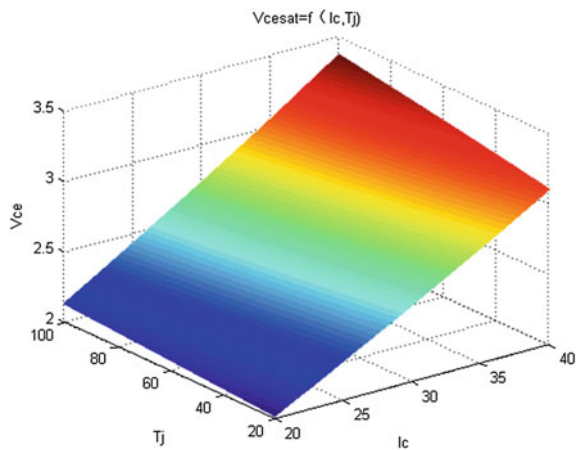
3.2.1 The Shortcomings of Past Evaluation of Aging

The above analysis shows that junction temperature as a key variable plays an important role in aging analysis. Junction temperature is not only accelerating aging of IGBT but also the influence factors of other aging characteristic parameters. However, the past aging analysis only uses saturation voltage drop to describe the aging of IGBT and does not fully consider the influence of junction temperature and collector current [9].

The last section analyzes the relationship between saturation voltage drop, collector current, and junction temperature. Because of the effect of temperature on saturation voltage drop, it is difficult to determine whether the temperature causes the change of saturation voltage drop or the IGBT is actually damaged. This is the shortcomings of past evaluation of aging.

So the following will consider I_c , T_j 's effect on saturation voltage drop and use I_c , T_j , and V_{cesat} to describe the aging of IGBT to make it more accurate and persuasive.

Fig. 8 $V_{cesat} = f(I_c, T_j)$
three-dimensional surface



3.2.2 Simulation of the Breakage of Bonding Wire

Use Saber to simulate the whole process of bonding wires damage online. Bonding wires damage is simulated by changing the value of the conductance resistance. The aging of IGBT is described by I_c , T_j , and V_{cesat} .

1. The bonding wires fall off 1/2. The conductance resistance (R_s) is twice as large as the original. The three-dimensional relationship diagram of V_{cesat} , I_c , and T_j is as shown in Fig. 9a.
2. The bonding wires fall off 3/4. The conductance resistance (R_s) is four times as large as the original. The three-dimensional relationship diagram of V_{cesat} , I_c , and T_j is as shown in Fig. 9b.
3. The bonding wires fall off 7/8. The conductance resistance (R_s) is eight times as large as the original. The three-dimensional relationship diagram of V_{cesat} , I_c , and T_j is as shown in Fig. 9c.
4. Four three-dimensional relationship diagrams of the saturation voltage drop, collector current, and temperature are shown in Fig. 9d.

It can be seen through the simulation results, V_{cesat} increases with the increase in I_c and T_j . In order to compare the relationship between the saturation voltage drop,

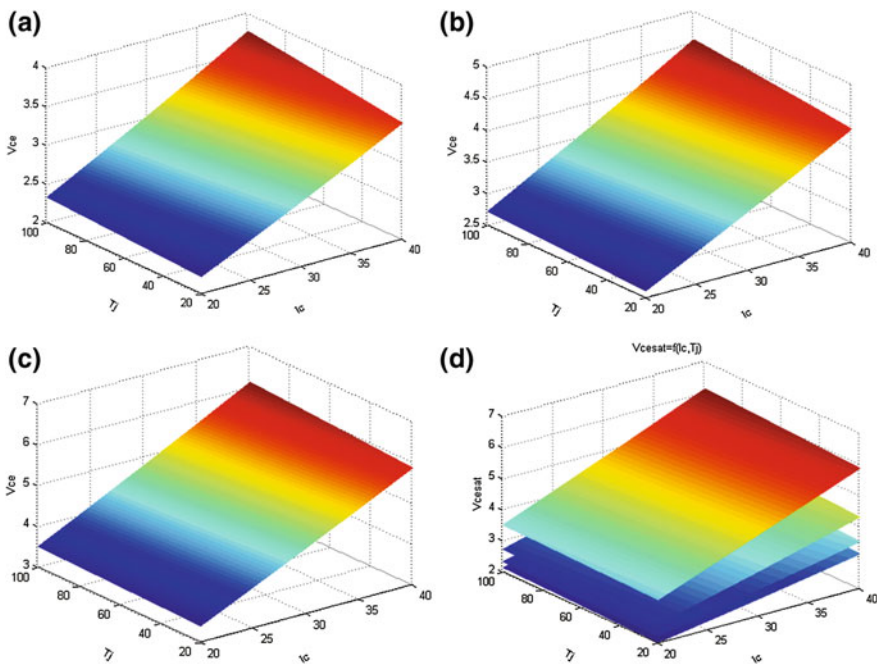


Fig. 9 **a** The 3D diagram of $V_{cesat} = f(I_c, T_j)$ when the bonding wires fall off 1/2; **b** the 3D diagram of $V_{cesat} = f(I_c, T_j)$ when the bonding wires fall off 3/4; **c** the 3D diagram of $V_{cesat} = f(I_c, T_j)$ when the bonding wires fall off 7/8; and **d** four 3D diagrams of $V_{cesat} = f(I_c, T_j)$

collector current, and junction temperature before and after the breakage of the bonding wires more intuitive, clearly and convenient, put four kinds of condition into a three-dimensional diagram. It is as shown in Fig. 9d. The surface is higher after the breakage of bonding wires, and the more bonding wires break, the higher the surface is. Such as at $T_j = 20\text{ }^\circ\text{C}$ and $I_c = 40\text{ A}$, $V_{\text{cesat}} = 2.1281\text{ V}$ when bonding wires are in good condition, V_{cesat} is 2.3247, 2.7166, and 3.4956 V when the bonding wires fall off 1/2, 3/4, and 7/8. It is clear that the increasing amplitude of IGBT saturation voltage drop increases with the increase in the roots of bonding wire. Because current can only flow through less bonding wire, namely, the effective diversion area is reduced. The less the bonding wires are, the greater the increasing amplitude of IGBT saturation voltage drop is.

It can truly reflect the aging of the IGBT module when the bonding wires fall off through the simulation of saturation voltage drop V_{cesat} along with the change of I_c and T_j . That is to say, we can use this way to assess the state of the IGBT module and predict life.

4 Conclusion

IGBT is a power switch device that is very important in modern power electronics. The reliability of IGBT concerns the security of circuits and even the whole system, so it is necessary to study the aging of IGBT module.

The breakage of bonding wires is the main failure mode of the aging of IGBT module. Use saturation voltage drop which is easy to measure as a state characteristic parameter and consider the influence of junction temperature and collector current to the saturation voltage to evaluate the aging failure of IGBT module. The results of simulation show that saturation voltage drop, collector current, and junction temperature have a good linear relationship between each other. So we can describe the aging of IGBT by saturation voltage drop with the influence of collector current and junction temperature.

Acknowledgements This work was supported by The National Key Research and Development Program of China (2016YFB1200504-C-02), the Fundamental Research Funds for the Central Universities of China (2016JBM062), and Satellite Power Supply Failure Feature Extraction and Health Degree Calculation Modeling Project (WY - YY/D022011324JY015).

References

1. Liu Y (2013) Research on the reliability analysis of power electronic circuits based on complex network theory. South China University of Technology (in Chinese)
2. Sankaran VA, Chen C, Avnat CS, Xu X (1997) Power cycling reliability of IGBT power modules. In: IEEE industry application society annual meeting, pp 1222–1227

3. Oukaour A, Tala-Ighil B, Pouderoux B, Tounsi M, Bouarroudj-Berkani M, Lefebvre S, Boudart B (2011) Ageing defect detection on IGBT power modules by artificial training methods based on pattern recognition. *Microelectron Reliab* 51:386–391
4. Xiang D, Ran L, Tavner P, Yang S, Bryant A, Mawby P (2010) Monitoring solder fatigue in a power module using the rise of case—above-ambient temperature. In: *Energy conversion congress and exposition (ECCE), IEEE*, pp 955—962
5. Zhou L, Zhou S (2010) Effects of wire-bond lift-off on gate circuit of IGBT power modules. In: *Power electronics and motion control conference (EPE/PEMC), 2010 14th International 6–8 Sept 2010*, pp 45–47
6. Yang X (2012) Method of state evaluation of IGBT power module based on saturation voltage drop. Chongqing University (in Chinese)
7. Mi W, Yan Y (2015) The reliability of the power IGBT module. *Electron Devices* 01:27–31 (in Chinese)
8. Xiong Y, Shen Y, Jiang J, He X (2006) IGBT loss calculation and loss model research. *Power Technol Appl* 05:55–60 (in Chinese)
9. Xia J, Han W, Gu L (2011) The application of Saber and Matlab in the dynamic simulation of IGBT. *Electron World* 23:68–69 (in Chinese)

Evaluation of Emergency Evacuation Capacity of Subway Station Based on M/G/c/c

Yacui He, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan
and Jian Zhang

Abstract In the case of an emergency, accurately assessment of evacuation capability is the key to safe and efficient evacuation. M/G/c/c represents a queuing model where customers' arrivals obey Poisson distribution, the service time obeys general distribution, the system has c service desks, and its capacity is c. It is suitable for the analysis of pedestrian walking process and the calculation of corresponding indicators in various spaces. The evacuation time, passenger flow density and facility bottleneck are chosen as evaluation indexes to analyze pedestrian walking characteristics and connection structure of station divergence nodes. Based on the M/G/c/c queuing model, a pedestrian divergence node probabilistic selection optimization model is constructed. Then, we use lingo software to solve the model. And Jianguomen Subway Station is used to verify its effectiveness.

Keywords Subway station · Dividing node · M/G/c/c · Lingo
Emergency evacuation capability

1 Introduction

Subway stations are often crowded with people; an accurate assessment of evacuation capability is the key to ensuring safe and efficient evacuation in case of emergency. Pre-assessment of emergency evacuation capability will help to provide improved solutions for passenger evacuation under emergencies.

The assessment of evacuation capacity evaluation preserves the status quo:
(1) The analysis and modeling of microscopic evacuation behavior: Study pedes-

Y. He (✉) · J. Xu · L. Jia · Y. Qin
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: 16120813@bjtu.edu.cn

K. Zhan · J. Zhang
Operations Division, Guangzhou Metro Group Co., Ltd,
Guangzhou 510330, China

trian flow and establish the model of route choice [1]; a simulation model based on cellular automata [2]; social force model [3] et al. (2) Modeling or emulating by using mathematical physics and computer: system service and capacity allocation problem based on M/G/c/c model [4]; a bionic cellular automaton model based on pedestrian dynamics, [5], etc. Although the relative research is relatively rich, it does not take into account the characteristics of passenger movement, facilities, service characteristics, and the relationship between them from the overall station.

2 Evaluation of Emergency Evacuation Capability

In the book of *Subway Design Code* (GB 50157-2013) [6], the definition of station evacuation capacity is mainly reflected in the evacuation time; the formula is as follows:

$$T = 1 + \frac{Q_1 + Q_2}{0.9[A_1(N - 1) + A_2B]} < 6 \text{ min} \quad (1)$$

where Q_1 and Q_2 indicate the maximum passenger flow in section and the largest waiting passengers on the platform, respectively, A_1 is the carrying capacity of escalator, A_2 is the carrying capacity of evacuation staircase, N is the number of escalators, and B is the total width of the evacuation staircase.

In formula (1), the evacuation time is mainly related to the passenger flow, width, and number of stairs and escalators. In practice, the connection structure of facilities, the number of exports [7], and the pedestrian selection behavior [8] will also have an impact on evacuation capacity. In order to facilitate the capacity assessment, select evacuation time [6], passenger density [9], and facility bottlenecks for analysis.

3 M/G/c/c Queuing Model for Evacuation Capacity

3.1 Single-Facility M/G/c/c Queuing Model

M/G/c/c is a state-dependent model, which is suitable for analyzing the characteristics of single-facility service under dynamic environment [10]. When there are n pedestrians, the model is shown as [11]

$$P(n) = \frac{[\lambda E(T_1)]^n P_0}{n! f(n) \dots f(2) f(1)} = P_0 \prod_{n=1}^c \frac{\lambda E(T_1)}{n \cdot f(n)} \quad (2)$$

where

$$P_0 = \left[1 + \sum_{n=1}^c \frac{[\lambda E(T_1)]^n}{n! f(n) \dots f(2) f(1)} \right]^{-1} = \left[1 + \sum_{n=1}^c \prod_{i=1}^n \frac{\lambda E(T_1)}{i \cdot f(i)} \right]^{-1} \tag{3}$$

$$f(n) = v_n / v_1 \tag{4}$$

$$E(T_1) = l / v_1 \tag{5}$$

$$c = \text{INT}(\rho \cdot l \cdot w) \tag{6}$$

P_0 is the probability of no pedestrians in the facility, $f(n)$ is passenger service rate, $E(T_1)$ is the expected service time for a single pedestrian within the facility, C is facilities maximum capacity, and λ is the arrival rate.

And l is the corridor length, w is the corridor width, v_n is the average walking speed for n pedestrians in a corridor, v_1 is the individual walking speed, $v_1 = 1.5$ m/s [12], ρ is the passenger limit density within the facility, $\rho = 5$ person/m² [12].

In essence, $f(n)$ is a function of the number of pedestrians in the facility, and the relationship of v_n and v_1 can be described by the velocity–density function, which is shown as [11]

$$v_n = v_1 \exp \left[- \left(\frac{n - 1}{\beta} \right)^\gamma \right] \tag{7}$$

where β and γ are the scale and shape parameters, respectively [11]:

$$\beta = \frac{(a - 1)}{\left[\ln \left(\frac{v_1}{v_n} \right) \right]^{1/\gamma}} = \frac{(b - 1)}{\left[\ln \left(\frac{v_1}{v_b} \right) \right]^{1/\gamma}} \tag{8}$$

$$\gamma = \frac{\ln \left[\frac{\ln(v_a/v_1)}{\ln(v_b/v_1)} \right]}{\ln \left(\frac{a-1}{b-1} \right)} \tag{9}$$

v_a is the average walking speed when crowd density = 2 person/m², v_b is the average walking speed when crowd density = 4 person/m², and $a = 2lw$, $b = 4lw$ [11].

Revising the pedestrian speed as follows in the case of an emergency [13]:

$$v_e = v \cdot \mu_e \tag{10}$$

In formula (10), v_e is the speed under emergency situations, v is the speed under normal circumstances, μ_e is the correction factor, and in upstairs $\mu_e = 1.26$; in downstairs $\mu_e = 1.212$; in channel $\mu_e = 1.49 - 0.36D$, where D means flow density.

$$D = nf / W_f L_f \tag{11}$$

In formula (11), f is the horizontal projection area of a single pedestrian (0.18 m²/person [14]), W_f is the flow width, L_f is the flow length.

According to the Little Formula of the queuing theory, we can get the formula of the relevant index [13]:

$$p_c = P[n = c] \tag{12}$$

$$\theta = \lambda(1 - p_c) \tag{13}$$

$$E[N] = \sum_{n=1}^c np_n \tag{14}$$

$$E[T] = E[N] / \theta \tag{15}$$

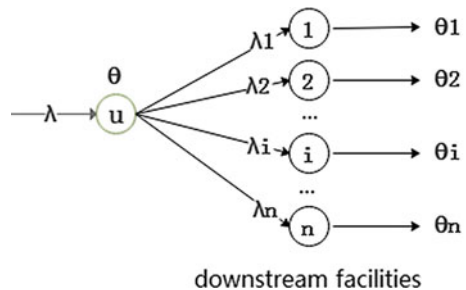
p_c is the probability of congestion on a staircase or passageway, Θ is the output people per second (person/s), $E[N]$ is the expected service number in the channel or staircase, and $E[T]$ is expected service time of passenger, including waiting and service time.

3.2 Optimal Model of Probability Selection for Split Node

In the emergency evacuation, the stations take guided measures to evacuate pedestrians into safe areas as soon as possible, that is, the number of evacuation per unit more the better. Hence, we can base on M/G/c/c model, constructing node probabilistic selection optimization model for pedestrian on the dividing facilities.

In Fig. 1, the node u represents the uplink facility; nodes 1, 2, ..., n represent the downstream facilities, their collection is denoted as V_d ; θ_u represents the output rate of upstream facility; $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the arrival rate of downlink facility; the p_c^n and p_0^n , respectively, indicate the congestion rate of downstream facility and the

Fig. 1 Chematic diagram of separate networks



probability that there is no pedestrian in the facility; and p_i represents the probability that the pedestrian chooses the downstream facility i .

The probability selection problem of the dividing nodes is modeled as (M1)

$$(M1) \max Z = \sum_{i \in V_d} \theta_i \tag{16}$$

s.t.

$$\sum_{i \in V_d} p_i = 1 \tag{17}$$

$$p_i \cdot \theta_n = \lambda_i, \quad \forall i \in V_d \tag{18}$$

$$\lambda_i \cdot (1 - p_c^i) = \theta_i, \quad \forall i \in V_d \tag{19}$$

$$p_c^i = p_0^i \cdot \prod_{k=1}^c \frac{\lambda_i \cdot E(T_1)}{k \cdot f(k)}, \quad \forall i \in V_d \tag{20}$$

$$p_0^i = \left[1 + \sum_{j=1}^c \prod_{k=1}^j \frac{\lambda_i \cdot E(T_1)}{k \cdot f(k)} \right]^{-1}, \quad \forall i \in V_d \tag{21}$$

$$0 \leq p_i \leq 1, \quad \forall i \in V_d \tag{22}$$

From the model M1, it can be seen that the pedestrian optimal selection probability is related to the upstream facility output rate and the downstream facility congestion rate. The congestion rate is related to the geometric attribute of the facilities and arrival rate, so the optimal probability is the result of the output rate of upstream facilities and the geometric attribute of downstream facilities.

3.3 Model Solution

First, determining p_c and p_0 , for $E(T_1) = \frac{l}{V_1}$ and $f(n) = \frac{V_n}{V_1}$, so:

$$p_0 = \left[1 + \sum_{j=1}^c \prod_{k=1}^j \frac{\lambda_i \cdot E(T_1)}{k \cdot f(k)} \right]^{-1} = \left[1 + \sum_{j=1}^c \lambda^j \cdot \prod_{k=1}^j \frac{l}{j \cdot V_k} \right]^{-1} \tag{23}$$

where $g(j) = \prod_{i=1}^j \frac{l}{i \cdot V_i}$, then:

$$p_0 = \left[1 + \sum_{j=1}^c \lambda^j \cdot g(j) \right]^{-1} \quad (24)$$

Similarly,

$$p_c = \frac{\lambda^c \cdot g(c)}{1 + \sum_{j=1}^c \lambda^j \cdot g(j)} \quad (25)$$

Stage I: Finding out the $g(j)$ value when the number of pedestrians is different, $g(j)$ can be expressed as follows:

$$\begin{cases} g(j) = g(j-1) \cdot \frac{\lambda}{j \cdot V_j}, & j \geq 2 \\ g(j) = \frac{\lambda}{V_1}, & j = 1 \end{cases} \quad (26)$$

Stage II: Model solution.

For Lingo Software can be used to solve linear programming and some linear or nonlinear equations, it is the best choice for solving the optimization model.

4 Case Study

Taking the Jianguomen Subway Station in Beijing as a case, the optimal model for probabilistic selection of dividing node is applied to the actual station, and the simulation method is used to analyze the results.

4.1 The Evacuation Route Planning of Jianguomen Station

Evacuation path uses the platform as a starting point, with the exit at the end; pathways include stairs, escalators, and corridors. As can be seen from Fig. 2, there are 14 evacuation routes, see Table 1 for details.

4.2 M/G/c/c Model Calculation Results Analysis

Select different arrival rates to calculate the relevant parameters of the station facilities.

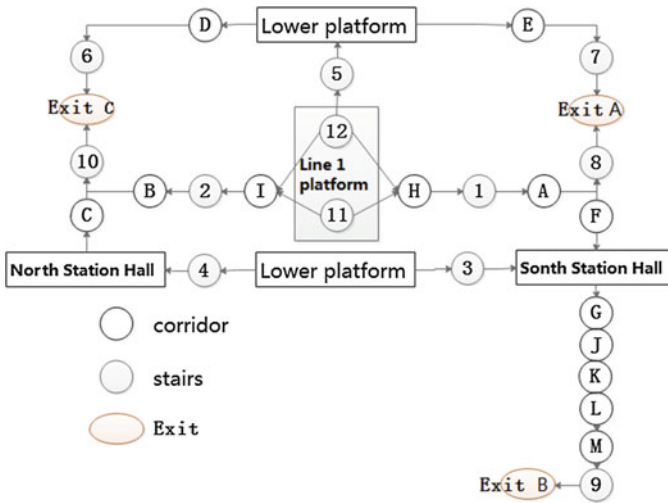


Fig. 2 The topology of Jianguomen station

4.2.1 Number of Evacuation

Based on the model, we can get the output rate of each exit at different arrival rates. The results are shown in Table 2, the proportion of evacuated persons at C, B, and A exits is 0.431: 0.306: 0.263, which is consistent with the actual.

When the arrival rate is 13, as the number of people waiting in the station is too crowded, the total output of exits will begin to decrease. The definition of evacuation time is 5 min (excluding response time), and the maximum evacuation number is 4103.

4.2.2 Evacuation Time

The evacuation time of each path is as Table 3, which is the total expected service time of the included facilities. From Table 1, the evacuation time is not strictly related to the path length. When the length is shorter, the path will attract more pedestrians and lead to congestion. Instead, the evacuation time will increase. As $\lambda = 13$, the maximum evacuation time is 315 s (including reaction time).

4.2.3 Passenger Density

The average density of facilities is got from the expected service number and the area of travel facilities. As $\lambda = 13.0$, Jianguomen stairs evacuation capabilities are in the three-level or four-level, and the channels are almost two-level, part of the three [14]

Table 1 Evacuation paths of Jianguomen station

Path number	The number of stairs/escalators	The number of corridors	The narrowest width of the path	Total length of the path	Path number	The number of stairs/escalators	The number of corridors	The narrowest width of the path	Total length of the path
Path 1	2	1	2.8	86.67	Path 8	2	1	1	89.35
Path 2	2	1	1	83.29	Path 9	3	2	2.8	101.46
Path 3	2	5	3.2	110.9	Path 10	3	2	1	98.08
Path 4	2	5	1	105.15	Path 11	3	2	3	105.55
Path 5	2	1	2.8	96.92	Path 12	3	2	1	97.87
Path 6	2	1	1	90.44	Path 13	3	8	3.8	225.97
Path 7	2	1	2.8	97.33	Path 14	3	8	1	220.22

Table 2 Output rate of exits

Single station arrival rate	Output rate				Total output/ 5 min	Evacuation efficiency (%)
	Exit A	Exit B	Exit C	Total		
4	1.631	2.475	3.219	7.325	2198	91.60
5	1.902	2.52	3.717	8.139	2442	81.40
6	2.429	2.583	4.449	9.461	2838	78.80
7	2.567	2.868	4.536	9.971	2991	71.20
8	2.749	3.157	4.611	10.517	3155	65.70
9	2.974	3.446	4.691	11.111	3333	61.70
10	2.941	3.666	5.386	11.993	3598	60.00
11	3.235	3.96	5.592	12.787	3836	58.10
12	3.522	4.231	5.799	13.552	4066	56.50
13	3.602	4.183	5.892	13.677	4103	52.60
13.5	3.513	4.144	5.815	13.472	4042	49.90
14	3.497	4.127	5.804	13.428	4028	48.00

Table 3 The service time of each path

Path	Evacuation ($\lambda = 13.0$)	Path	Evacuation time ($\lambda = 13.0$)
Path 1	242.855	Path 8	135.033
Path 2	237.297	Path 9	253.363
Path 3	177.686	Path 10	247.804
Path 4	200.150	Path 11	131.973
Path 5	142.951	path 12	131.568
Path 6	135.596	Path 13	232.286
Path 7	142.388	Path 14	154.750

It can be seen from Table 4 that when the arrival rate is 13.0, the probability of congestion in stairs 3, 4, 5, 10, 11, 12, escalator 10 and channel G appears, and the evacuation levels of these congestion facilities are all four.

4.3 Comparison of Model Results with Simulation Results

Select the Pathfinder simulation, verify the model calculation results from the evacuation number, evacuation time and passenger density.

In Pathfinder simulation, the number of evacuations was 3596 in 6 min (including reaction time). In contrast, the model results are larger, for the construction model needs to make some assumptions and dilute some of the constraints. In the simulation results, the proportion of evacuated population at C, B, and A was 0.409: 0.319: 0.271, which was consistent with the calculated results.

Table 4 Summary of calculation results

Facilities	Congestion rate	Output rate (per/s)	Expected service number	Expected service time/s	Facilities	Congestion rate	Output rate (per/s)	Expected service number	Expected service time/s
Stair 1	0.000	3.459	95.923	27.728	Escalator 9	0.000	0.724	47.676	65.869
Stair 2	0.000	3.459	97.856	28.287	Escalator 10	0.553	0.864	136.198	157.657
Stair 3	0.347	4.243	228.144	53.774	Corridor A	0.000	3.459	13.739	3.972
Stair 4	0.401	3.892	210.524	54.095	Corridor B	0.000	3.459	13.739	3.972
Stair 5	0.042	5.129	221.897	43.263	Corridor C	0.000	3.892	41.852	10.754
Stair 6	0.000	1.890	139.041	73.580	Corridor D	0.000	2.565	66.954	26.108
Stair 7	0.000	1.890	139.041	73.580	Corridor E	0.000	2.565	65.511	25.545
Stair 8	0.000	0.778	33.078	42.497	Corridor F	0.000	2.422	26.111	10.783
Stair 9	0.000	3.459	150.156	43.405	Corridor G	0.544	3.040	91.134	29.979
Stair 10	0.545	2.464	402.169	163.215	Corridor H	0.000	3.459	82.459	23.836
Stair 11	0.095	3.459	117.416	33.941	Corridor I	0.000	3.459	82.847	23.948
Stair 12	0.095	3.459	117.416	33.941	Corridor J	0.000	3.040	66.559	21.895
Escalator 6	0.000	0.675	44.694	66.225	Corridor K	0.000	3.040	24.670	8.115
Escalator 7	0.000	0.675	44.694	66.225	Corridor L	0.000	3.040	34.301	11.283
Escalator 8	0.000	0.259	10.921	42.091	Corridor M	0.000	3.040	52.745	17.350

Table 5 Maximum density of various regions in the station

Facility number	Maximum density	Evacuation level	Facility number	Maximum density	Evacuation level
Stair 1	1.139	III	Escalator 9	1.722	III
Stair 2	1.185	III	Escalator 10	4.799	IV
Stair 3	4.353	IV	Corridor A	0.687	II
Stair 4	3.871	IV	Corridor B	0.462	II
Stair 5	3.860	IV	Corridor C	0.627	II
Stair 6	1.113	III	Corridor D	0.345	II
Stair 7	1.287	III	Corridor E	0.339	II
Stair 8	0.284	I	Corridor F	0.392	II
Stair 9	2.334	III	Corridor G	3.768	IV
Stair 10	3.281	IV	Corridor H	0.644	II
Stair 11	4.332	IV	Corridor I	0.457	II
Stair 12	4.552	IV	Corridor J	1.497	III
Escalator 6	1.378	III	Corridor K	0.217	I
Escalator 7	1.195	III	Corridor L	0.740	I
Escalator 8	0.367	I	Corridor M	0.808	III

The time for the pedestrian to evacuate the station is 365 s (including the reaction time) in the simulation, and the model result is 315 s. It is because the model only considers the running time within the facility and does not consider the facility’s connection.

The maximum passenger density obtained from the simulation is as Table 5. From Table 5 the stairs 3, 4, 5, 10, 11, 12, the escalator 10 and the channel G have a higher density, and the results are in good agreement with the model results. Similarly, these facilities will also become the evacuation bottlenecks.

To sum up, through the simulation, we can prove the effectiveness of the queuing model.

5 Conclusion

Based on the queuing model, the evacuation capacity of the subway is evaluated, and the evacuation capacity indexes such as evacuation time, number of people and average density are obtained, and the evacuation bottleneck is simulated effectively. In addition, Pathfinder software is used to validate the model results, which provides a simulation basis for improving the facility evacuation capacity.

However, there are still some shortcomings in the model, which will continue to be studied in future work:

- (1) Do not consider pedestrian psychology, whether to carry baggage, baggage dimensions, and travel purpose.
- (2) When calculating the evacuation bottleneck, the influence of the toll gate and other facilities on the carrying capacity is not considered.
- (3) The model makes certain assumptions about constraints and environment and has limitations.

Acknowledgements The authors gratefully acknowledge the support provided by China National “13th Five-Year” key research project “Safety assurance technology of urban rail system” (Grant No.2016YFB1200402) and “National Engineering Laboratory for System Safety and Operation Assurance of Urban Rail Transit”.

References

1. Gipps PG, Marksjo B (1985) A micro-simulation model for pedestrian flows. *Math Comput Simul* 27(2):95–105
2. Wolfram S (1986) *Theory and applications of cellular automata*. World Scientific Pub. Co Inc., Hackensack, New Jersey, US
3. Heilbing D, Molnar P (1995) Social force model for pedestrian dynamics. *Phys Rev E* 51 (5):4282–4286
4. Cruz FRB, Smith JMG, Queiroz DC (2005) Service and capacity allocation in M/G/c/c state-dependent queueing networks. *Comput Oper Res* 32(6):1545–1563
5. Kirchner A, Schadschneider A (2002) Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. *Phys A* 312:260–276
6. GB 50157-2013 (2013) *Subway design code*. China planning Publishing House, Beijing (in Chinese)
7. Xu K, Zhai J, Xia Ch et al (2010) Research on evacuation plan of personnel evacuation based on queuing theory. *Secur Technol* 12:35–38 (in Chinese)
8. Zhou K (2013) *Study on transfer behavior analysis and facility allocation of high-speed rail passenger hub*. Unpublished master thesis, Harbin Institute of Technology (in Chinese)
9. Fruin J (1992) *Designing for pedestrians*. Public Transportation United States
10. Zhang L (2007) *Optimization of container terminal equipment configuration based on queuing network theory*. Unpublished master thesis, Tongji University (in Chinese)
11. Cheah JY, Smith JM (1994) Generalized M/G/c/c state dependent queueing models and pedestrian traffic flows. *Queueing Syst* 15(1–4):365–386
12. Tregenza P (1976) *The design of interior circulation*. Van Nostrand Reinhold
13. Predtechenskii VM, Milinskii AI (1978) *Planning for foot traffic flow in buildings*. National Bureau of Standards, US Department of Commerce, and the National Science Foundation, Washington, DC
14. Chen Q, Xie X (2007) Simulation of pedestrian dissipative simulation of large—scale activities. In: *The 3rd China intelligent transportation annual meeting*, Nanjing, pp 215–222 (in Chinese)

Research on the Abrasion Detection Technology of the Pantograph Slipper of Urban Rail Train

Xuebing Hu, Yejian Chen, Yong Zhang and Zongyi Xing

Abstract In order to realize the online noncontact detection of the Abrasion of the Pantograph Slipper of Urban Rail Train, the abrasion detection technology of pantograph slipper of urban rail train is studied. In this paper, an abrasion detection method of pantograph slipper of urban rail train based on image processing is proposed, and the steps are as follows: First, the mixed noise of the original image from the acquisition system is filtered by image filtering. Then, the edge of Pantograph Slipper in the image is detected by the adaptive Canny edge detection based on histogram concavity analysis after image filter. Finally, the abrasion curve of the Pantograph Slipper is obtained by camera calibration and curve blending, and which is used to judge whether the abrasion of the pantograph slipper is transfinite. The results of the experiment showed that the proposed method is effective to detect the pantograph slipper abrasion curve in the original image. The proposed method is effective to realize the online non-contact detection of the Abrasion of the Pantograph Slipper of Urban Rail Train.

Keywords Pantograph slipper · Image filter · Adaptive Canny edge Detection · Camera calibration

1 Introduction to Image Processing Technology

The image is the main source of obtaining and exchanging information, digital image processing technology has the advantages of good reproducibility, high processing accuracy, high flexibility, and so on. Therefore, it is applied to all aspects of human life and work, such as aerospace, aviation technology, communication engineering, etc. This paper applies image processing technology to railway safety transportation. On the test site, four industrial cameras based on CCD

X. Hu (✉) · Y. Chen · Y. Zhang (✉) · Z. Xing
School of Automation, Nanjing University of Science and Technology,
Nanjing 210094, China
e-mail: 34445721@qq.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_33

333

(Charge-coupled Device) sensitive chips are installed, each two of which is used to shoot a pantograph on the roof of the train. The slider image captured by the industrial camera is transmitted to the image processing system through Gigabit Ethernet to detect the abrasion of pantograph slider after being processed by filtering, edge detection, camera calibration, and other image processing techniques.

2 The Algorithm on the Abrasion Detection Technology of the Pantograph Slider

In this paper, the pantograph slider abrasion is analyzed mainly by the half bow image taken by two CCD cameras, and the digital image processing technology is used to extract the upper edge of pantograph slider and the lower edge of the bracket. The distance between the upper edge of the slider and the lower edge of the bracket is calculated by camera calibration, and the residual abrasion curve of the half bow slider is obtained by subtracting the thickness of the bracket. The curves of the two semi-bow slider curves are fused, and the residual abrasion curve of the whole slider is obtained. The detection process of pantograph slider abrasion is shown in Fig. 1.

2.1 Image Filtering

Due to the influence of the external environment, there are a lot of mixed noises in the original image collected by camera, mainly pulse and Gauss noise [1]. This paper proposes a filtering algorithm to remove mixed noise first, median filter for impulse noise suppression, then using the mean filter algorithm to restrain the Gauss noise, the two filtering algorithms works step by step to filter the mixed noise in the pantograph image.

2.2 Adaptive Canny Edge Detection

Adaptive Canny edge detection algorithm is proposed to achieve an automatic detection of edge with high- or low-threshold settings by the image gradient histogram concavity analysis [2]. The specific steps of the adaptive Canny edge detection algorithm are as follows:

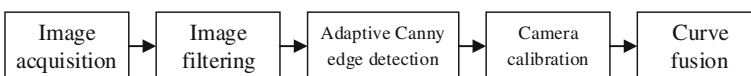


Fig. 1 Flow of abrasion detection

2.2.1 Gradient Calculation

In the image after filtering, the gradient of the point is calculated by using 3*3 convolution module in the pixel point and its 8 neighborhood. The image pixels (i, j) is defined in horizontal and vertical directions as follows [3]:

$$f_x(i, j) = f(i+1, j-1) + 2f(i+1, j) + f(i+1, j+1) - f(i-1, j-1) - 2f(i-1, j) - f(i-1, j+1), \quad (1)$$

$$f_y(i, j) = f(i-1, j+1) + 2f(i, j+1) + f(i+1, j+1) - f(i-1, j-1) - 2f(i, j-1) - f(i+1, j-1). \quad (2)$$

The gradient magnitude $M(i, j)$ is defined as

$$M(i, j) = \sqrt{f_x(i, j)^2 + f_y(i, j)^2} \quad (3)$$

The gradient direction is defined as

$$\theta(i, j) = \arctan(P_x(i, j)/P_y(i, j)) \quad (4)$$

The gradient magnitude $M(i, j)$ reflects the edge intensity of image pixels (i, j) , and $\theta(i, j)$ is the normal vector of pixels (i, j) on the image, which is perpendicular to the direction of the edge.

2.2.2 High-Low Threshold Selection Based on the Histogram Concavity Analysis

The gradient magnitude histogram concavity analysis is used to select the high and low threshold. The selection process is as follows [4]:

- (1) The 256-gradient magnitude histogram $h(i)$ is extracted, and the nonzero starting point $(i_{start}, h(i_{start}))$ and nonzero terminal $(i_{end}, h(i_{end}))$ of histogram $h(i)$ are extracted.
- (2) Calculate the slope of the gradient magnitude from i_{start} to i_{end} in the histogram according to Eq. (5).

$$s(i) = \frac{h(i) - h(i_{start})}{i - i_{start}}, \quad i_{start} < i < i_{end} \quad (5)$$

Among them, $s(i)$ is the slope, and i is the gradient magnitude.

- (3) The maximum value of the slope $s(i)$ is extracted and the corresponding gradient magnitude is i_{peak} . $(i_{peak}, h(i_{peak}))$ is a bump of the gradient histogram, if $i_{peak} < i_{end}$ makes $i_{start} = i_{peak}$ and turns to step 2, if $i_{peak} = i_{end}$, step 4).
- (4) The N convex points of the pantograph image histogram can be obtained by the above steps $(i_{p1}, h(i_{p1}))(i_{p2}, h(i_{p2})) \dots (i_{pn}, h(i_{pn}))$, a minimum convex polygon of the envelope histogram $(\bar{h}(i))$ is obtained by connecting the start point $(i_{start}, h(i_{start}))$, the destination point $(i_{end}, h(i_{end}))$, and the N convex points in a straight line.
- (5) Calculate the concave residuals of the histogram $c(i) = \bar{h}(i) - h(i)$. When the concave residuals take the maximum value, the corresponding gradient magnitude is the high threshold T_h , the lower threshold $T_l = 0.4T_h$.

2.2.3 Marginal Point Judgment

Judgment method of edge point is according to the gradient magnitude feature image, if the gradient magnitude is larger than the T_h , it is the edge point; if the gradient is less than points T_l , the point is not the edge point. If the gradient points in T_l and T_h , the possible edge points need further judgment, the magnitude is judged according to the gradient of the adjacent point, if there exists a gradient magnitude greater than T_h points, the point is the edge point, the edge is not otherwise [5, 6].

2.2.4 Edge Join

For the image $N(i, j)$ after non-maximum suppression, the edge images $T_h(i, j)$ and $T_l(i, j)$ are obtained after high- and low-threshold detection respectively. Image $T_h(i, j)$ is an image obtained by high-threshold processing; $T_l(i, j)$ is an image obtained by low-threshold detection. According to the connectivity of edges, based on image $T_h(i, j)$, the possible edge points are searched in image $T_l(i, j)$ for edge linking, and the final edge image is obtained.

Similarly, continue searching in image $T_h(i, j)$ to trace the contour line with point R as the starting point. Repeat the above steps and cycle until the points that can be joined to this contour are not found in images $T_h(i, j)$ and $T_l(i, j)$. Thus, the contour line of the included point P is completed. Then follow each of the lines in $T_h(i, j)$ until the new contour is not found in $T_h(i, j)$.

In this paper, an objective evaluation index is used to evaluate the edge detection algorithm, the objective evaluation index mainly includes the reconstruction of similarity (*MSSIM*) and continuity index (*CI_{dx}*), and the weighted value as the final objective evaluation index *EIdx* of the edge detection, as shown in Eq. (6). The larger the value of *EIdx* is, the better the image edge detection is, the better the detection performance is.

$$EIdx = \omega_1 \times MSSIM + \omega_2 \times CIdx. \tag{6}$$

Among them, ω_1 and ω_2 are the weights of reconstruction similarity and continuity index, respectively, and $\omega_1 + \omega_2 = 1$.

2.3 Camera Calibration

After the edge image is obtained by adaptive Canny edge detection, according to the camera internal and external parameters with camera calibration, the distance between the upper edge of the pantograph slider and the lower edge of the slider is converted into a distance value in the real-world coordinate system, this distance reduces the thickness of the slider and gets the pantograph slider curve [6]. Because the Zhang Zhengyou calibration method has good robustness, and the calibration board requirements is not high, with the use of strong, so this paper chooses the Zhang Zhengyou calibration method for camera calibration.

2.4 Curve Fusion

In this paper, two groups of cameras are used to collect the left and right half bow image of pantograph, Therefore, in order to obtain the residual thickness curve of the whole bow slider, the residual abrasion curve of the half bow of the two groups of cameras should be fused. Suppose that the coordinate system of the left half residual abrasion curve is $o_l - x_l y_l$, and the coordinate system of the right half residual abrasion curve is $o_r - x_r y_r$. According to the characteristics of the slide, the two coordinate systems are fused, and the coordinate system after fusion is $o - xy$ [7].

The data points of the left and right half bow slider abrasion are fused according to Eq. (7), and the data of the slider abrasion of the coordinate system $o_l - x_l y_l$ and $o_r - x_r y_r$ are fused into the coordinate system $o - xy$ [8], so the complete residual contour abrasion curve of the slider is obtained.

$$\begin{cases} x = x_l, y = y_l \\ x = x_r + \Delta x, y = y_r + \Delta y \end{cases} \tag{7}$$

In Eq. (7), (x, y) is the point in the fusion coordinate system, the origin of coordinate of the fusion coordinate system $o - xy$ coincides with the origin of coordinate of the left half bow coordinate system $o_l - x_l y_l$ [9], Therefore, the offset of the left half bow data relative to the origin of coordinate of the fusion coordinate system is zero, the offset of the coordinate origin of the right half bow coordinate system $o_r - x_r y_r$ in the X axis and the Y axis relative to the coordinate system $o - xy$ are Δx and Δy .

3 Experiment and Analysis

This paper presents an adaptive Canny edge detection algorithm that is better than the traditional Canny algorithm, and then applies it to the abrasion detection method of the pantograph slider, and the pantograph image is processed to verify the effectiveness of abrasion detection technique of the pantograph slider proposed in this paper.

3.1 Analysis of Adaptive Canny Edge Detection

Based on the MATLAB platform, the detection performance of the traditional Canny detection algorithm and the adaptive Canny edge detection algorithm are compared and analyzed. Because the edge detection effect is related to the structure of the object under test [10], the edge detection results are showed in Fig. 2

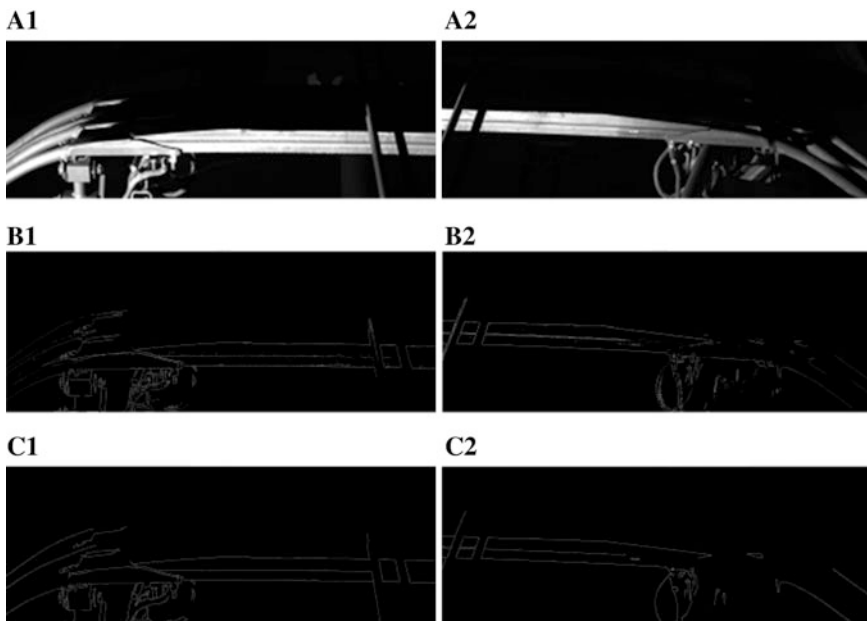


Fig. 2 Edge detection results, (A1) original image of left half bow, (A2) original image of right half bow, (B1) image of the detection result of the left half bow using the traditional Canny edge algorithm, (B2) image of the detection result of the right half bow using the traditional Canny edge algorithm, (C1) image of the detection result of the left half bow using the edge algorithm proposed in this paper, (C2) image of the detection result of the right half bow using the edge algorithm proposed in this paper

Table 1 Objective index of edge detection

Objective index	<i>MSSIN</i>		<i>CIdx</i>		<i>EIdx</i>	
	Traditional Canny	Adaptive Canny	Traditional Canny	Adaptive Canny	Traditional Canny	Adaptive Canny
Left half of bow	0.709	0.812	0.718	0.736	0.712	0.782
Right half of bow	0.721	0.792	0.693	0.720	0.710	0.763

The traditional Canny algorithm edge detection and the adaptive Canny edge detection proposed in this paper are calculated, respectively, by objective indexes, and the results are shown in Table 1.

As shown in Table 1, the values of the *MSSIM*, *CIdx*, and *EIdx* in the images of left and right half bow of pantograph obtained by the edge algorithm proposed in this paper are all larger than the values of the *MSSIM*, *CIdx*, and *EIdx* in the images of left and right half bow of pantograph obtained by the traditional Canny edge algorithm. This shows that the edge detection algorithm proposed in this paper is superior to the traditional Canny edge algorithm.

3.2 Realization of Abrasion Detection Algorithm of Pantograph Slider

In this paper, four CCD high-definition cameras are used to shoot the front slider and the rear slider. Two cameras for a group, a total of two groups, these cameras are installed in the symmetrical running direction of the train, with a certain angle shooting left and right bow of pantograph. The slider flashlamps install on both sides of the track, to fill the pantograph in elevation and make the side surface of pantograph slider lighter, comparing to the upper surface and back of the slider, the upper and lower edges of the slide can be prominent as shown in Fig. 3.

Locate the slider and slider bracket based on the contour of the pantograph edge. The thickness curves of the slider of the left and right half bow and its bracket are obtained by camera calibration. According to the definition of the slider abrasion, minus the thickness of the slider bracket, the curve of the slider residual abrasion is obtained, as shown in Figs. 4 and 5.

The curves of the left half and right half bow are fused by curve fusion algorithm, and the fusion curve is obtained. For ease of analysis, the lower edge of the slide is taken as the coordinate axis, as shown in Fig. 6.

The residual abrasion curve of the slider was statistically calculated, and the minimum surplus of the slide plate was got, so as to get the maximum surplus of the slider as shown in Fig. 6.

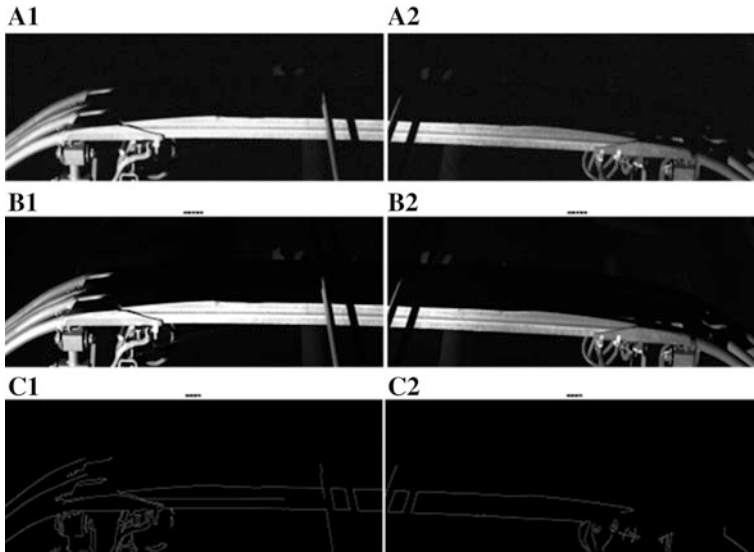


Fig. 3 Edge detection process, (A1) original image of left half bow, (A2) original image of right half bow, (B1) filtered image of left half bow, (B2) filtered image of right half bow, (C1) edge extraction results of left half bow, (C2) edge extraction results of left half bow

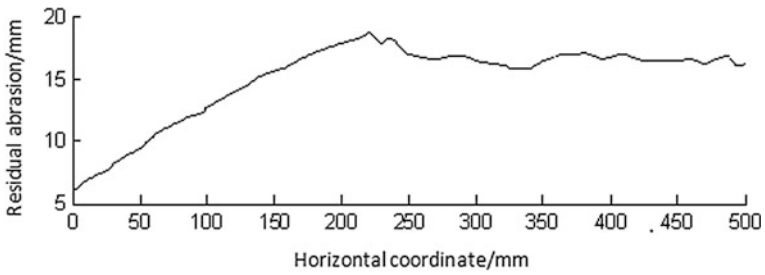


Fig. 4 Residual abrasion curve of left half bow

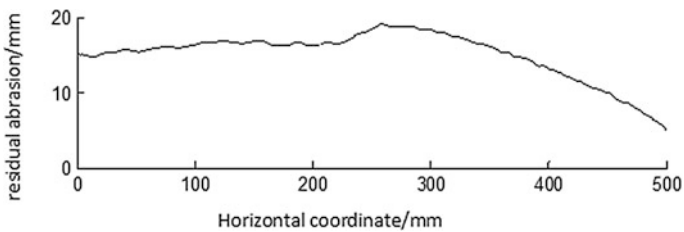


Fig. 5 Residual abrasion curve of right half bow

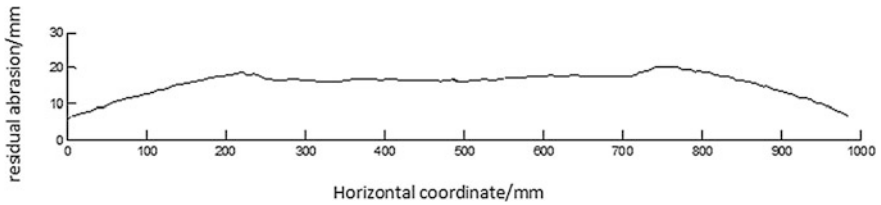


Fig. 6 Residual abrasion curve of pantograph slider

4 Conclusion

In this paper, the abrasion detection of the pantograph slider is studied, and a pantograph abrasion detection method is proposed. The original image acquired by the acquisition system is filtered to reduce the mixed noise caused by electromagnetic interference. The edge of the slider is detected by the adaptive Canny edge detection algorithm. The curve of pantograph's residual abrasion is achieved by camera calibration and curve fusion algorithm. The detection result shows that the detection technology of pantograph slider proposed in this paper can effectively draw the abrasion curve of pantograph slider, and verify the effectiveness of this method. It is of great significance in practical application.

Acknowledgments This work is supported by the National Key Research and Development Project (2017YFB1201201).

References

1. Xin W (2011) Study on arc discharge mechanism of pantograph/catenary system. Southwest Jiao Tong University (in Chinese)
2. Fenghui S, Boming W (2011) Ultrasonic testing method for pantograph wear of double slide pantograph. *Mech Electr Prod Dev Innov* 3:129–131 (in Chinese)
3. Fenghui S (2011) Design and debugging of pantograph skateboard wear system. Southwest Jiao Tong University (in Chinese)
4. Chengfen J, Boming W, Fenghui S (2010) Image design of pantograph wear inspection system based on OpenGL. *Mech Eng Autom* 6:21–22, 25 (in Chinese)
5. Anzhi Y, Zhongming Z, Chengyi W (2013) Active shape model of the pantograph automatic detection method of China stereology and image analysis. Based on the society conference on Chinese stereology society thirteenth, 7 (in Chinese)
6. Xuerong F, Siyang L, Ning L et al (2012) A pantograph measurement system based on laser ranging. *China Electron Commer* 9:39–40 (in Chinese)
7. Aydin I (2015) A new approach based on firefly algorithm for vision-based railway overhead inspection system. *Measurement* 74:43–55

8. Judek S, Jarzebowicz L (2014) Algorithm for automatic wear estimation of railway contact strips based on 3D scanning results. International conference and exposition on electrical and power engineering. IEEE, 724–729
9. Li M, Zeyong W, Xiaorong G, Li W, Kai Y (2010) Detection of residual wear of pantograph. LV2600 Railway Rolling Stock 3:93–97 (in Chinese)
10. Na C (2011) Survey of edge detection algorithms in image processing. China Sci Technol Inf 4:130–131, 149 (in Chinese)

Research on Fault Diagnosis of Vehicle Equipment for High-speed Railway Based on Case-Based Reasoning

Lixuan Chen, Dongxiu Ou and Hongjing Yao

Abstract Vehicle equipment plays an important role in high-speed railway train control system for safety protection and efficiency improving. The role of vehicle equipment for high-speed railway can be affected by a number of uncertain faults and these faults have characteristics of diversity, concealment, and overlap. To solve the above problems, this paper proposes a new algorithm to detect faults of vehicle equipment based on Case-based Reasoning. A database is created based on previous successful cases and features of these cases are extracted. Then fault causes of vehicle equipment are derived by matching the information of target cases to cases in the database. Then maintenance suggestions are given according to fault causes of vehicle equipment. The test result shows that the algorithm can diagnose faults of vehicle equipment for high-speed railway accurately and effectively.

Keywords Railway vehicle equipment · Fault diagnosis · Case-based reasoning
Similarity match

1 Introduction

As core equipment of train control system, vehicle equipment plays an important role in protection of train operation [1]. Faults of vehicle equipment may threaten the safety of train operation. Therefore, it is an important job for the technical department of Railway Administration to diagnose faults of vehicle equipment timely and accurately. At present, fault diagnosis of vehicle equipment mainly relies on manual experience. Maintenance staff analyzes phenomena of faults and log files of vehicle equipment to judge causes of faults [2]. The cycle of diagnosis is long

L. Chen · D. Ou (✉)

Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai, China
e-mail: ou.dongxiu@tongji.edu.cn

H. Yao

Electricity Section, Jinan Railway Administration, Jinan, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_34

343

and the efficiency is low. Moreover, it may affect the safety of train when personnel experience is insufficient.

In order to solve the above problems, some scholars have studied different intelligent methods of fault diagnosis for vehicle equipment. Fault tree was proposed to detect faults in the early time [3]. Bayesian network was adopted by some researchers to calculate the probability of failures and analyze the safety of vehicle equipment [4]. Other studies introduced some other methods to the field, such as neural network, support vector machine, and so on [5, 6].

Currently most studies on fault diagnosis of vehicle equipment for high-speed railway concentrate on data of fault records. The fault records are written by maintenance staff in the technical department of Railway Administration and it includes fault description, fault causes, maintenance conditions, and other information [4, 5]. In 2013, some researchers using one intelligent algorithm to diagnose faults of vehicle equipment for high-speed railway with data of fault records. The results have shown that the accuracy is 92.4% on types of fault and 67.2% on specific fault [4].

This paper analyzes fault modes and maintenance status of vehicle equipment for high-speed railway. Features of fault data are extracted and Case-based Reasoning model is introduced to search for previous successful cases from the database created. Thus, we can diagnose fault causes of vehicle equipment quickly and give maintenance suggestions accurately.

2 Composition and Fault Classification of Vehicle Equipment

The main functions of vehicle equipment for high-speed railway are to generate dynamic speed curve and to monitor train according to some parameters. A type of train called CTCS3-300T is the most widely used in the train control system for Chinese high-speed railway. Vehicle equipment in this kind of train is mainly composed of ATPCU (Automatic Train Protection Control Unit), C2CU (Chinese Train Control System-2 Control Unit), SDP (Speed and Distance Processing), SDU (Speed and Distance Unit), VDX (Digital Input and Output Unit), STU-V (Safe Transmission Unit-Vehicle), GSM-R (Global System for Mobile Communications for Railway), TCR (Track Circuit Reader), BTM (Balise Transmission Module), CAU (Compact Antenna Unit), DMI (Driver-machine Interface), JRU (Judicial Record Unit), GCD (General Crypt Device), and so on [7]. The structure of vehicle equipment is shown in Fig. 1.

ATPCU and C2CU are used as control units and responsible for core functions during normal operation. SDP receives original information of pulse from SDU and gets speed and distance of train after processing the information. BTM and CAU are responsible for receiving information of balise. TCR is responsible for receiving information of track circuit. VDX is used to input and output safety related signals.

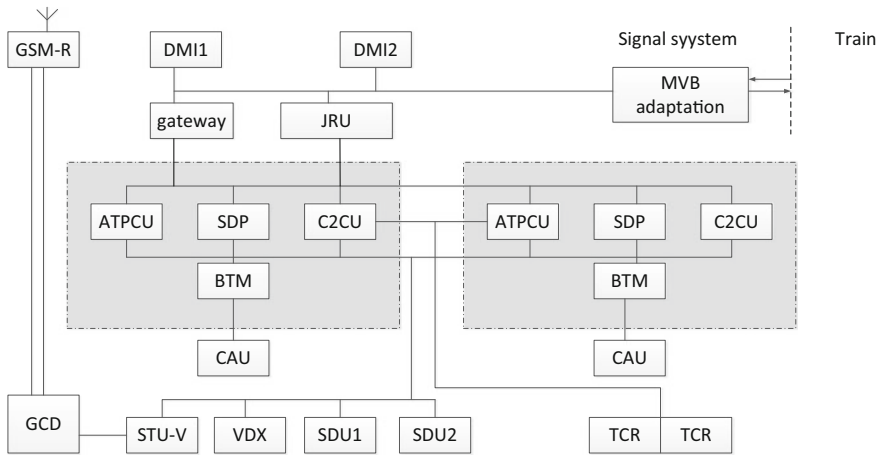


Fig. 1 Structure of vehicle equipment for high-speed railway

GSM-R is used for receiving and coping with information from RBC. GSM-R is connected with GCD which interacts data with vehicle equipment through STU-V.

The failures of vehicle equipment come from uncertain factors associated with software failures, hardware failures, human errors, and environmental conditions. According to function modules where fault occurs, faults can be divided into the following 9 types, including fault of ATPCU, fault of BTM, fault of C2CU, fault of DMI, fault of interfaces, fault of speed detection, fault of TCR, fault of wireless timeout, and other faults. According to demands of maintenance, the 9 types of faults are further subdivided. For example, fault of BTM can be subdivided into “zero—responder”, “BSA permanent error in operation”, “BSA permanent error at boot” and “port status invalid”. Fault classification of vehicle equipment for high-speed railway is partly shown in Fig. 2.

3 Database Creation

3.1 Original Fault Data of Vehicle Equipment

Original fault data of vehicle equipment mainly includes phenomenon of faults, causes of faults and suggestions for maintenance. Among them, phenomenon of faults is divided into two categories according to their sources. One is DMI alarm texts which are recorded by JRU and can be seen in the user interface by drivers or operators. It contains some alarm information, such as “Emergency braking”, “Speed sensor failure”, and so on. The other is the logs of some vehicle equipment including ATPCU, C2CU, and SDP. These logs are symbolic statements or error codes written by computer when a particular or important event occurs and need to

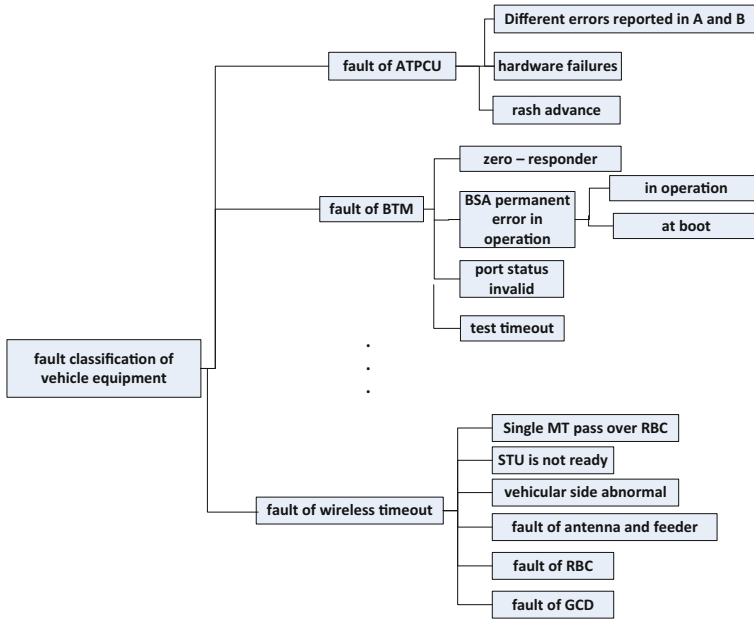


Fig. 2 Fault classification of vehicle equipment for high-speed railway

be downloaded from the train. It is recorded by textual language. One piece of record contains current date and time, device name and number, error codes, and event description. Causes of faults include fault of ATPCU, fault of BTM, and so on, as shown in Fig. 2. Suggestions for maintenance mainly include software upgrade, hardware replacement, unit test, and so on.

3.2 Feature Extraction of Fault Phenomenon Data

Fault phenomenon data is numerous and descriptions of data are textual language. Extracting feature words from fault phenomena data is helpful for computer to understand and recognize.

As for fault phenomenon data from DMI alarm texts, it is recorded in the following format. For example, “A train with ID 6502 had converted from C3 mode to C2 mode” is recorded and shown in the user interface. Then we can extract the key word “Level conversion”.

As for fault phenomenon data from logs of some vehicle equipment, it is recorded in the following format. For example, “12-05-16 16:16:02 116 LID: 246 TID:SMGM_LogTask 000032071; FID:btmc_status_ ATP A 4E2FS081 StatusPort invalid in BTM1” is recorded by ATPCU log when a fault occurs. Then we can extract the key word “StatusPort invalid”.

According to data sources, fault phenomenon data can be divided into 4 categories, including DMI alarm text, logs of ATPCU, logs of SDP, and logs of C2CU. Each category consists of multiple keywords as features. Features of fault phenomena are partly shown in Table 1.

3.3 Description of Single Case

Each case of vehicle equipment failure mainly consists of three parts, including phenomenon, cause, and maintenance suggestion. The cause of fault is divided into 9 types and 30 specific faults, as shown in Fig. 2. The phenomenon of fault is divided into 4 categories and each category includes features which is described in Table 1. The maintenance suggestion of fault is a further analysis of the failure and it is decided by causes. Some other information including ID number of case, time of case, and source of case is added to the description of single case as auxiliary information. The description of single case is shown in Table 2.

Table 1 Features of fault phenomena data

Category	Feature
DMI alarm text	Communication interrupts with DMI, emergency braking, ground equipment failure, speed sensor fault, level conversion...
ATPCU log	Different errors reported in A and B, BSA permanent error, BSA temporary error...
SDP log	BT15XXXX, BT26XXXX, BT26XXXX, BTM1 status message not valid...
C2CU log	Start-up NVMEM or PAMEM test not properly executed...

Table 2 Description of single case

Attribute	Value
Cause	Types of fault
	Specific fault
Phenomenon	DMI alarm text
	ATPCU log
	SDP log
	C2CU log
Maintenance suggestion	Maintenance suggestion
Other	Source of case
	Time of case
	ID number of case

3.4 Organization of Database

In order to improve efficiency and accuracy of fault diagnosis, cases are organized according to causes of fault, which are shown in Fig. 2. Each specific fault contains at least one typical case and the description of each typical case is shown in Table 2.

4 Case-based Reasoning Model for Fault Diagnosis

Case-based Reasoning model is a method for finding similar and successful “past” cases [8, 9]. In this paper, Case-based Reasoning model is introduced to search for the most similar case from database containing successful cases.

4.1 Similarity Calculation

4.1.1 Local Similarity Calculation

According to data sources, fault phenomenon data can be divided into 4 categories, including DMI alarm text, logs of ATPCU, logs of SDP, and logs of C2CU. Local similarity represents the similarity of features in one fault phenomenon category. We calculate local similarity as follows.

If we take fault classification of vehicle equipment as a tree, each bottom node in the tree represents a specific fault. If one feature in one fault phenomenon category may appear when some specific fault occurs, we add this feature to the tree to form new underlying nodes belonging to this specific fault. For example, if we assume feature 1–5 belongs to the same fault phenomenon category, we create a tree for calculating the local similarity. The tree is shown in Fig. 3.

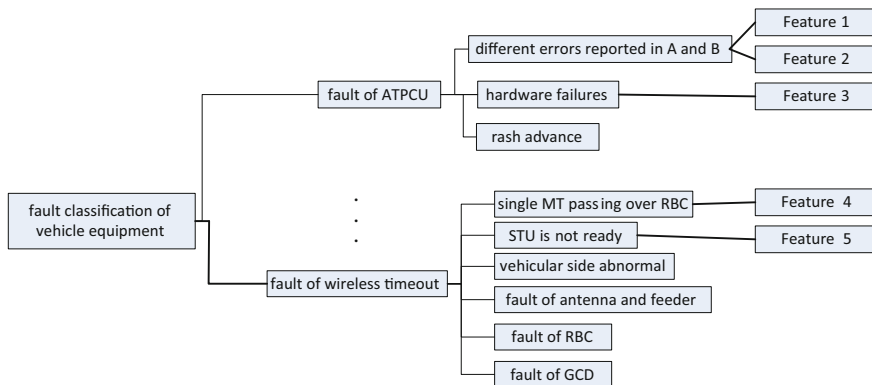


Fig. 3 Tree for calculation of local similarity

Define the calculation of local similarity as Formula (1) which is as follows:

$$s(i_m, j_m) = \frac{d(\minl(i_m, j_m))}{d(\text{tree})}. \tag{1}$$

where

- $s(i_m, j_m)$ the local similarity between feature i and feature j in the m th fault phenomenon category
- $\minl(i_m, j_m)$ the position of the smallest common ancestor of feature i and feature j in the tree
- $d()$ the depth of the position in the tree
- $d(\text{tree})$ the depth of the tree.

If we use Formula (1) to calculate local similarity in Fig. 3, the results of local similarity calculation are shown in Table 3.

4.1.2 Global Similarity Calculation

We calculate global similarity as follows. The first work is to calculate weight of fault phenomenon category. The second work is to calculate global similarity by combing local similarity and weight of fault phenomenon category.

If we remove one fault phenomenon category from a case, feathers belonging to the case will be the same as feathers belonging to another case which has different fault cause. Then we think that the case may be diagnosed wrongly. Define the calculation of weight of fault phenomenon category as Formula (2) according to data in the database.

$$w_i = \frac{F_{(1,2,\dots,i-1,i+1,\dots,n)}}{\sum_{i=1}^{i=n} F_{(1,2,\dots,i-1,i+1,\dots,n)}} \tag{2}$$

where

- $F_{(1,2,\dots,i-1,i+1,\dots,n)}$ the number of cases which may be diagnosed wrongly after removing the i th fault phenomenon category.

Global similarity represents the similarity between cases. Define the calculation of global similarity as Formula (3) which is as follows:

Table 3 Calculation results of local similarity

Feature	Local similarity
Feature 1, Feature 2	2/3
Feature 1, Feature 3	1/3
Feature 1, Feature 4	0

$$s(A, B) = \frac{\sum_{m=1}^n \{w_m \times \text{MAX}[s(A_m, B_m)]\}}{\sum_{m=1}^n w_m} \quad (3)$$

where

$s(A, B)$ the global similarity between case A and case B
 w_m the weight of the m th fault phenomenon category
 $\text{MAX}[s(A_m, B_m)]$ the maximal local similarity in the m th fault phenomenon category between features which belongs to case A and features which belongs to case B.

4.2 Case-Based Reasoning Procedures

When information of a target case is input, the Case-based Reasoning procedures are as follows.

Step 1: Information of the target case is input. In the information of target case, fault time, DMI alarm texts and logs of some vehicle equipment are the key data we need. DMI alarm texts and logs of some vehicle equipment are dealt with and features are extracted from them.

Step 2: The target case is matched with cases in the database. Similarity between the target case and cases in the database is calculated and the most similar case is chosen.

Step 3: The case which has highest similarity with the target case is selected as the final diagnosis. The fault cause and maintenance suggestions of the target case are given according to the most similar case.

5 Validation

The fault diagnosis system is developed using Microsoft Visual Studio. The system uses the above algorithm to make inferences according to the information inputted by users. Fault time and DMI alarm texts are inputted by users. Position of logs are chosen by users. Features of the target case are extracted by the system and can be seen in the interface. The fault cause and maintenance suggestions of the target case are given in the system interface. The system interface is shown in Fig. 4.

119 practical cases of Jinan Railway Administration in 2012 are selected as test samples to verify the accuracy of our algorithm. In each case, the test data includes fault time, the DMI alarm texts, log of ATPCU, log of SDP and log of C2CU, and the actual diagnosis result. Among them, fault time, the descriptions of DMI alarm texts and the actual diagnosis results are given by fault records, which are written

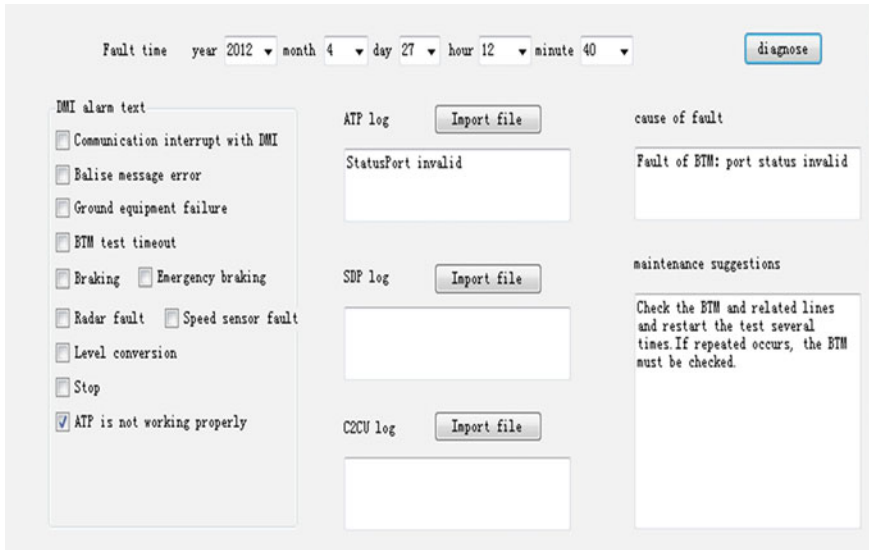


Fig. 4 System interface

by maintenance staff in the technical department of Jinan Railway Administration. The fault records are partly shown in Table 4. The fault time is given by the column “Month”, the column “Day” and the column “Time” in the fault record. The descriptions of DMI alarm texts are given by the column “Description” in the fault record. The actual diagnosis results are given by the column “Type of fault” and the column “Cause of fault” in the fault record. Logs of ATPCU, logs of SDP and logs of C2CU are given in the file format. The diagnosis results of our system are compared with the actual results of test samples. The result is shown in Table 5.

From the distribution of test samples, fault of BTM and fault of wireless timeout occurs the most frequently, accounting for more than 70% in one year. Fault of

Table 4 Fault record

Month	Day	Time	Location	Train	Type of fault	Cause of fault	Description
5	16	16:10	Beijing South Station	6201-01	BTM	Port status invalid	ATP1 is not working properly
2	8	18:25	Beijing–Shanghai Railway	6201-00	Wireless timeout	Fault of RBC	A train had converted from C3 mode to C2 mode
3	10	23:46	Depot	6205-01	ATPCU	Different errors reported in A and B	Communication interrupts with DMI

Table 5 Test results

Type of fault	Number of test cases	Number of successful diagnosis	Accuracy (%)
ATPCU	2	2	100
BTM	38	34	89.4
C2CU	9	7	77.8
DMI	5	2	40
Speed detection	5	3	60
TCR	2	0	0
Wireless timeout	48	47	97.9
Other	10	4	40

C2CU takes the second place, accounting for about 7% in one year. The number of other types of fault is very few. From the test results, our algorithm achieves high fault diagnosis rate in fault of ATPCU, BTM, C2CU, and wireless timeout. The overall accuracy of the algorithm is 80.1% on specific fault, which shows that the algorithm is practical.

6 Conclusion

In this paper, Case-based Reasoning model is introduced into fault diagnosis of vehicle equipment for high-speed railway. The algorithm can give accurate causes of fault and maintenance suggestions by calculating similarity in the support of database containing existing successful cases. It can provide some help for fault diagnosis and maintenance of vehicle equipment for high-speed railway.

Acknowledgments This work is supported by National Key R&D Program of China (2016YFB1200401).

References

1. Zhang SG (2008) General technical scheme of CTCS-3 train control system. China Railway Press (in Chinese)
2. An YH (2015) The failure analysis and research of the train control on-board equipment (CTCS3-300T). China Academy of Railway Sciences (in Chinese)
3. Chen CY, Zhang D, Ren P, Xu HD (2003) Study on automobile fault diagnosis expert system based on fault tree analysis. *Trans Chin Soc Agric Mach* 34(5):130–133
4. Cheng Y, Xu T, Yang L (2013) Bayesian network based fault diagnosis and maintenance for high-speed train control systems. In: International conference on quality, reliability, risk, maintenance, and safety engineering. IEEE, 1753–1757

5. Wang F, Xu T, Tang T et al (2016) Bilevel feature extraction-based text mining for fault diagnosis of railway systems. *IEEE Trans Intell Transp Syst* 18(1):49–58
6. Marquez FG, Lewis WR, Tobias MA, Roberts C (2008) Life cycle costs for railway condition monitoring. *Transp Res Part E* 44(6):1175–1187
7. Chen ZQ (2014) Research on CTCS-3 and CTCS-2 information fusion of vehicle equipment. *Railway Signal Commun Eng* 11(5):1–3 (in Chinese)
8. Du XM, Yu YL (1999) Case-based reasoning for multi-attribute evaluation. *Syst Eng Electron* 21(9):45–48
9. Bergmann R, Stahl A (1998) Similarity measures for object-oriented case representations. In: *European workshop on advances in case-based reasoning*, Springer, Berlin, pp 25–36

A Bayesian-MCMC Model to Assess Metro Train Collector Shoes Slider Degradation Under Different Materials

Yue Pan, Guoqiang Cai and Xi Li

Abstract This paper presents a Bayesian-MCMC model to assess collector shoes slider degradation under different materials. A Markov Chain Monte Carlo (MCMC) method, based on the Bayesian decision model, is put forward and built up a framework in case of the life cycle of collector shoes under different materials forecast. All of inspection data is gathered from Beijing metro lines, and WinBUGS software are used to predict the slider's wear rate. Result shows that the difference between the predicted value and the real one is less than 10% of the later one. Consequently, in case of new metro equipment parts, the newest method is able to ensure the safety operation in the metro by providing a valid device to the equipment manufacturers, the maintenance department as well as the purchasing department of the metro equipment.

Keywords Metro maintenance · Bayesian-MCMC model · Bayesian method MCMC · Collector shoes slider

1 Introduction

A slider is considered as one of the most important components of the third rail collector shoes system in the metro [1]. Two main indispensable materials, including both carbon and copper, are used to produce the slider. Especially, its wear has, explicitly, made a huge effect on the safety of vehicle travelling [2]. Besides, some approaches have been used to ensure the timeline of collector shoes

Y. Pan · G. Cai
Beijing Jiaotong University, Beijing 10044, China

Y. Pan · G. Cai (✉)
State Key Lab of Rail Traffic Control and Safety, Beijing 100044, China
e-mail: gqcai@bjtu.edu.cn

X. Li
R&D Center, Beijing Mass Transit Railway Operation Corporation Limited,
Beijing 100044, China

maintenance by taking experiences as a main drive, though the actual situation is rarely taken into serious consideration. Most of the studies are on the subject of collector shoes and tend to use the result of the wear test experiment in order to make a prediction about the wear rate of collector shoes slider as well as determine the replacement cycle. Currently, only two kinds of receiving modes, including overhead contact network and third rail power supply, are used for metro vehicles. The Beijing metro has been using third rail power supply mode to deliver the power of traction power supply system through the contraction with the third rail to the vehicle. An occurrence of severe frictional loss existed at the sliding contact friction, between the collector shoes and third rail. Therefore, based on a theory of ensuring the third rail service life, it highly lists the necessary elements, consisting of a match of contact rail and collector shoes materials, abrasion resistance as well as collector dynamic performance, of choosing a better slider, without losing any performance condition.

The abilities of both copper alloy [3] and the carbon alloy [4] materials are exposed as good electrical conductivity, wear-resistant, arc resistance and high strength, and simultaneously certain impact and heat resistance as well. The divide of the slider in Beijing metro line are copper alloy and carbon alloy materials. As the slider is one of the most important components of the third rail collector shoes system in the metro, it is important to improve the collector service life. Moreover, obtaining the information of degradation rate or the whole life cycle of slider also can make an effect to reduce the amount of maintenance and operating costs.

Having said that, this paper puts forward a Bayesian Markov Chain Monte Carlo (MCMC) model for collector shoes slider degradation in order to assess the degradation rate through the whole life cycle of the infrastructure. This model runs using the inspection data from the Beijing metro lines. The Beijing metro lines have a total length of 574 km and contain 19 lines. The renewal and maintenance works of the slider performed included collector shoe complete renewal and thorough improvement of the slider. The sample analyzed in this paper includes a series of inspection data and nine lines of Beijing metro. Unfortunately, reliable inspection data is limited because of the measured error.

2 Bayesian and MCMC Method

Bayesian approaches [5] were considered as interesting alternatives to the ‘classical’ statistical theory in stochastic modelling, before the late 1980s. Statisticians discovered MCMC methods [6] in the early 1990s, Bayesian statistics suddenly became the latest fashion modelling method in all areas of science.

Because the Bayesian approaches consider parameters as random variables that follow a prior distribution, which later combined with a traditional likelihood to obtain the posterior distribution of the parameters of interest, therefore, the calculation of the posterior distribution $f(\theta | x)$ of the parameters θ given the observed data x can be computed as:

$$f(\theta | x) = \frac{f(x | \theta) f(\theta)}{f(x)} \tag{1}$$

According to the given calculation above, we are now able to build the PDF of the degradation rate. We can obtain simple distribution through basic sampling algorithms, but when the analysis equation of the distribution is so complex, we cannot do sampling directly. The MCMC methods can solve this kind of problem, and can be applied in practical maintenance. The basic idea of MCMC methods is to construct a Markov chain that starts sampling from an arbitrary state and transfers according to the Markov chain. After a period of sampling, the sample will be very close to the target distribution, finally, choose the results as the final sample. There are two common methods to produce large amounts of eligible data, they are as follows.

2.1 Gibbs Sampling Algorithm

This algorithm is suitable for processing incomplete information when the joint distribution of multiple variables is not clear, but the conditions of each variable in the distribution are known. The basic idea is to use the sampling of the conditional distribution to approach the sampling of the joint probability distribution, and the statistical characteristics of the final sample can reflect the characteristics of the joint probability distribution.

2.2 Metropolis–Hasting Sampling Algorithm

Metropolis–Hasting algorithm, referred as M–H algorithm, in this theory, we assumed the distribution of the samples is $p(x)$, $Q(X^* | X)$ is transition probability function from sample x to x^* , according to this theory, given an initial value $x^{(0)}$ as the sampling point, then produce sample $x^{(1)}$ through transition probability function. Using random function to generate a number u from $[0,1]$, the computing processing is as follows:

$$\begin{cases} x^{(1)} = x^*, & u > 1 \\ x^{(1)} = x^*, & u < 1 \text{ and } u < \frac{p(x^*)}{p(x)} \times \frac{Q(x | x^*)}{Q(x^* | x)} \\ x^{(1)} = x^{(0)} & \text{else} \end{cases} \tag{2}$$

After a large number of iterations, we can obtain the statistic characteristic of the samples which is the target distribution $p(x)$. In this paper, the Metropolis–Hasting algorithm is used to generate a large number of data which are consistent with the posterior distribution.

3 Prior Distribution (Log-normal Distribution)

Having introduced Bayesian idea, the log-normal distributions [7, 8] are put forward to describe the life period of metro components. The log-normal distribution is a complete distribution which is non-negative. Above all, we can assess the value of the current shoes slider's geometry parameters C_1 and C_0 by sampling from Beijing metro lines' inspection data. It was found that C_0 has a log-normal distribution that was not rejected at 5% significance level. So, it is suitable to use this distribution to describe the degradation of the sliders that was caused by the travelled distance. Moreover, the calculation of $p(x)$ of the random variable X , given the parameters σ , μ can be computed as follows:

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad (3)$$

where σ , μ is the mean value of deviation (mm); σ is the standard deviation measured after renewal (mm).

4 Collector Shoes Slider Geometry Degradation Model

From the respect of physical significance, slider degradation is always non-negative, and it has some mistakes since the slider was produced. In the past, many experimental studies have validated a linear relationship between the sliders degradation and the accumulated travelled distance. Therefore, the evolution of the standard deviation can be estimated using the following linear relationship [9]:

$$\delta_{LDi} = c_1 + c_0 T_i, \quad (4)$$

where δ_{LDi} is the standard deviation (mm); c_1 is the initial standard deviation measured after renewal (mm); c_0 is the deterioration rate (mm/3000 km, mm/5000 km or mm/10,000 km, determined by different actual inspected value); T_i is the accumulated travelling miles after the latest renewal.

Having said that, this paper puts forward a Bayesian-MCMC model for the slider to assess the degradation phenomenon under different materials, and a new method to predict the life of the metro equipment parts is established.

5 Model Established

5.1 Prior Distribution

It was found that the hypothesis, the deterioration rate c_0 followed a log-normal, was rejected at 5% significance level. The same does not apply to the initial

standard deviation (c_1), the hypothesis is not rejected at 10% significance level. There are also other distributions that were fitted to both variables, but the log-normal distribution has shown the highest p -value for the Kolmogorov–Smirnov (K–S) goodness of fit test. Moreover, we supposed that the value of the initial standard deviation is zero so that a log-normal distribution can be defined as a prior distribution for the degradation parameters.

5.2 Bayesian Model

Having confirmed the prior distribution, we are now able to specify the model for the current collector shoes blocking degradation phenomenon as follows:

$$\sigma_{LD_i} = c_1 + c_0 T_i \quad i = 1, 2, 3, \dots, n \sim N(c_1 + c_0 T_i; \sigma^2) \tag{5}$$

$$f(c_0, c_1, \sigma^2) = f(c_0, c_1) \cdot f(\sigma^2) \tag{6}$$

$$\log(c_0) \sim N(\mu, \sigma^2) \tag{7}$$

$$\sigma^2 \sim \text{IG}(a, b) \tag{8}$$

We will assume that (c_0, c_1) and σ^2 are independent, and therefore, their joint prior density function factorizes into their marginal prior density functioned as formulated above. Therefore, we will assume that (c_0, c_1) will follow a priori a bivariate log-normal distribution, whereas σ^2 will follow an inverse-gamma distribution. Now that the normal model is difficult to be established, the MCMC simulation is needed to assess the posterior distributions.

5.3 MCMC Processing

In order to obtain specific parameter values for the posterior function, we adopt some methods in accordance with the PDF sampling, when the number of samples is very large, the samples show statistical characteristic which contain some information of the posterior function formula. Finally, we know that in the inverse-gamma distribution, the parameters were chosen so that it can be considered a vague prior distribution. A typical method in normal models will be used to obtain the parameter $\sigma^2 \sim \text{IG}(a, b)$. The $\tau = \sigma^{-2}$ is used for calculating conveniently, and we can find $\tau \sim \text{IG}(a, b)$. Therefore, as the precision should be vague, parameters were set as $a = b = 0.1$.

Having specified the model and conducted proper solicitation of the priors, we are able to perform a MCMC simulation to obtain the posterior distribution of the

parameters of interests which in this case is the deterioration rate (c_0). Therefore, in order to simulate it, the model specified above was built in WinBUGS [10] software.

6 Experiments

While as the model is established, the inspection data is gathered from 10 lines of Beijing metro which has less amount of error. As the log-normal distribution is not convenient to calculate, we do a simple transformation and present the log one after the original one. Therefore, we can compute the specific value of μ, σ^2 of different lines. Divided into carbon alloy and copper alloy, the slider degradation data was analysed. Note that the standard deviation is expressed in Eq. 4, the intimal standard deviation and the accumulated travelling miles can be calculated. Therefore, the degradation in life cycle can be running a Monte Carlo simulation. Divided into carbon alloy and copper alloy, the slider degradation data was analysed. As Table 1 shows, the degradation rates in life cycle were predicted.

7 Results

The result of the Bayesian-MCMC model is summarized in Table 2. In order to compare the differences between the inspection data series of the degradation and the predicted ones under different materials, the dissimilarity between the real value and predicted one also be seen in Fig. 1.

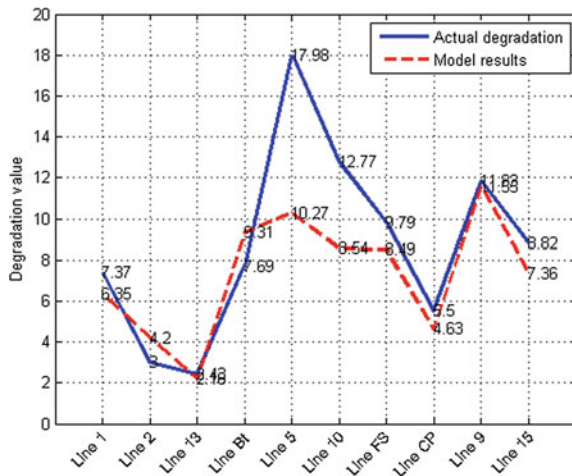
Table 1 The result of degradation rates prediction

Copper alloy		Carbon alloy	
Line	Degradation rate	Line	Degradation rate
mm/5000 km		mm/5000 km	
Line 1	0.6054	Line 5	0.88
Line 2	0.3897	mm/3000 km	
Line 13	0.2945	Line 15	1.142
Line 10	0.7335		
Line Fangshan	0.9764		
Line Changping	0.6933		
Line 9	0.9264		
mm/10,000 km			
Line Batong	0.9566		

Table 2 Comparison between actual degradation and results

Line	The accumulated travelling miles (km)	Predicted degradation rate (mm/5000 km)	Actual degradation (mm)	Results (mm)
<i>Copper alloy</i>				
Line 1	52425.4	0.6054	7.37	6.35
Line 2	53,911	0.3897	3	4.2
Line 13	36,965	0.2945	2.43	2.18
Line Batong	97,394	0.4783	7.69	9.31
Line 10	58,219	0.7335	12.77	8.54
Line Fangshan	43,495	0.9764	9.79	8.49
Line Changping	33,358	0.6933	5.5	4.63
Line 9	62,356	0.9264	11.83	11.55
<i>Carbon alloy</i>				
Line 5	58,324	0.88	17.98	10.27
Line 15	19,323	1.9033	8.82	7.36

Fig. 1 Comparison between actual degradation and results



8 Discussion

In this model, the average thickness and the wear of the slider is used—the average thickness is an average of eight sliding blocks on one car. The discussion is as follows:

(1) This model makes same travelling mileage measured value as the sampling interval, which is different from the original data sampling interval, therefore,

approximate data processing (e.g. make 15,000 km as an approximate value of 14,825 km) is used. (2) The small amount of missing data is estimated according to the wear rate to ensure the integrity of the data. (3) Due to the difference in load, current density and sliding speed, the wear rate of the slider is different, but it will not cause calculation error of the predicted value.

9 Conclusion

To sum up, in this paper, a Bayesian-MCMC model for collector shoes slide degradation is put forward. A log-normal distribution is used as a prior probability distribution to model degradation parameters. Moreover, the results show that the predicted degradation rate of this model is based on the expected value after a large number of iterations in the MCMC process. Thus, it can be used to predict the most worn slider in the actual life cycle, which is regarded as the highest degree of cruelty. Consequently, this model can be used to predict the degradation of a slider according to the metro running mileage.

Acknowledgements This work was supported by the Independent Research Project of State Key Lab of Rail Traffic Control & Safety of Beijing Jiaotong University (NO. RCS2017ZJ001).

References

1. Dong L, Chen GX, Zhu MH et al (2007) Wear mechanism of aluminum–stainless steel composite conductor rail sliding against collector shoe with electric current. *Wear* 263 (1):598–603
2. He DH, Manory RR, Grady N (1998) Wear of railway contact wires against current collector materials. *Wear* 215(1):146–155
3. Mo D, He G, Hu Z et al (2007) Study on Cu-based slider-crank for collector shoe of maglev vehicle. *Powder Metall Technol* 25(5):336–335
4. Tsuchiya Z (1950) On the carbon slider of the pantograph current collector. *TANSO*, 1
5. Probability IT, Edition N (2002) Introduction to probability, 2nd Edition, introduction to probability. *Athena Scientific* 2002:625–627. (Sect. 4.7)
6. Chernozhukov V, Han H (2003) An MCMC approach to classical estimation. *J Econometrics* 115(2):293–346
7. Fu S (2016) A hierarchical Bayesian approach to negative binomial regression [J]. *Methods App Anal* 22(4)
8. Lai C, Rayner JCW, Hutchinson TP (1998) Properties of the sample correlation of the bivariate Lognormal distribution. *ICOTS* 5
9. Andrade AR, Teixeira PF (2012) A Bayesian model to assess rail track geometry degradation through its life-cycle. *Res Transp Econ* 36(1):1–8
10. Fryback DG, Stout NK, Rosenberg MA (2001) An elementary introduction to Bayesian computing using WinBUGS. *Int J Technol Assess Health Care* 17(1):98–113

Design of High Power Factor Charger for Transportation Electrification

Fucun Li, Liang Guo, Hao Li and Guodong Qu

Abstract The electrification of transportation equipment have a high impact on the power grid, especially when the electric energy is used as its power source, the power conversion of the charging equipment will produce a large amount of harmonics. In order to reduce harmonics and improve operating power factor of the charging system, we select a specific circuit topology based on analysis of the operating characteristics of this type of charger, and carry out theoretical analysis and simulation verification. On these foundations, the prototype of charger is made and closed-loop debugging is carried out, the experimental results show that the output voltage of the prototype is adjustable, the grid-side current is undistorted, and realized a unit power factor operation.

Keywords Traffic charger · Unit power factor · Prototype · Experimental verification

1 Introduction

At present, in most electrified transportation, the charging equipment is usually composed of uncontrolled devices, the utility model has the advantages of low cost and simple control, but its drawback is that when charging, it will cause grid-side current distortion, the harmonic pollution caused by charger will not be ignored when the charging power is large enough, at the same time, these harmonics will also affect the accuracy of the energy measurement.

According to different classification methods, there are many different rectifier circuit, combined with different charging occasions and operating modes, each structure has different application occasions. In electrified transportation, charging equipment is usually powered by specially laid lines, the charging operation time is

F. Li (✉) · L. Guo · H. Li · G. Qu

State Grid Shandong Electric Power Research Institute, No. 2000 Wangyue Road, Shizhong District, Jinan, Shandong, People's Republic of China
e-mail: fengxiao0516@yeah.net

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_36

363

long, and the bidirectional flow of electric energy is not necessary. In order to adapt to the charging characteristics of electrified transportation and reduce the influence on the quality of distribution grid, the requirements for the functional characteristics of the charger are as follows:

- (1) The AC side voltage and current phase should be consistent when charging, and the grid-side current should undistorted, that is to say, unit power factor operation is required on the grid side;
- (2) While achieving high power factor operation, the charger can output a DC voltage with adjustable amplitude;
- (3) In the later stage of charging, the voltage can be changed further and a constant current charging can be realized through a DC/DC transformation;
- (4) The production cost of the charger needs to be reduced as much as possible.

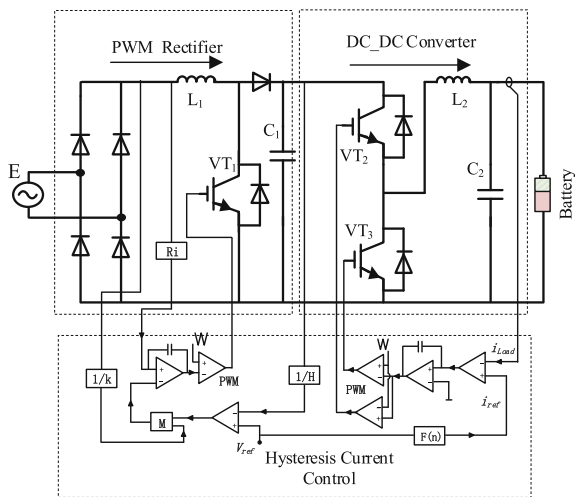
2 Topology Analysis and Simulation

2.1 Topology and Theoretical Analysis

In combination with the requirements of this type of charger, we select a AC/DC + DC/DC mode, in the AC/DC part, we use a rectifier bridge and a controllable device to form a boost rectifier, in the DC/DC section, we use two controllable devices consisting of a Bi-DC/DC converter, the whole charger contains only three controllable devices only, the system structure is shown in Fig. 1.

When the continuous inductance current is continuous, the inductance voltage meets the requirements of formula (1) as follows:

Fig. 1 Charger system diagram



$$\begin{cases} L \frac{di_L}{dt} = u_{in} & 0 \leq t \leq dT \\ L \frac{di_L}{dt} = u_{in} - u_o & dT \leq t \leq T. \end{cases} \quad (1)$$

According to different switching states, formula (1) can be converted to

$$L \frac{di_L}{dt} = u_{in} - (1 - d)u_o \quad d \in \{0, 1\}. \quad (2)$$

The capacitance current meets the requirements of formula (3) as follows:

$$\begin{cases} C \frac{du_o}{dt} + \frac{u_o}{R} = 0 & dT \leq t \leq T \\ C \frac{du_o}{dt} + \frac{u_o}{R} = i_L(t) & 0 \leq t \leq dT. \end{cases} \quad (3)$$

According to different switching states, formula (3) can be converted to

$$C \frac{du_o}{dt} + \frac{u_o}{R} = (1 - d)i_L(t) \quad d \in \{0, 1\} \quad (4)$$

By Laplace transform, formula (2) can be transformed into

$$sLi_L = u_{in} - (1 - d)u_o. \quad (5)$$

In combination with (4), the output voltage is

$$u_o = \frac{Ri_D}{RCs + 1}. \quad (6)$$

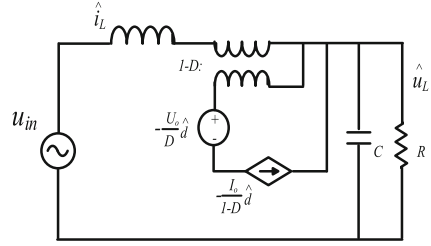
According to (2)–(6), we enable the $s = 0$, and DC operating points can be obtained as follows:

$$\begin{cases} U_o/U_{in} = 1/(1 - D) \\ I_o = \frac{U_o}{R} \\ I_o = (1 - D)I_L. \end{cases} \quad (7)$$

At the same time, according to the above analysis, equation of state for linear small signal model of the system can be obtained as follows:

$$\begin{cases} \widehat{i}_{in}(s) = \widehat{i}_L(s) \\ sLi_L(s) = \widehat{u}_{in}(s) + U_o \widehat{d}(s) - (1 - D)\widehat{u}_o(s) \\ \widehat{u}_o(s) = \frac{R}{1 + sRC} \left[(1 - D)\widehat{i}_L(s) - \frac{I_o}{1 - D} \widehat{d}(s) \right]. \end{cases} \quad (8)$$

Fig. 2 AC small signal equivalent model



The AC small signal equivalent model is shown in Fig. 2.

According to the formula (8), we can get the relation between output voltage, inductance current, input voltage, and duty cycle as follows:

$$\begin{bmatrix} \hat{u}_o \\ \hat{i}_L \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} G_{yv} & G_{dv} \\ G_{vi} & G_{di} \end{bmatrix} \begin{bmatrix} \hat{u}_{in} \\ \hat{d} \end{bmatrix}. \tag{9}$$

Here,

$$\Delta = s^2LC + s\left(\frac{L}{R} + DC\right) + \frac{D}{R} + 1.$$

$$\begin{aligned} G_{yv} &= 1 - D & G_{dv} &= \frac{U_o[(1-D)^2(2+sRC)-\Delta]R}{(1-D)(sRC+1)} \\ G_{vi} &= sC + \frac{1}{R} & G_{di} &= U_o\left(\frac{2}{R} + sC\right). \end{aligned}$$

2.2 Simulation

In terms of power factor improvement, the current control has peak current control, average current control, and current hysteresis control, and so on, among them, the average current control has the characteristics of constant switching frequency, no ramp compensation, low noise sensitivity, and so on. The specific process is first to sample and filter the inductance current, compared with the reference value and then output after current error amplifier, this signal drives PWM pulse. The inner current loop is to keep the input current consistent with the reference current; the outer voltage loop ensures that the output voltage follows the instruction requirements.

The main factors affecting the converter conversion are grid-side inductance, output side capacitance, switching frequency, and so on. For the grid-side inductance, it plays the role of storage, conversion, filtering, and so on, it determines the high-frequency ripple current content of the grid side, the inductance current may also be interrupted if the ripple current is too large.

The size of the output capacitor mainly considers the following factors: high-frequency ripple current, DC bus voltage on the output side, maintenance

time, and so on, the maintenance time is especially critical, it means that the output voltage remains stable after the input of the grid is cut off. Its value is determined by the energy stored in the capacitor, the output power, the output DC bus voltage, and the minimum allowable output voltage of the load.

The selection of switching frequency mainly includes the size of filter inductance, the requirement of electromagnetic compatibility, and the loss of power switch. We know that the higher the switching frequency, the smaller the volume of the magnetic devices, but the increase of switching frequency also means that the switching loss and the diode reverse recovery loss increase, at the same time, in order to reduce the difficulty of EMC design, the switching frequency cannot be too high. Here, we built a simulation model under the input voltage $V_s : 220\sqrt{2} \sin 50 t$, the output voltage of $U_o: 550\text{--}650 \text{ V}$, the switching frequency $f_s = 140 \text{ kHz}$, the output capacitance $C = 3000 \mu\text{F}$, and the grid-side inductance $L = 3 \text{ mH}$, the simulation results are shown in Fig. 3.

From the simulation results, we can see that the grid-side current is not distortion, the output voltage can be adjusted within a certain range and the system achieved a unit power factor operation.

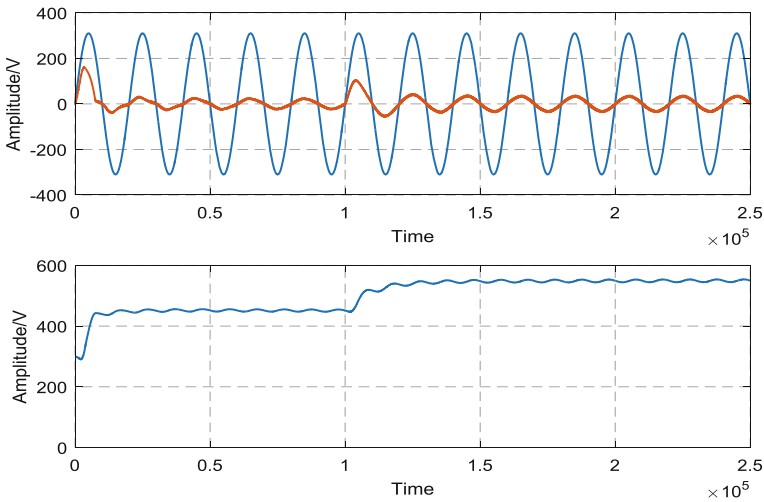


Fig. 3 System simulation results

3 Manufacture and Experiment of Prototype

3.1 Prototype Manufacture

The charger prototype mainly consists of two parts, the main circuit part and the control circuit part, according to the specific parameters set in the previous section, for controllable devices, we chose 6 MOSFET models for IRFP460, for fast recovery diodes, the reverse voltage is mainly considered, here we choose 4 MUR20100 whose voltage withstand capacity is 1000 V. The prototype of the charger is shown in Fig. 4.

3.2 Experiment

According to the charger, we first make the control part not enabled, that is, there is no closed-loop control for the charging system, the experimental waveforms of the grid-side voltage and current of the system are shown in Fig. 5.

From the waveform shown in Fig. 5, we can see that the grid-side current distortion is serious when the system is not closed loop, its harmonic pollution to the grid will not be ignored when the charging power is large enough. Then we have double closed-loop debugging of the system, the debugging results are shown in Fig. 6.

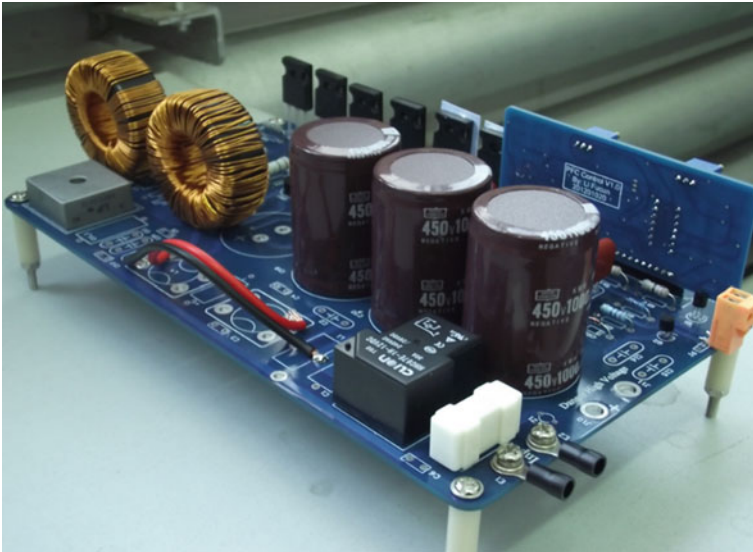


Fig. 4 Prototype of charger

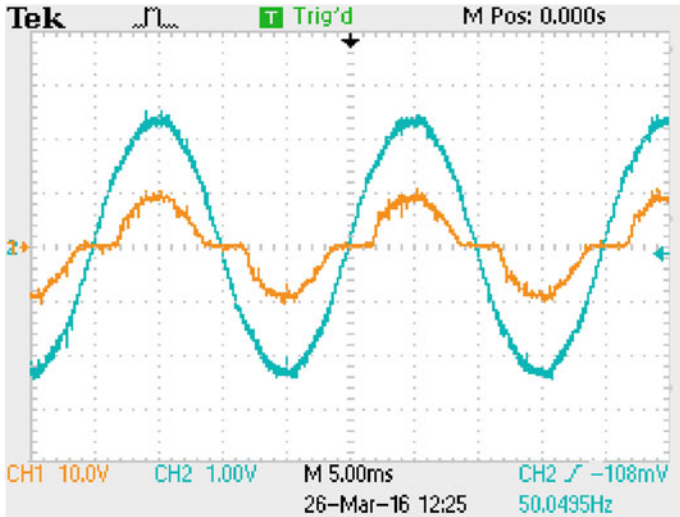


Fig. 5 The experimental waveforms without closed loop

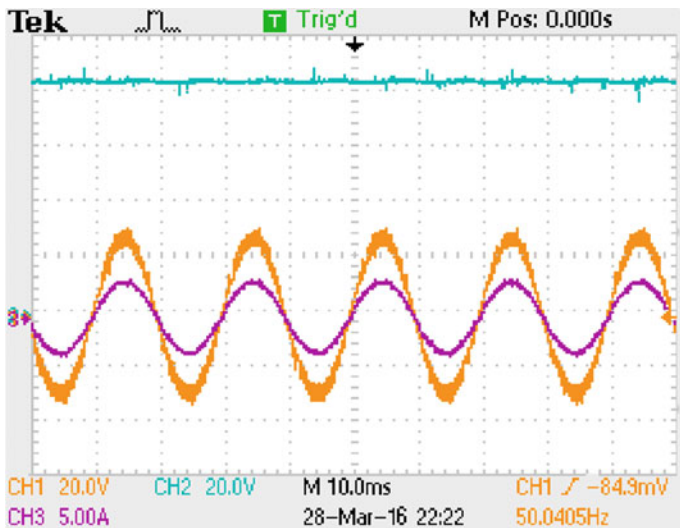


Fig. 6 Experimental waveforms in double closed loop

As can be seen from Fig. 6, the grid-side voltage and current phase of the prototype the same, through FFT analysis, we can get that the power factor is 0.99 and the harmonic distortion rate is 2.7%. From the above experimental results, we can see that the system can realize the controllable output voltage and realize a unity power factor operation.

4 Conclusions

In this paper, according to the characteristics of the charger for transportation electrification, we first analyze its theory combined with specific topological structure, and then set some important parameters of the structure, and following a system simulation is carried out according to the set parameters, simulation results verify the feasibility of this method.

On the basis of theoretical analysis and simulation verification, a prototype of the charger is made and closed-loop test is carried out, the experimental results show that grid-side current does not distort and the output voltage of the prototype is controllable, the charging system realized a unity power factor operation.

References

1. Thangavelu T, Shanmugam P, Raj K (2015) Modelling and control of VIENNA rectifier a single phase approach. *IET Power Electron* 8(12):2471–2482
2. Cheng L, Kai Y, Xiaoguang W (2016) A design of high power factor single phase inverter. In: 2016 IEEE 8th international power electronics and motion control conference (IPEMC-ECCE Asia), pp 3111–3115
3. Khrwat LMDA (2015) FACTS devices modeling in real-time dynamic control of power systems. In: 2015 16th international conference on sciences and techniques of automatic control and computer engineering (STA), pp 438–444
4. Gadekar S, Kulkarni N, Mhetre S, Kulkarni HH (2015) Design and development of passive filter and comparative study of simulation results of passive and active filter. In: 2015 international conference on energy systems and applications, pp 324–328
5. Shete AR, Gavade SS (2014) The case study of automatic power factor controller on distorted system with overview of harmonics reduction technique. In: 2014 international conference on computation of power, energy, information and communication (ICCPEIC), pp 336–341
6. Gang Z, Yongfang G, Hongbo S (2012) The simulation research of the high power factor charger in vehicle. In: 2012 third international conference on digital manufacturing & automation, pp 939–942
7. Su W, He M (2011) Research and simulation of active power factor correction. In: 2011 international conference on consumer electronics, communications and networks (CECNet), pp 5156–5158
8. Xianjin Z, Chenglong G, Jianhua W (2010) Input admittance analysis based on one cycle control in boost single-phase pfc converters, In: 2010 international conference on electrical and control engineering, pp 3838–3841
9. Huang H, Fu J (2012) Design and experiment of 90 W switching power supply with PFC funtion. *J Power Supply*, 5–9
10. Xu R, Fang W, Liu X, Liu Y, Hu Y, Liu Y (2014) Design and experimental verification of on-board charger for electric vehicle. In: 2014 international power electronics and application conference and exposition, pp 1422–1427

Comparative Study of Fault Detection Algorithm Based on Multivariate Statistical Analysis

Shuyu Zhang

Abstract This paper has studied several methods based on multivariate statistical analysis, DPCA, CPCA, and MBPLS, which are all the extension of PCA, mainly introducing the principles and steps. At the same time, a method for determining the threshold in practice is proposed in this paper. Besides, we verify the effectiveness of the detection method by the data of train suspension system from simulation experiment. And then we make a comparative analysis of the results through the effect and time. According to the results, we can find it is obvious that CPCA and MBPLS are superior to DPCA in detecting faults.

Keyword Fault detection · Multivariate statistical analysis · Comparative study

1 Introduction

Currently, international and domestic researchers have obtained some research achievements about the online monitoring of train key components and fault diagnosis technique, which have drawn a lot of attentions. In recent years, more and more scholars have studied the problem about the data-driven fault diagnosis method such as Dong Hua L et al. and Liu H [1, 2], and one of the most important methods is based on the multivariate statistical analysis of fault detection methods.

The basic approach is to decompose the multivariable sample space into lower dimension projection space and residual subspace, in which the subspace is composed of the main element variables. Then, we construct two statistics in the space of the decomposition to reflect the change of space. We can realize the monitoring of the process by projecting the observation vector onto the subspace that is decomposed and calculating the value of the statistics which mentioned in four articles by Jiang Ch, Zhou DH, Wei Q, Wen B and He Y respectively [3–6].

S. Zhang (✉)

School of Traffic and Transportation, Beijing Jiaotong University,
No. 3, Shangyuan Village, Haidian District, Beijing, China
e-mail: jade0135@163.com

This paper briefly introduces the basic principles of DPCA, CPCA, and MBPLS which have been studied by Sang WC and Wei X et al. [7, 8] based on multivariate statistical analysis, and makes a comparison and analysis by verifying their application in fault detection with the simulation data of train suspension system.

2 Principle of Several Fault Detection Algorithms

2.1 Principal Component Analysis

Principal component analysis (PCA) is the core and foundation of fault detection and diagnosis based on multivariate statistical analysis. The basic idea of PCA is to obtain a new low-dimensional space by reducing dimension of the original information, where it contains the main information of the original space. The new variable is a linear combination of the original variable. The primary and secondary statuses can be obtained according to variance, and then it is possible to get the first principal component, the second principal component, etc. which are independent and arranged in order of priority.

In the fault detection, PCA uses the statistical method to establish the principal component model according to the historical data under the normal working condition, that is, finding the low-dimensional principal component of the original variable. It can determine the monitoring process failure, if it finds any difference between the measured data and the model during the monitoring process. The modeling process is as follows:

1. Historical data sets under normal operating conditions: $X = (x_{ij})_{n \times m}$
2. Normalize X , M is the mean vector, and $D_\sigma^{1/2}$ is the standard deviation matrix

$$\bar{X} = [X - (11 \dots 1)^T M] D_\sigma^{1/2} \quad (1)$$

3. To find the covariance matrix of the normalized matrix \bar{X} ,

$$\sum = \frac{1}{n-1} \bar{X}^T \bar{X} \quad (2)$$

4. Find the eigenvalue of the covariance matrix, $\lambda_1 \geq \lambda_2 \dots \geq \lambda_m$, and find its unit eigenvector p_1, p_2, \dots, p_m :

$$\sum = \sum_{i=1}^m \lambda_i p_i p_i^T \quad (3)$$

5. Determine the number of principal elements by the eigenvalue variance cumulative contribution rate method. When the cumulative contribution rate of l principal elements exceeds the threshold (selected according to the actual situation), such as 90%, the corresponding l is the number of principal elements that need to be retained.

$$CPV(l) = 100 \left(\frac{\sum_{i=1}^l \lambda_i}{\sum_{i=1}^m \lambda_i} \right) \% \quad (4)$$

6. Principal component score matrix and principal component vector are

$$\hat{T} = \bar{X}\hat{P} \quad (5)$$

$$t_i = \bar{X}p_i, i = 1, 2, \dots, l \quad (6)$$

2.2 Two Important Statistics and Control Limits

To determine whether there is a fault in the process, mainly, we carry out hypothesis testing through the establishment of statistics, from the perspective of statistical analysis. For statistical testing, the *Hotelling T²* statistics would be established in the main subspace, and the square prediction error (SPE) statistics would be established in the residual subspace.

Hotelling T² statistics reflect the multivariate changes by the fluctuation of the main component vector in the principal component model: t_i is the i th row in the principal component score matrix \hat{T} ; Λ is the diagonal matrix of $l \times l$ consisting of the eigenvalues corresponding to the first l principal elements; and \hat{P} is the principal component matrix.

$$T_i^2 = t_i \lambda^{-1} t_i^t = \bar{X}_i \hat{P} \Lambda^{-1} \hat{P} \bar{X}_i^T \quad (7)$$

The control limits for the *Hotelling T²* statistics are available through the F distribution. It can only verify the variation of certain variables in the principal subspace through the test based on *Hotelling T²*. At the same time, the fault cannot be detected when the measured variables are not good in the main element model.

The SPE statistic can also be called a Q statistic, which mainly reflects the situation where the measured value deviates from the principal component model at that time. It can monitor multiple variables and express information that is not interpreted by the principal component. \bar{X}_{ij} is the measured value of the j th variable at the time i of the normalized data, and \hat{X}_{ij} is the data reconstruction value of the j th variable at time i , $i = 1, 2, \dots, m$. The control limit of the SPE can be calculated by the approximate distribution.

$$SPE(i) = \sum_{j=1}^m (\bar{X}_{ij} - \hat{X}_{ij})^2 \quad (8)$$

SPE statistics and *Hotelling T²* statistics represent the degree of change about projection of the data in the subspace and the main subspace. In terms of the size of

the control limit, the T^2 control limit is greater than the SPE. Besides, to detect the fault by SPE and T^2 , there are four test results:

1. T^2 and SPE statistics exceed the control limit.
2. T^2 exceeds the control limit and the SPE does not exceed the control limit.
3. T^2 do not exceed the control limit and the SPE exceeds the control limit.
4. T^2 and SPE statistics did not exceed the control limit.

The first and the third correspond to the occurrence of the fault; the second result may be disturbed; the fourth indicates that the process is normal.

2.3 DPCA Algorithm Principle

Dynamic PCA or DPCA method is proposed on the basis of traditional PCA, in order to solve the problem that the process data contains autocorrelation in addition to the strong cross-correlation.

In a dynamic system, the data value at the current time depends in part on the value of the past time. The original observation data is added to the data at the time of the previous s to obtain the broadened data matrix. s is the maximum continuous impact of sampling times between the sampled data.

$$X(s) \begin{bmatrix} x_t^T & x_{t-1}^T & \dots & x_{t-s}^T \\ x_{t-1}^T & x_{t-2}^T & \dots & x_{t-s-1}^T \\ \vdots & \vdots & \vdots & \vdots \\ x_{t+l-n}^T & x_{t+s-n-1}^T & \dots & x_{t-n}^T \end{bmatrix}, \tag{9}$$

where x_t^T is the data collected by m sensors at time t , and then follow the steps of the PCA algorithm that can be calculated in turn.

2.4 CPCA Algorithm Principle

CPCA (Consensus PCA) method is aimed at highly nonlinear and significant interference system, monitoring the data at different stages of a cycle, without predicting future data to realize real-time monitoring. CPCA is the extension of the PCA, and the data is expanded into three-dimensional batch data (batch number \times variable \times time).

CPCA expands the 3D data matrix X into a 2D matrix X , and then performs normal principal component analysis and blocks the two-dimensional data X , in accordance with the time divided into K blocks (each sampling time for one block), and then each block would be processed. Producing the number of principal

component (PCs) of each block, we can get the largest $maxpc$ and then we should find $maxpc$ PCs from each block to model CPCA.

2.5 MBPLS Algorithm Principle

PLS is called “Partial Least Squares” or “Projection to Latent Structure”, according to the idea of the main component extraction to extract the largest degree of correlation in the original data to the new data space. And then using the biased least squares method to make regression, compared with the ordinary least squares, PLS not only reduces the dimension of the original data but also solves the problem of data dependency.

Significant variables can be measured directly, including the independent variable matrix X and the dependent variable matrix Y .

Hidden variable cannot be measured directly; the process is meaningful.

$$X = \hat{X} + E_k = \sum_{i=1}^k t_i p_i' + E_k = TP' + E_k \quad (10)$$

$$Y = \hat{Y} + F_k = \sum_{i=1}^k b_i t_i q_i' + F_k = TBQ' + F_k \quad (11)$$

\hat{X} , \hat{Y} is the fitting matrix of the matrix X , Y ; t_i and u_i are the scoring vectors of the i th PLS hidden variable; and p_i and q_i are the corresponding load vectors).

MBPLS (Multi-block PLS) is the PLS block data expansion, the advantage is that, in addition to the entire process of monitoring space, we can also get each processing block of the monitoring space. When the process goes wrong, it can be easier to detect, separate, and identify. The basic idea of MBPLS is that the hidden variables cannot be directly measured in the system where there is a driving effect, and these hidden variables can be observed through the process variables and linear combination.

3 Algorithm Comparison

3.1 Processing Comparison

See Table 1.

Table 1 Processing comparison

Heading	DPCA	CPCA	MBPLS
Input	Extension matrix of two-dimensional matrix	Contains the three-dimensional data of the batch, into two-dimensional data, divided into K blocks by time	Including two parts of data, part of the data into B block
Whether it is looping	No	Yes	Yes
Score vector matrix calculation	By the load matrix to find out directly	Calculate the k time score vector according to the number of principal elements	Find the score vector of each block and calculate with Y , then compress it

3.2 Comparison of Experimental Results

The train suspension system fault simulation platform built by Matlab and Simpack software can realize the situation simulation of suspension system components at any time any failure, so that we can get a simulation of a series of springs and a series of dampers with 25% performance degradation and 75% performance degradation. Experiments were carried out using the simulation data and comparing their experimental results. The fault detection sampling time is 0.1 s, and faults are introduced into the suspension system after 15 s of Simpack–Matlab co-simulation. The measurement data of the first 250 sampling times of the vehicle running cycle are used for the experiment. The fault information for the simulated data is as follows:

- $K_{15\%}$: Primary spring with 5% reduction
- $K_{25\%}$: Primary spring with 25% reduction.

Black solid line is the threshold; the red solid line means the data of suspension system in the normal; the blue line represents the data of primary spring with 5% attenuation; the green-dotted line represents the data of primary spring with 25% attenuation.

In practical applications, the threshold is set, taking into account the speed of the train and the impact of changes in the load of the train, usually using the value larger than the largest peak of the normal data results, taking a constant.

Figure 1 shows the results of the fault detection using DPCA, CPCA, and MBPLS with SPE and T^2 indicators. Through SPE, we can see that DPCA can successfully detect a 25% attenuation, while a 5% failure cannot be detected by DPCA. However, CPCA and MBPLS can detect both failures of 5 and 25% immediately after the fault occurring at 1 s. By the index of T^2 , we can find that CPCA and MBPLS can quickly detect faults with 5 and 25% attenuation, while DPCA does not detect a 5% attenuation and a 25% failure test obvious, which is similar to the SPE indicator (Table 2).

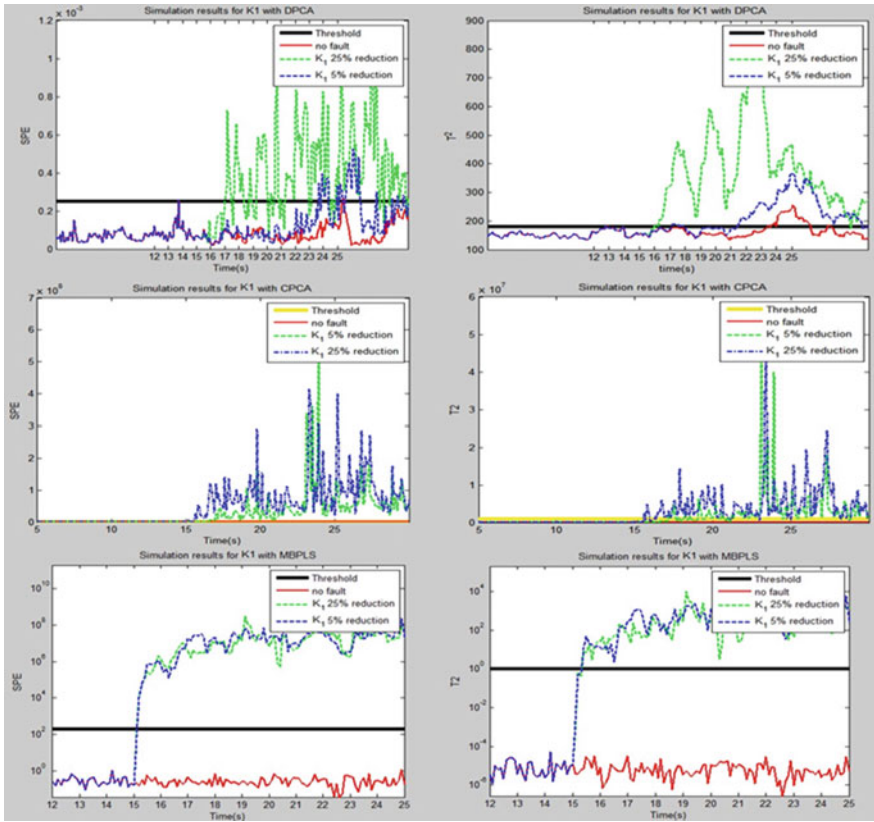


Fig. 1 Fault detection result

Table 2 Effect and time

	DPCA	CPCA	MBPLS
Effect	The fault detection with less attenuation is slower and less noticeable	Fault detection with less attenuation is faster and noticeable	Detection faster and less attenuation of the fault detection is obvious
Calculating time	0.004852 s	0.00347 s	0.003653 s

4 Conclusion

It is obvious that CPCA and MBPLS are superior to DPCA in detecting faults according to the results. Because a system fault may be interpreted as residual noise, it is difficult to detect. At the same time, we can see that the CPCA's and MBPLS's indexes SPE and T^2 are also very sensitive to subtle noise. The method involved in this paper is only to study whether there is a fault, not to diagnose the fault occurred, which deserves further study.

References

1. Donghua Z, Gang L, Yuan L (2011) Data driven industrial process fault diagnosis technology. Science Press (in Chinese)
2. Liu H (2013) Research on the fault diagnosis of urban rail vehicle suspension system. Beijing, China (in Chinese)
3. Jiang CH, Zhou DH (2005) Fault detection and identification for uncertain linear time-delay systems. *Comput Chem Eng* 30:228–242
4. Wei Q (2014) Fault diagnosis of urban rail suspension system based on an experimental platform. Beijing, China (in Chinese)
5. Wen B (2011) Study on fault detection and diagnosis based on principal component analysis. Nanjing, China (in Chinese)
6. He Y (2014) Research on fault detection method based on multiblock PLS. Shenyang, China (in Chinese)
7. Sang WC, Lee IB (2015) Multiblock PLS-based localized process diagnosis. *J Process Control* 15(3):295–306
8. Wei X, Guo Y, Jia L, et al (2013) Fault detection of rail vehicle suspension system based on CPCA. In: Conference on control and fault-tolerant systems. IEEE 2013:700–705

Research on Cycling Energy Saving Based on Improved Violence Search

Liangliu Bai, Yong Zhang, Zongyi Xing, Xinrong Liu
and Xuejin Wang

Abstract Aiming at the energy-saving optimization problem of metro train, this paper presents a new energy-saving research method based on improved violence search method. First, the dynamic model and the motion equation of single train are analyzed, and the model is applied to the timing energy-saving model which is closest to the real situation. Second, the paper analyzes the multiple constraints of single-train running and puts forward the goal of three single-train runnings. Finally, an improved violence search method combining constraint method and violent search method is used to solve the multi-objective problem. Based on the violent search method, this paper transforms the multi-objective problem into the single-objective problem through the target constraint method, and then solves the single-objective problem through the violent search method to make higher solution speed of multi-objective optimization problems. After applying the method, the “speed–distance” and the “power–distance” curves are simulated and calculated with the actual line data of Beijing–Yizhuang Line Rong Jing East Street to Wanyuan Street as input.

Keyword Urban rail transit · Cycling energy saving · Improved violence search method · Timing energy-saving model

1 Introduction

As a new type of urban transportation, the metro train has the advantages of a large carrying capacity, take up almost no space of the urban surface, punctual and efficient and green environmental protection; it grows very fast. Compared with other modes of transportation, it has a great advantage in energy conservation and environmental protection. But its huge energy consumption has become the bottleneck

L. Bai · Y. Zhang (✉) · Z. Xing · X. Liu · X. Wang
School of Automation, Nanjing University of Science and Technology,
No. 200 Xiao Ling Wei Street, Xuan Wu District, Nanjing, China
e-mail: 34445721@qq.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_38

379

that restricts the further development of urban rail transit. Therefore, it is of great importance to reduce transportation cost and improve energy utilization by means of quantitative analysis of the energy consumption impact on urban rail transit trains and finding the energy-saving breakthrough.

The operation optimization of cycling is originated from railway transportation. There are some studies at home and abroad. Yanyan [1] studied the influence of line conditions, vehicle selection, and forecasting traffic on the energy consumption of rail transit and analyzed quantitatively their impact on the station energy consumption factors and influence level through the investigation of operating conditions on subway lines in Beijing, Shanghai, and other cities. Howlett [2] proposed an energy consumption calculation model based on continuous varying slope interval. Hongguo [3] mainly carried out research on optimization of the train running process, constructed a simulation model of multi-target train operation targeting energy consumption, parking accuracy and time-division error, and expressed that constituted a “maximum traction, uniform, idling, maximum braking, the most economical” operation strategy while the train is moving, as the operating speed fluctuates, the lower the energy consumption. The above studies have made great contributions to the energy-saving operation of single trains, but there is a problem that multi-objective optimization is slow to solve.

This paper presents a new energy-saving research method based on improved violence search method. The “speed–distance” curve and the “power–distance” curve are simulated and calculated with the actual line data of Beijing–Yizhuang Line Rong Jing East Street to Wanyuan Street as input, and provides technical support for train energy-saving operation.

2 Train Operation Dynamics Analysis and Motion Equation

2.1 Traction Analysis

In the traction stage of the train, the train’s traction and travel speed are similar to an inverse proportional function. As the speed increases, traction gradually reduced. When the train speed is small, the traction reaches the maximum. Using traditional interpolation method to solve traction when programming and executing programs, we need calling the database repeatedly to query the speed and traction values [4]. Therefore, this paper uses the method of the piecewise curve-fitting [5]. In the process of train traction, sometimes the metro train is not towed by full traction; hence, we introduce the traction coefficient μ_f to demonstrate the ratio of current traction F to total traction F_{\max} , namely

$$F = \mu_f F_{\max} \quad (1)$$

Among them, $\mu_f \in [0, 1]$.

2.2 Braking Force Analysis

The direction of the braking force is the opposite of the forward direction of the train, which prevents the train from moving forward. However, the train will not use all the braking force during the daily operation. Therefore, we need to introduce a braking force coefficient μ_b . μ_b indicates the ratio of the actual braking force B to the maximum braking force B_{\max} , and is a variable coefficient. The train braking force is shown below:

$$B = \mu_b B_{\max} \quad (2)$$

Among them, $\mu_b \in [0, 1]$.

2.3 Resistance Analysis

2.3.1 Basic Resistance

The basic resistance of the train refers to the line, locomotive, running speed, and even the impact of climate on the train when running on a flat track, and it is difficult to calculate the exact result; we usually measure with actual operating conditions through the Davis experience formula [6] to express it:

$$w_0 = A + Bv + Cv^2 \quad (3)$$

Among them, w_0 is the basic resistance coefficient, A , B and C are the resistance polynomial coefficients, which are calculated by the actual test results, and v is the speed of a train.

2.3.2 Additional Resistance

The additional resistance to the train when passing through a non-straight route consists of two parts: the first is the curve resistance of lines and the another is the ramp resistance. We usually introduce the multiparticle model to calculate their values. In the multiparticle model, the additional resistance coefficient of the curve w_c is determined by the following formula:

$$w_c = \frac{c}{R} \cdot \frac{l}{L} \quad (4)$$

Among them, c is an empirical constant that reflects the condition of the curve resistance, R is the radius of curvature, l is the length of the train in the curve section, and L is the total length of the train.

Therefore, the total additional resistance coefficient of the train w_1 is

$$w_1 = w_c + w_i \quad (5)$$

Among them, w_i is the coefficient of unit gradient resistance.

The total running resistance of the train W is

$$W = (w_0 + w_1) \cdot g \cdot M / 1000 \quad (6)$$

Among them, W is the total running resistance of the line, the unit is N ; w_0 is unit basic resistance, the unit is N/kN ; w_1 is unit additional resistance, the unit is N/kN ; g is gravity acceleration, the unit is N/kg ; and M is the total mass of the train, the unit is kg .

2.4 Equation of the Train Motion

The train has four conditions that are traction [7], uniform speed, idle running, and brake during operating. From the perspective of the resultant force C , the status of the train is as follows: Traction condition: $C = F - W$; Uniform speed condition: $C = 0$; Idle running condition: $C = -W$; Brake condition: $C = -(W + B)$. In the equations above, F is the total traction, W is the total running resistance of the line, and B is the braking force.

When the train leaves from the A_n station to the A_{n+1} station, the train is always starting and stopping under the action of resultant force. We consider the running time Δt of the train as the minimum time interval according to the simulation step size, and the force condition and acceleration value of the train have been maintained at the moment Δt of the initial situation within Δt time periods. When the train leaves from the moment t_1 , starting point S_1 , initial speed v_1 to the moment $t_2 = t_1 + \Delta t$, terminal point S_2 , speed v_2 , and then within the time range Δt , we have

$$\begin{cases} a = \frac{C(t_1)}{M} \\ v_2 = v_1 + a\Delta t \\ S_2 = S_1 + \frac{v_2^2 - v_1^2}{2a} \end{cases} \quad (7)$$

Among them, a is the acceleration of the train when running and M is the total mass of the train.

3 Establishment of Timing Energy-Saving Model of the Train

3.1 Time Quasi-objective Function

The longer the train runs in the interval, the lower the energy consumed by the interstation operation, but unlimited time extension is of no practical significance. Therefore, we must limit time when discussing energy-saving situation of the train.

When the train leaves from the station A_n to the station A_{n+1} , the train will experience traction condition a , uniform condition v , idle running condition c , and brake condition b . There are three state turning points: the turning point S_{a-v} of traction condition and uniform speed condition; the turning point S_{v-c} of uniform speed condition and idle running condition; and the turning point S_{c-b} of idle running condition and brake condition. The experience moment of the train in each case is t_a, t_v, t_c , and t_b . Then, the actual total running time $t_{all} = t_a + t_v + t_c + t_b$, the planned running time from the station A_n to station A_{n+1} , is T .

We consider the on-time situation as one of the objective functions; then,

$$f_T(S) = |t_{all} - T| \tag{8}$$

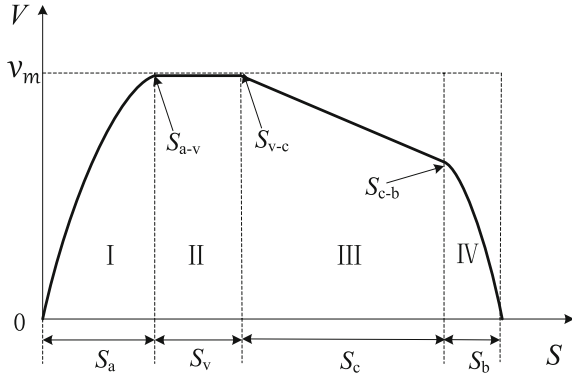
3.2 Fixed-point Parking Objective Function

Because when the metro train is on parking, the error of parking distance between the train and safe screen door on platform must be controlled within the scope of ± 15 cm, which achieved the constraints on fixed-point parking of the train. The fixed-point parking objective function is [8]

$$f_s(S) = |S_n - (S_a^n + S_v^n + S_c^n + S_b^n)| \tag{9}$$

Among them, S_a^n refers to travel distance from the station A_n to station A_{n+1} under traction conditions; S_v^n refers to travel distance from the station A_n to station A_{n+1} under uniform speed conditions; S_c^n refers to travel distance from the station A_n to station A_{n+1} under idle running conditions; S_b^n refers to travel distance from the station A_n to station A_{n+1} under brake conditions; and S_n refers to total travel distance from the station A_n to station A_{n+1} . The typical velocity–distance curve is shown in Fig. 1.

Fig. 1 Typical velocity–distance curve



3.3 Energy-Saving Objective Function

Calculate the energy consumption within a Δt time period according to the equation of the train motion in Sect. 4, then the energy consumed in the whole-line traction is

$$E_{\text{all}} = \sum_0^{T/\Delta t} E_n = \sum_0^{T/\Delta t} v(t_n) \cdot F(t_n) \tag{10}$$

Among them, T is the planned running time from the station A_n to station A_{n+1} . E_n is the consumed energy within the time Δt , $v(t_n)$ is the speed of the train at the moment t_n , and $F(t_n)$ is the traction at the moment t_n .

The brake energy that the train produced is

$$E_{\text{reg}} = (E_{\text{msch}} - E_W) \cdot \eta_{\text{reg}} \tag{11}$$

Among them, η_{reg} refers to the ratio of the regenerative energy that returning the catenary to the total braking energy that produced, and one of the typical values is 95% [9]. Simply subtract the brake regeneration energy from the traction energy consumption to get the total consumption in the interval, the smaller the value, the better the energy-saving effect; the expression is as follows:

$$f_E = E_{\text{all}} - E_{\text{reg}} \tag{12}$$

3.4 Timing Energy-saving Model

After considering three objectives, that is, time punctuality, parking fixed point, and energy saving, we can establish the following objective functions:

$$\begin{aligned}
 & \min \{f_E, f_T, f_S\} \\
 & \left\{ \begin{aligned}
 & \frac{dt}{ds} = \frac{1}{v} \\
 & v \frac{dv}{ds} = F(\mu_f, v) - W(s, v) - B(\mu_b, v) \\
 & t(0) = 0, t(S_n) = T \\
 & v(0) = 0, v(S_n) = 0 \\
 & 0 \leq v \leq V_{\text{line}} \\
 & \left| \frac{dv}{dt} \right| \leq a_{\text{max}} \\
 & \mu_f \in [0, 1], \mu_b \in [0, 1], t \in N^*
 \end{aligned} \right. \tag{13}
 \end{aligned}$$

4 Using Improved Violence Search Method to Solve the Timing Energy-saving Model

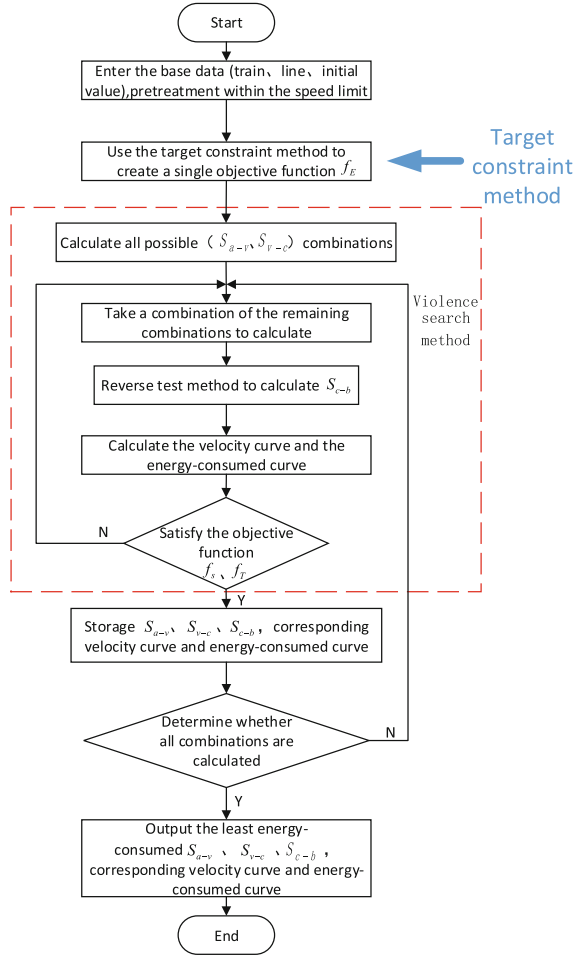
Based on the violent search method, this paper first transforms the multi-objective problem into the single-objective problem through the target constraint method, and then solve the single-objective problem through the violent search method to make higher solution speed of multi-objective optimization problems. The specific application process of the improved violent search method has been shown in Fig. 2.

The calculation process of the improved violent search method in this paper is based on the timing energy-saving model, think of Δt as time step, divided into four calculation steps according to acceleration, uniform speed, idling, and braking. The train’s speed, distance and energy consumption are gradually accumulating in accelerated motion, so we need to judge whether the train arrives at a constant turning point S_{a-v} . When the train runs at a constant speed, we need to calculate both traction energy consumption and brake energy consumption. Under idle running condition, the train slows down naturally only when being under the action of resistance; it is necessary to see if it exceeds the speed limit and reaches the idling turning point S_{c-b} . When the train brakes, the train is only affected by resistance and braking force, and braking force produces the brake energy, the whole-line calculation will not be completed until the train is decelerating to a speed of zero.

5 Simulation Analysis

In order to verify the correctness and accuracy of the model and algorithm, we simulate and calculate with the actual line data of Beijing–Yizhuang Line Rong Jing East Street to Wanyuan Street as input after choosing the master’s thesis of

Fig. 2 The algorithm flow of the improved violent search method



Beijing Jiaotong University authored by Xuesong [10]. Then, we enter the simulation line data and train parameters, and solve the problem in the interval through the improved violent search method. The whole-line “speed–distance” curve of the train from Rong Jing East Street to Wanyuan Street is shown in Fig. 3, and the “power–distance” curve is shown in Fig. 4.

The turning points of the state from Rong Jing East Street to Wanyuan Street are S_{a-v} : 269.6 m; S_{v-c} : 692.1 m; and S_{c-b} : 1030 m. The run time and maximum running speeds from South Village Station to South College Town Station are shown in Table 1.

In the master’s thesis of Beijing Jiaotong University authored by Xuesong Y, Yizhuang line in the actual operation test, we can see that when run time is 97.8 s, energy consumption is 19.09 kWh between stations, and simulation time is 98.3675 s. According to the simulation results in this paper, when we set the

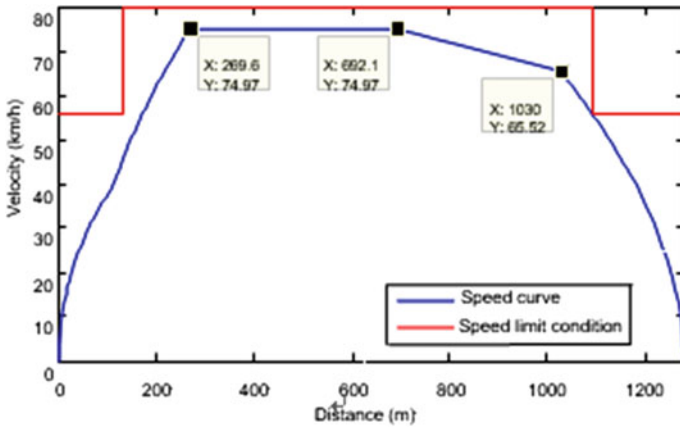


Fig. 3 The optimized “speed–distance” curve of the train from Rong Jing East Street to Wanyuan Street

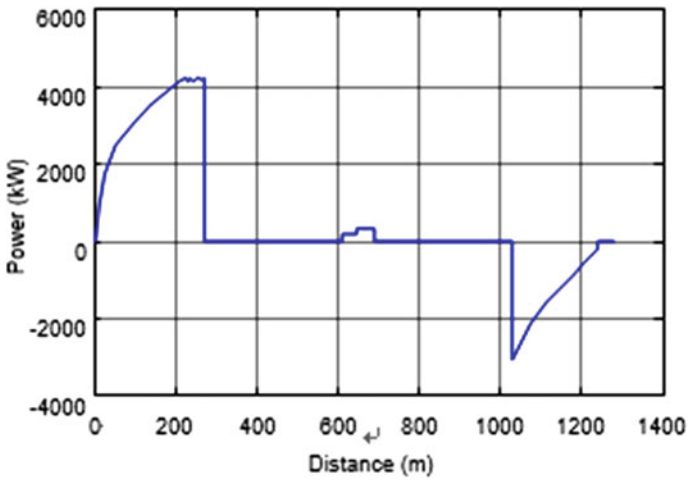


Fig. 4 The optimized “power–distance” curve of the train from Rong Jing East Street to Wanyuan Street

Table 1 Target run from South Village Station to South College Town Station

Data	Numerical value
The optimum speed (km/h)	74.97
Run time (s)	97.9
Traction energy consumption (km/h)	24.03
Brake energy consumption (km/h)	5.31
Global energy consumption (km/h)	18.72

interstation run time, we could see that the optimum result of that run time is 97.9 s, total energy consumption caused by traction in the interval is 24.03 kWh, available braking energy is 5.31 kWh generated in brake process, and actual power consumption is 18.72 kWh.

Suppose that we adopt another method to obtain the two curves, we could need large amounts of iterative operation to get the optimal solutions, when we set multiple parameters by means of some heuristic algorithm that may influence the convergence of the function and the accuracy of optimization. Nevertheless, the improved violence search method converts a multi-objective function into a single-objective function which simplified problems; then, the search scope is further reduced by means of solving constraints; and finally, we directly solve the single-target problems using the violence search method. In the simulation example above, we can accurately judge the constraints of the line using simulation programs and algorithms in a shorter optimization time, and effectively look for a more energy-efficient turning point in the case of the constant run time between stations.

6 Conclusion

This paper presents a new energy-saving research method based on improved violence search method so that we could solve the multi-objective optimization problem of single-train operation. Take the actual data of Yizhuang line in Beijing as an example, and then carry out the simulation research; we obtain “speed–distance” curve and “power–distance” curve after the train has run in the interval by accurate calculation. Time spent in simulation is close to the actual running time, and the energy consumption generated in the interval is slightly higher than simulation results in literature, but it comes closer to the actual test results, which explains that the algorithm presented in this article is more practical and reliable.

Acknowledgements This work is supported by Guangzhou Science and Technology Project (No.201604030061).

References

1. Yanyan Z (2009) Study on traction and energy consumption of urban rail transit system. Beijing Jiaotong University (in Chinese)
2. Howlett P (1996) Optimal strategies for the control of a train. *Automatica* 32(4):519–532
3. Hongguo S (2006) Train running process simulation and optimization research. Southwest Jiaotong University (in Chinese)
4. Ruixia L, Jian Y, Na L, Ruigang S (2012) Development of energy saving method for urban rail vehicles. *J Shanghai Univ Eng Technol* 01:10–14 (in Chinese)
5. Yong W, Sining L (2004) Research on train marshalling scheme based on energy conservation and passenger service. *Urban Rail Transit Res* 06:27–31 (in Chinese)

6. Yanbing X, Dawei M, Lie W (2007) Calculation method of traction energy consumption in trains. *China Railway Sci* 03:84–87 (in Chinese)
7. Jian X (1994) Saving the optimal operating strategy of traction energy consumption. *Mod Energy Conserv* 04:25–29 (in Chinese)
8. Weidong J, Zili W, Chongwei L, Xiantai G, Fan J (1997) Research on optimization method of train energy saving. *J China Railway Soc* 06:59–63 (in Chinese)
9. Lee H, Lee C, Jang G et al (2006) Harmonic analysis of the Korean high-speed railway using the eight-port representation model. *IEEE Trans Power Deliv* 21(2):979–986
10. Xuesong Y (2012) Urban rail transit train energy saving optimization and energy consumption assessment. Beijing Jiaotong University (in Chinese)

Analysis of Running Stability of High-Power Locomotive Under Harmonic Wear Wheel

Qian Xiao, Zhixiang Luo and Jifeng Zheng

Abstract Harmonic wear is one of the most common forms of out-of-round wheel (OOR), especially in heavy locomotives, potentially affecting the running stability. In this paper, a locomotive wheel-rail interaction coupled dynamic model was established for numerical simulation, and common harmonic wear types of locomotive wheels were employed as excitations of the dynamic model. Then, accelerations and Sperling indices of the locomotive in different harmonic wear types were calculated and compared. The result shows that harmonic order has a great influence on locomotive lateral vibration, and has a little effect on the vertical vibration in the same harmonic wave depth; harmonic wave depths have a larger effect on the lateral and vertical Sperling indices in the same harmonic order; compared with no harmonic wear, the lateral stability in different harmonic wear types will change obviously while the vertical stability will not change significantly.

Keyword Locomotive · Wheel–rail interaction · Harmonic wear Stability

1 Introduction

Out-of-round wheel (OOR) is a long-standing problem in railway operations. And harmonic wear is one of the most common forms of OOR, which results in high-frequency vibration and impact of wheel–rail contact. In addition, the locomotive load is particularly large, which seriously affects the locomotive running stability.

A. Johansson and J. C. O. Nielsen studied the influence of polygonal wheels on the dynamic force of wheel and rail by field test and numerical simulation [1]. Zhang Xue-shan et al. analyzed the effect of polygonal wheels on the lateral

Q. Xiao (✉) · Z. Luo · J. Zheng

Key Laboratory of Ministry of Education for Conveyance and Equipment,
East China Jiaotong University, Nanchang 330013, Jiangxi, China
e-mail: jxralph@foxmail.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_39

391

stability of the vehicle [2]. Zhang Jie presented a detailed investigation conducted into the relationships between wheel polygonal wear and wheel/rail noise, and the interior noise of high-speed trains through extensive experiments and numerical simulations [3]. Wang Yijia and others used the established vehicle-track coupling system dynamics model and measured data to study the polygonal wheels on the dynamic performance of the vehicle [4]. Wang Chen found that the torsional vibration of the wheelset was a major factor in the wheel polygonization in the study of the locomotive wheel tread damage [5].

So far, the research on stability caused by OOR has mainly focused on vehicles, while heavy-loaded locomotives have seldom been studied. Once the locomotive appears snake instability, the running quality is deteriorated, leading to strong wheel–rail interaction. A greater force between the wheel and rail accelerates wear and fatigue, damages the railways, and even leads to derailment. Therefore, the analysis of locomotive running stability is of great significance.

2 Harmonic Wear Wheels of Locomotive and Calculation of Riding Indices

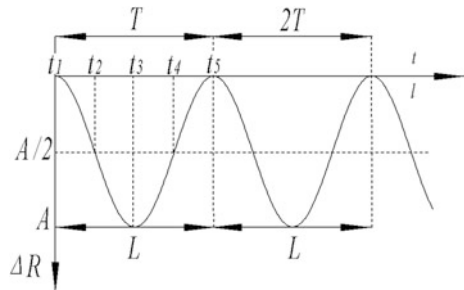
Wheel circumference of circular harmonic wear can be used a method of the harmonic wave function to define [6], as shown in Fig. 1, at one cycle of the wheel rolling, the wheel diameter difference is considered as a harmonic function described by the following formulas [7]:

$$\Delta R = \begin{cases} \frac{A}{2}(1 - \cos \frac{2\pi}{T}t) & 0 \leq t \leq nT \\ 0 & nT < t \leq \frac{2\pi}{\omega} \end{cases} \quad (1)$$

$$\frac{t}{T} = \frac{R\varphi}{L} \quad (2)$$

Where A is wave depth; t is wheel rolling time; T is a harmonic wear cycle; n is the order of harmonic wear; ω is rolling angular velocity; L is wave length; φ is the phase angle.

Fig. 1 Schematic diagram of harmonic wear



The radius of the rolling circle of the harmonic wear wheel varies with the rolling angle as shown in Fig. 2, described as:

$$\begin{cases} R(\varphi) = R_0 + \Delta R(\varphi) \\ \varphi(t + \Delta t) = \varphi(t) + \omega \Delta t, \end{cases} \tag{3}$$

where R_0 is nominal rolling radius; φ is phase angle; ω is rolling angular velocity; Δt is time interval.

Vertical_Sperling_Index and Lateral_Sperling_Index calculate Sperling’s riding indices for vertical and lateral directions [8]. The riding index for the lateral direction is calculated as

$$W_z = 10 \sqrt{\int_{0.5}^{30} a^3 B_w^3 df}, \tag{4}$$

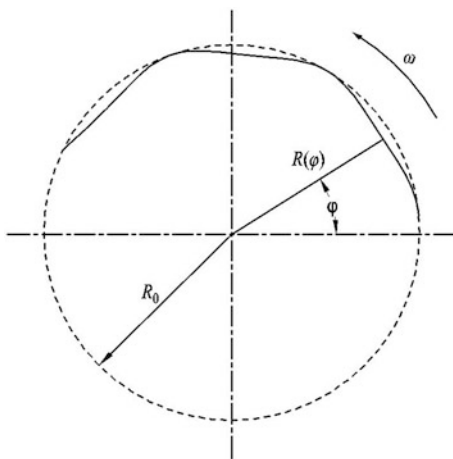
where a is acceleration in the frequency domain (cm/s^2), f is the frequency (Hz), B_w is defined by the following equation:

$$B_w = 0.737 \left(\frac{1.911f^2 + (0.25f^2)^2}{(1 - 0.277f^2)^2 + (1.563f - 0.0368f^3)^2} \right)^{1/2}. \tag{5}$$

The ride comfort index for the vertical direction is calculated as

$$W_z = 10 \sqrt{\int_{0.5}^{30} a^3 B_s^3 df}, \tag{6}$$

Fig. 2 Schematic diagram of rolling radius of harmonic wear wheel



where B_s is defined as

$$B_s = 0.588 \left(\frac{1.911f^2 + (0.25f^2)^2}{(1 - 0.277f^2)^2 + (1.563f - 0.0368f^3)^2} \right)^{1/2}, \quad (7)$$

3 Establishment of Numerical Model

With high-power locomotive as the research object, a model of locomotive was established with the multibody dynamics software UM referring to locomotive structure topology diagram as shown in Fig. 3 [9]. This model was simplified into three components, namely, car body, bogie, and wheelset, and each was regarded as rigid. The model consisted of 50 degrees of freedom. Many nonlinearities, such as wheel/rail contact geometry, lateral stop, and primary suspension, were considered. German railway spectra of low irregularity were set as the track irregularity. Table 1 shows the values of the simulation parameters involved in the UM model. The resulting model is shown in Fig. 4.

4 Results and Analysis

4.1 Lateral and Vertical Vibration Accelerations

Taking the second wheel of the locomotive front bogie as an example, the order of harmonic is 1, 6, 11, and 15, and the wave depth is 0.1 and 0.3 mm, which runs continuously for 7 seconds. The comparisons of lateral vibration acceleration in different harmonic orders are shown in Figs. 5 and 6. In the case of the same

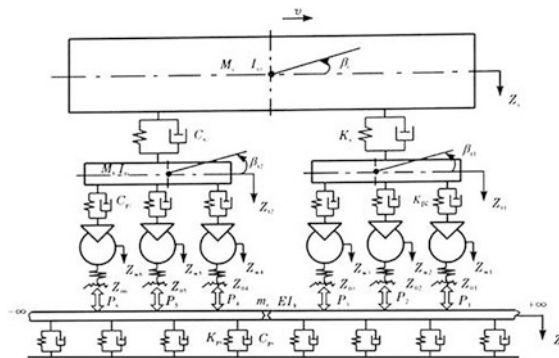


Fig. 3 Locomotive structure topology diagram

Table 1 Parameters involved in the UM model

Parameter	Value	
Axle load	24 t	
Wheelset mass	3400 kg	
Wheel profile	JM3	
Rail profile	60 kg/m	
Primary suspension	Longitudinal stiffness of steel spring	750,000 N/m
	Lateral stiffness of steel spring	750,000 N/m
	Vertical stiffness of steel spring	785,000 N/m
Secondary suspension	Longitudinal stiffness of air spring	65,000 N/m
	Lateral stiffness of air spring	65,000 N/m
	Vertical stiffness of air spring	561,000 N/m
Anti-yaw damper	80,000 N/m	

Fig. 4 Parameterized high-power locomotive model setup in UM

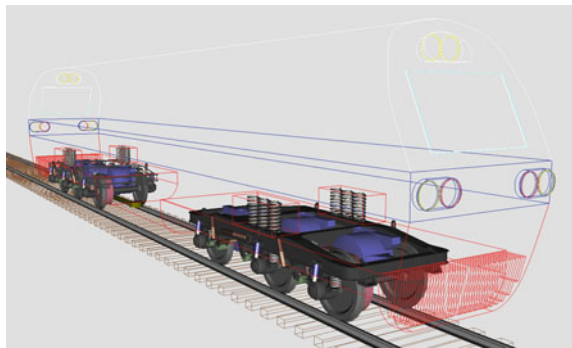


Fig. 5 Lateral vibration acceleration of different orders in 0.1 mm wave depth

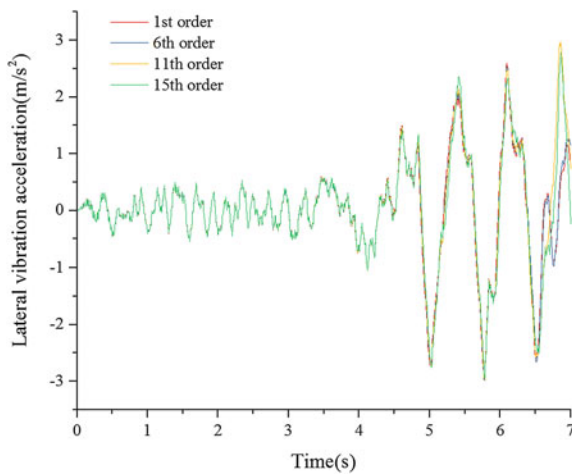
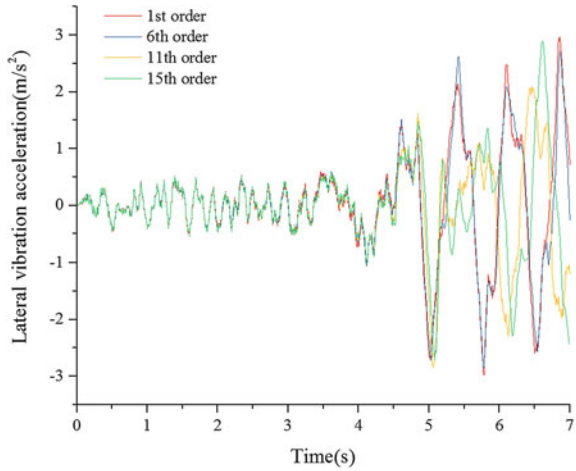


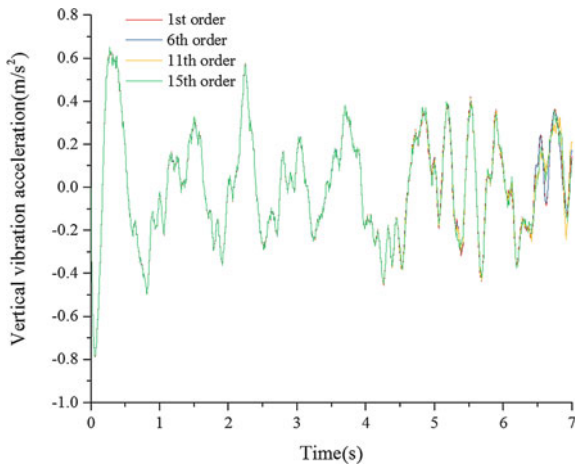
Fig. 6 Lateral vibration acceleration of different orders in 0.3 mm wave depth



harmonic wave depth, with the extension of the running time, the greater the harmonic order, the greater the lateral vibration acceleration changes, indicating the different harmonic order of the same wave depth has a great influence on the lateral stability.

Similarly, the comparisons of vertical vibration acceleration in different harmonic orders are depicted in Figs. 7 and 8. With greater harmonic order, the vertical vibration acceleration does not change a lot in the case of the same harmonic wave depth, which demonstrates harmonic orders have little effect on locomotive vertical vibration.

Fig. 7 Vertical vibration acceleration of different orders in 0.1 mm wave depth



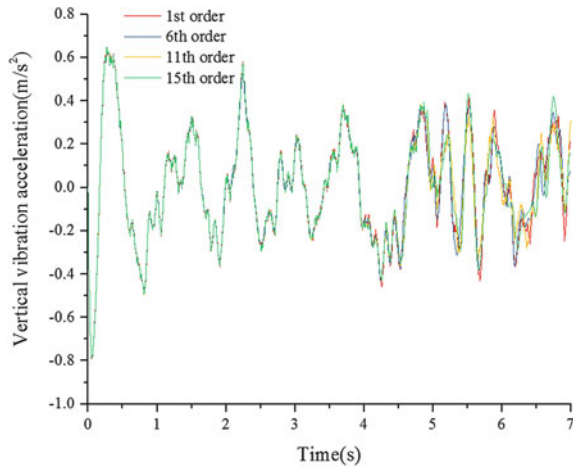


Fig. 8 Vertical vibration acceleration of different orders in 0.3 mm wave depth

4.2 Lateral and Vertical Sperling Indices

The lateral and vertical Sperling indices of the same order in different wave depth are shown in Table 2. At the same harmonic wheel wear order, the Sperling indices value of the lateral and vertical decrease with the increase of the wave depth, which indicates that the wave depth of the different harmonic wear in the same order has a great influence on the Sperling indices of the locomotive both in lateral and vertical.

Table 2 Lateral and vertical Sperling indices in different harmonic wear types

Sperling indices	Normal wheel	Wave depth (mm)	The order of harmonic wear			
			1	6	11	15
Lateral	3.052	0.1	3.052	3.05	3.049	3.047
Vertical	2.321		2.321	2.321	2.322	2.322
Lateral	3.052	0.3	3.05	3.042	3.04	2.983
Vertical	2.318		2.318	2.317	2.318	2.317

5 Conclusion

- (1) In the same harmonic wave depth, harmonic order has a great influence on the lateral vibration of the locomotive, and has a little effect on the vertical vibration of the locomotive.
- (2) In the same harmonic order, harmonic wave depths have a larger effect on the lateral and vertical Sperling indices.
- (3) Compared with no harmonic wear, the lateral stability of the locomotive in different harmonic wear types will change obviously while the vertical stability will not change significantly.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 51565013).

References

1. Johansson A, Nielsen JO (2003) Out-of-round railway wheels—wheel—rail contact forces and track response derived from field tests and numerical simulations. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 217(2):135–146
2. Zhang X, Xiao X, Jin X (2008) Influence of high speed railway wheels ovalization on vehicle lateral stability. *Chin J Mech Eng* 44(3):50–56 (in Chinese)
3. Zhang J, Han G, Xiao X, et al (2014) Influence of wheel polygonal wear on interior noise of high-speed trains. *J Zhejiang University A*, 15(12):1002–1018 (in Chinese)
4. Wang Y, Zeng J, Luo R et al (2013) Effect of polygonal wheel on vehicle dynamic performance. *J Sichuan Univ (Eng Sci Ed)* 45(3):176–182 (in Chinese)
5. Wang C, Ma W, Luo S et al (2016) Study on damage mechanism of locomotive vehicle tread. *J Vibr Measur Diagn* 36(5):890–896 (in Chinese)
6. Xiao Q, Zhang H, Zhou X (2016) Influence of wheel harmonic wear on creep characteristics of high speed wheel-rail contact on straight line. *China Railway Sci*, 37(6):60–68 (in Chinese)
7. Pogorelov D (2010) Universal mechanism user's manual. Bryansk State Technical University, Bryansk
8. Garg V (1984) Dynamics of railway vehicle systems. Academic Press
9. Zhai W (2015) Vehicle-track coupled dynamics, 4th edn. Science Press (in Chinese)

Detection of Wheel Tread Wear Based on Laser Displacement Sensor

Jianyu Zhang, Zhi Yang, Yong Zhang and Zongyi Xing

Abstract In order to solve the problems of heavy workload and low accuracy existing in manual measurement, a method based on laser displacement sensor (LDS) is proposed to detect wheel tread wear. First, the 2D laser displacement sensors installed on the lateral side of the track are used to collect the wheel tread data. Then, techniques including data preprocessing, coordinate transformation, data fusion, and so on are applied to process the wheel tread data combined with standard wheel tread data. Finally, the wheel tread wear can be calculated by using geometric relationship. Error analysis and field experiments show that the accuracy of the proposed method is less than 0.2 mm and has the strong anti-interference ability. The proposed method can greatly meet the actual requirements of measurement of wheel tread wear.

Keywords Wheel tread wear · LDS · Error analysis · Standard wheel tread data

1 Introduction

Wheel tread wear is an important parameter for the safe operation of urban rail vehicles. The inspection of wheel tread wear is of great importance to the safety of train operation. With the increase of train mileage, uneven wheel tread wear and increased wear rate lead to larger wheel diameter change. It reduces the safety and comfort of train operation [1]. Therefore, it is very important to detect the wheel tread wear at regular intervals, obtain accurate wear value, and determine whether it is overrun or not to ensure the safe operation of the train [2, 3].

At present, wheel tread wear detection can be divided into two methods: contact measurement and non-contact measurement [4]. Contact inspection is mainly based on wheel diameter measuring instrument, and the detection accuracy is extremely

J. Zhang · Z. Yang · Y. Zhang (✉) · Z. Xing
School of Automation, Nanjing University of Science and Technology,
No.200 Xiao Ling Wei, Xuanwu District, 210094 Nanjing, China
e-mail: 34445721@qq.com

affected by human factors. However, because of its simple operation, it is still the main method of tread wear measurement. Non-contact detection is mainly divided into image method and laser method. In the image method, the wheel tread is photographed by CCD camera, and the wheel tread profile information is obtained by image processing, and the wheel tread wear detection is realized [5], but the system structure is complex and the anti-interference ability is low. With the rapid development of laser technology, the research on tread wear detection of urban rail vehicles has made great progress. Haichun and Jing et al. [6, 7] put forward a portable wheel measuring instrument based on laser displacement sensor, through the combination of laser sensor and linear stepper motor to obtain discrete coordinate data and get the contour wheel size parameters, but the device cannot effectively extract the wheel, the detection precision is low. Qibo et al. [8, 9] proposed a wheel diameter detection method based on the laser triangulation principle. The wheel diameter was detected by a single laser sensor or two laser sensors. Haiyu and Changgeng et al. [10, 11] proposed a rim size online detection method based on 2D laser displacement sensor. By LDS which is arranged in the through rail on both sides of track, the wheel contour will be detected and the flange size parameters will be obtained.

A method for measuring wheel tread wear based on LDS is proposed in this paper: First, the 2D laser displacement sensors installed on the lateral side of the track are used to collect the wheel tread data. The actual contour of the wheel is reconstructed according to the standard wheel contour, and then the wheel flange height is obtained from the geometric relationship of the wheel contour, so the wheel tread wear detection is realized.

2 Tread Wear Detection Algorithm

When the wheel axle position sensor detects the train through the detection system, triggering LDS on electric work surface profile data on sensor wheel, and then the data after the data preprocessing, coordinate rotation, data fusion, and optimal selection algorithm for tread wheel tread wear detection.

2.1 Data Preprocessing

When the wheel axle position sensor detects the train through the system, the 2D laser displacement sensor based on 50 Hz sampling frequency of the wheel tread contour data acquisition, data acquisition of the wheel, effective data extraction 6 carriage 48 wheels corresponding to each wheel. It is found in practical engineering applications: When the train operator performs the braking operation, the sensor will detect the wheel brake and bogie vehicle device, this kind of data and data analysis of the influence of wheel profile independent of wheel profile data, should be in the data segment when the data filtering, data preprocessing.

2.2 Coordinate Rotation

In the detection system, the LDS has a deflection angle β between the laser line and the rail. The wheel tread data acquired by the sensor will inevitably be distorted, so that the original data of the sensor can be rotated in coordinate to correct the deformed tread contour curve. According to the actual position parameter value of the sensor, the coordinate transformation of the sensor's own coordinate system from xoy to the baseline coordinate system uov is realized by coordinate rotation [12]. The coordinate rotation diagram is shown in Fig. 1, and the laser displacement sensors are rotated in accordance with the formula. The coordinate rotation diagram is shown in Fig. 1, and the laser displacement sensors are rotated in accordance with the formula 1.

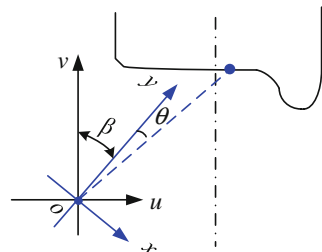
$$\begin{cases} u_n = \sqrt{(x_n)^2 + (y_n)^2} \sin(\theta + \beta) = x_n \cos \beta + y_n \sin \beta, \\ v_n = \sqrt{(x_n)^2 + (y_n)^2} \cos(\theta + \beta) = y_n \cos \beta - x_n \sin \beta. \end{cases} \quad (1)$$

In the formula, (x_n, y_n) is the coordinate of the detection point on the sensor's own coordinate system xoy . θ is the angle between the probe point of the sensor and the y axis. (u_n, v_n) is the coordinate value of the sensor detection point after the coordinate rotation in the tread reference coordinate system uov .

2.3 Data Fusion

This paper uses a set of 2D laser displacement sensor can obtain the wheel wear tread contour and the outside of the wheel baseline cannot detect the rim and medial line, and rarely wear top rim, shape remains unchanged, so it is necessary to wheel standard profile and sensor data acquisition of wheel tread data fusion to obtain complete wheel profile curve. The tread reference coordinate system uov of the LDS and the wheel and wheel standard contour coordinate system pwq are related to the coordinate system xoy , as shown in Fig. 2.

Fig. 1 Coordinate rotation transformation process



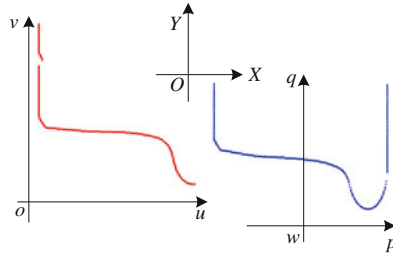


Fig. 2 Sketch map of each coordinate system

Data fusion is mainly divided into three parts: coordinate movement, baseline fusion, and data fusion.

2.4 Best Tread Selection

When the wheel is passing through the sensor’s effective detection area, the system will collect a plurality of tread wear data, and the key tread data needs to be selected to improve the system accuracy. When the wheel deviates from the middle, that is, the wheel detection surface does not pass through the wheel center, the measured tread shape will have a certain tensile deformation.

Assume that the wheels in a sampling time, laser detection sensor surface without the wheel center, but in the next sampling time than the round of heart, that is two times the sampling interval in sensor surface has been round of heart, as shown in Fig. 3, the laser detection surface after wheel center measuring flange height is BC, the height missing the wheel center is B₁C₁, and B₁C₁ > BC, so this

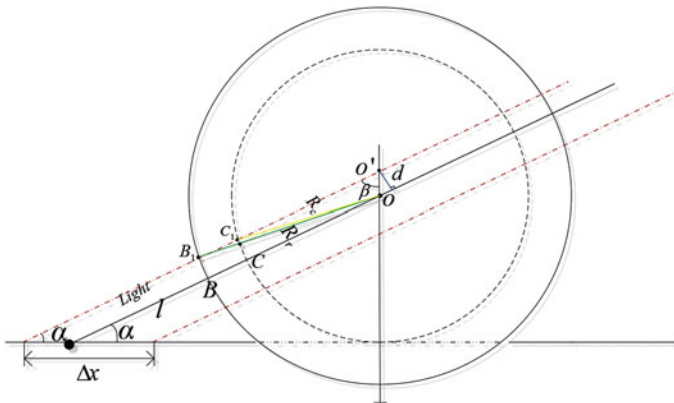


Fig. 3 Relation between laser line and ideal wheel center at adjacent sampling time

paper selects the minimum value of the high rim data set as the best tread data for the calculation of wheel tread wear.

3 Error Analysis of Wheel Flange Abrasion Detection

The error of the wheel wear detection system of LDS is composed of sensor deflection angle, laser detection surface missing wheel center and data fusion based on error. In order to meet the on-site detection precision of 0.2 mm, using the computer simulation method, the wheel tread wear analysis. The flange height of the wheel used for the simulation is 28.6 mm, and the standard flange height is 28 mm, that is, the tread wear is 0.6 mm.

3.1 Sensor Deflection Angle

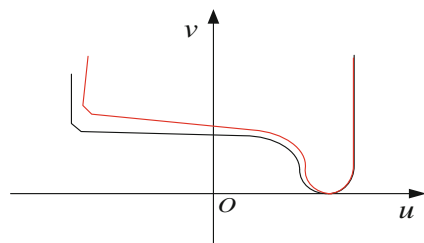
In the process of tread wear detection, it is necessary to coordinate the rotation of the sensor output data, in which the deflection angle β of the sensor is the key parameter for coordinate rotation. Due to the uncertainty of the machining accuracy and installation of the sensor base, the deviation of the actual deflection angle of the sensor will be biased, that is to say, there is a deviation of the deflection angle $\Delta\beta$. When the deflection angle of the sensor has a $\Delta\beta$ deviation, the output of the sensor is rotated by coordinate, and the overall offset and dislocation occur on the tread, as shown in Fig. 4, the tread wear test will have a greater impact.

Assuming that the sensor accuracy and other factors remain unchanged, only the deviation of the sensor deflection angle is changed, and the computer simulation is used to analyze the tread wear error. The deflection angle of the sensor is in the range of $[-0.1^\circ, 0.1^\circ]$, and the tread wear F_h is simulated at 0.01° .

When the deflection angle of the sensor exists deviation, the tread wear is positively related to the deflection angle, and the relationship between the tread wear deviation and the deflection angle deviation is fitted by the first order curve:

$$\Delta F_h = 2.6156\Delta\beta - 0.011. \tag{2}$$

Fig. 4 Overall misalignment of profile caused by deviation of sensor deflection angle



Therefore, in order to ensure that the tread wear detection error is less than 0.2 mm without considering other factors, the deviation of the sensor deflection angle should be guaranteed to be within the range of $\pm 0.03^\circ$.

3.2 The Detecting Surface Missing the Center of Wheel

According to the requirements of the detection process flange size, surface emitting laser detection and laser sensor through the round of heart [13], but in the actual project detection data of selected laser line to the wheel center distance is the shortest of the flange height calculation can lead to tension effect, influence of tensile effect on the flange height detection.

In Fig. 3, Δx is the distance in the interval of the sampling, O represents the ideal wheel center point laser detection surface over the wheel center, O' represents the reference center of laser detection surface missing the center of the wheel. So, Flange height error e caused by the tensile effect is as follows:

$$e = B_1C_1 - BC \quad (3)$$

In the formula, according to the geometric relation, B_1C_1 and OO' can be expressed as follows:

$$B_1C_1 = O'B_1 - O'C_1 = \sqrt{R_r^2 - OO'^2 \cos^2 \alpha} - \sqrt{R_c^2 - OO'^2 \cos^2 \alpha} \quad (4)$$

$$OO' = \frac{d}{\sin \alpha} \quad (5)$$

Simultaneous (3) and (4), the flange height error can be expressed as

$$e = \sqrt{R_r^2 - \frac{d^2}{\tan^2 \alpha}} - \sqrt{R_c^2 - \frac{d^2}{\tan^2 \alpha}} - R_r + R_c \quad (6)$$

In the formula, the distance d takes the maximum value of d_{\max}

$$d_{\max} = \Delta x \cdot \sin \alpha / 2.$$

In the actual use of the wheel, the flange height parameter range is [28–35 mm]. Therefore, Fig. 5 shows the corresponding description between the distance d and the wheel flange height error e in the case of different flange heights. In the figure, the top curve describes the corresponding 35 mm of the flange height, and the bottom curve describes the corresponding 28 mm of the flange height. When the distance $d = 0$, the flange height error $e = 0$, along with the distance d increases, the wheel flange error e also increases along with it.

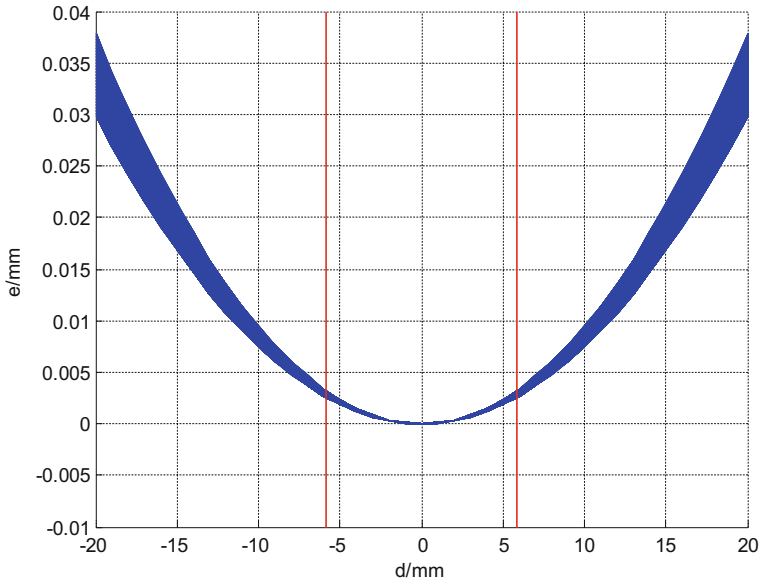


Fig. 5 Relationship between laser line to ideal wheel center distance D and measurement error of wheel rim height

The tread wear test system has a sampling frequency of 50 Hz and a speed of 3 km/h. The maximum distance of the system is $d = 5.869$ mm, and the maximum flange height is measured at 0.0033 mm.

3.3 Data Fusion

In the process of data fusion, ΔX will have horizontal distance deviation ΔD_x , and ΔY will have vertical distance deviation ΔD_y . When there is delta ΔD_x and deviation ΔD_y , the wheel tread data is stretched along the horizontal and vertical directions, which affects the accuracy of wheel rim height detection.

Assuming that the deflection angle and other factors remain the same, there is only data fusion error, and computer simulation is used to analyze the tread wear error.

(1) Error analysis of horizontal distance deviation of sensor

The horizontal distance deviation of ΔD_x in the range of $[-0.2 \text{ mm}, 0.2 \text{ mm}]$, with 0.02 mm simulation calculation of tread wear the error ΔF_h , the calculation results are shown in Table 1. As shown in Table 1, when the horizontal distance deviation exists in data fusion, the error of tread wear is 0.01 mm, so that the influence of horizontal distance deviation on tread wear is negligible.

Table 1 Effect of D_X on the horizontal distance deviation of the tread wear

ΔD_X	-0.2	0.18	-0.16	-0.14	-0.12	-0.1	-0.08
ΔF_h	0	0	0	0	0	0	0
ΔD_X	-0.06	-0.04	-0.02	0	0.02	0.04	0.06
ΔF_h	0	0	0	0	-0.01	-0.01	-0.01
ΔD_X	0.08	0.1	0.12	0.14	0.16	0.18	0.2
ΔF_h	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01

(2) Error analysis of sensor vertical distance deviation

The deviation of the vertical distance is ΔD_Y 0.05 mm in the range [-0.5 mm, 0.5 mm], and the tread wear error ΔF_h is calculated by the simulation of the step size. The result is shown in Table 2. From Table 2, the maximum deviation of tread wear is 0.25 mm when there is a deviation of vertical distance in data fusion, and the error difference is positively related to the deviation of disposal distance.

The first-order curve fitting is used to fit the relationship between the height deviation of the wheel flange ΔF_h and the deviation of the distance ΔD_Y as follows:

$$\Delta F_h = 0.4899\Delta D_Y - 0.0043. \tag{7}$$

Therefore, the deviation of the vertical distance should be within the range of ± 0.40 mm in order to ensure that the tread wear detection error is less than 0.2 mm without considering other factors.

4 Experiment and Analysis

In order to verify the effectiveness and accuracy of wheel tread wear test method, field tests were carried out on the Guangzhou Chisha depot, including standard wheelset test and vehicle passing test. The tread wear value is standard, the height of wheel flange is different from the actual wheel flange height difference. Therefore, the tread wear detection accuracy is equivalent to the flange height detection precision. For the convenience of study, and the experimental results are analyzed from wheel flange height.

Table 2 Effect of D_Y on the vertical distance deviation of the tread wear

ΔD_Y	-0.5	-0.45	-0.4	-0.35	-0.3	-0.25	-0.2
ΔF_h	-0.25	-0.22	-0.2	-0.18	-0.15	-0.13	-0.1
ΔD_Y	-0.15	-0.1	-0.05	0	0.05	0.1	0.15
ΔF_h	-0.08	-0.05	-0.03	0	0.02	0.04	0.07
ΔD_Y	0.2	0.25	0.3	0.35	0.4	0.45	0.5
ΔF_h	0.09	0.12	0.14	0.17	0.19	0.22	0.24

In the standard wheel experiment, the standard of wheelset flange height is 28 mm, the standard wheel placed in orbit, the artificial detection system through the slow wheel drive along the same direction, a total of 7 times of measurement system, the experimental results are shown in Table 3.

Table 3 obtained the standard wheelset experimental system, the maximum error was 0.12 mm, and the experimental results meet the tread wear detection accuracy of ± 0.2 mm.

In the experiment, the Guangzhou Metro Line 8 A2 train B, C two carriages of 8 wheels 7172 times high rim detection, multiple measurement results mean as the measurement system of the wheel rim of high value, the measurement results are shown in Fig. 6. In the graph, the difference between the mean measured value of the flange height and the artificial measurement mean is within the range of 0.2 mm. It is proved that the method proposed in this paper meets the requirements of field inspection.

Table 3 Experimental results of standard wheelset

Times	1	2	3	4	5	6	7
Height/mm	28.13	28.06	28.05	27.97	28.01	27.93	27.88
Error/mm	0.13	0.06	0.05	-0.03	0.01	-0.07	-0.12

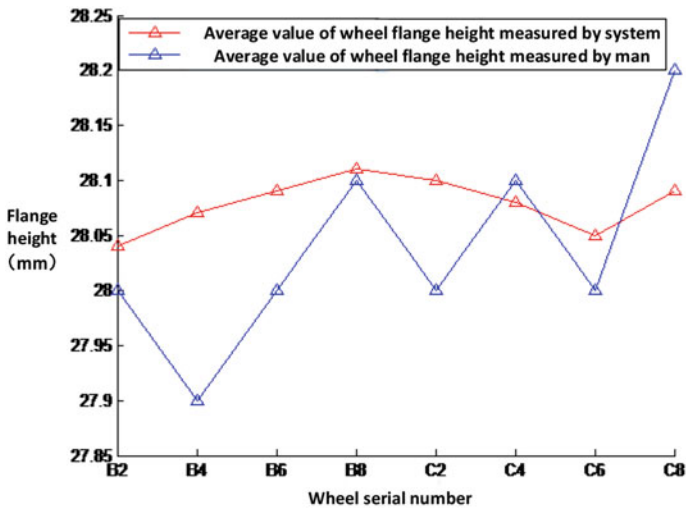


Fig. 6 Comparison of mean measurements of flange height system with artificial mean

5 Conclusion

In this paper, a method of wheel wear detection based on LDS is proposed to realize high precision measurement of wheel tread wear. By a group of 2D laser, displacement sensor is installed on the outer track detection of wheel tread contour, using coordinate rotation, data preprocessing, data fusion, and optimal selection algorithm of wheel tread actual contour reconstruction, and wheel flange height through the wheel profile geometry, and wheel tread wear value. The standard wheelset test and over vehicle test show that the proposed method can detect accuracy up to +0.2 mm, and can meet the needs of onsite inspection and repair.

Acknowledgements This work is supported by National Key R&D Program of China under 286 Grant (2017YFB1201201).

References

1. Yu F, Huadong M, Qi Z (2010) Analysis of abnormal tread on wheel tread of Shanghai Metro Line 3. *Locomotive Electr Drive* 2010(02):45–46 (in Chinese)
2. Kaihua W, Jianhua Z, Zuohua H, Heng Z (2004) Photoelectric image inspection method for wheel flange wear of wheel set. *Optics* 2004(05):535–537 + 540 (in Chinese)
3. Huijie Z, Mingwei P (2010) Discussion on tread wear law of wheel train of Metro. *J Dalian Jiaotong Univ* 2013(02):6–9 (in Chinese)
4. Zhaoyi S, Haiqiang Z, Zongyi X, Xiaohao W (2010) Research on measurement method of wheel flange size based on laser displacement sensor. *Mechatron Eng* 2015(06):793–797 (in Chinese)
5. Yan G (2014) Research on automatic and dynamic measurement method of wheelset geometry. Beijing Jiaotong University, Beijing (in Chinese)
6. Yang X, Li X, Gao Z, et al (2014) A cooperative scheduling model for timetable optimization in subway systems. *J IEEE Trans Intell Transp Syst* 2013, 14(1):438–447 (in Chinese)
7. Jing Y, Haichun X, Zongyi X (2017) Portable wheel set measuring instrument based on laser displacement sensor. *Instrum Technol Sens* 2017, (04):41–44 (in Chinese)
8. Qibo F, Shiqian C, Jianying C, Fengshan L, Yingjie Z (2008) Dynamic measurement system of wheelset geometric parameters. *China J Railway Sci* 2008, (05):138–144 (in Chinese)
9. Zhifeng Z (2008) Study on the method and system for on-line dynamic measurement of wheelset geometric parameters. Beijing Jiaotong University, Beijing (in Chinese)
10. Haiyu L, Xiaoqing C, Zhaoyi S, Zongyi X, Xiaohao W, Fei W (2016) On-line measurement system of wheelset size of urban rail vehicle based on laser displacement sensor. *Urban Rail Transit Res* 2016, (09):10–14 (in Chinese)
11. Changgeng W, Zongyi X, Wen H (2016) A new on-line measurement system for wheelset dimensions of metro vehicles. *Railway Comput App* 2016, (05):62–65 (in Chinese)
12. Xiaohao W, Zongyi X, Zhaoyi S, Zhaoxin L, Xudong G (2016) Contour fitting method of wheelset based on Lagrange multiplier method. *J Urban Rail Transit* 2015, (07):35–39 (in Chinese)
13. Kaihua W, Jianhua Z, Zuohua H (2005) Research on non-contact photoelectric inspection method for wheel tread wear of wheel pair. *Opto Electron Lasers* 2005, 16(3):332–335 (in Chinese)

A Hybrid Filtering Algorithm for Pantograph Image Denoising

Weiwei Pei, Zongyi Xing and Zhuang Chen

Abstract In the pantograph image detection method, the pantograph image is affected by electromagnetic interference and transmission medium, so there are some noises in the image, which are mainly the gauss noise and impulse noise. For the pantograph image denoising problem, the single denoising method is not effective. Generally, two types of filters are used to filter the noise step by step. In this paper, a hybrid noise filtering algorithm is proposed. First, ABDND-based median filtering algorithm is used to suppress the impulse noise. Then, an improved mean filtering algorithm is used to suppress the Gaussian noise. Two algorithms are used to filter out the hybrid noise in the pantograph image. Finally, the experimental results show that the proposed hybrid filtering algorithm can effectively remove the noise and protect the edge of the image, which is superior to the traditional filtering algorithm.

Keywords Pantograph · Gaussian noise · Impulse noise · ABDND
Mean filtering

1 Introduction

The pantograph is a replaceable equipment that is installed on the roof of an urban rail vehicle to get the current from the contact wire. When the train is running, the pantographic slipper is directly connected to the contact wire and this may generate the electric arc, which may even cause excessive abrasion, rupture, and so on. So the pantograph image detection method has been widely used and the high-quality image is one of the key factors that affects the detection precision.

With the development of image processing technology, the noise filtering problem has been concerned by scholars at home and abroad for a long time.

W. Pei · Z. Xing (✉) · Z. Chen
School of Automation, Nanjing University of Science
and Technology, Nanjing 210094, China
e-mail: xingzongyi@njjust.edu.cn

Xiancheng [1] proposed two different denoising methods for Gaussian-impulse hybrid noise filtering problem. However, when the variance of Gaussian noise and the density of impulse noise change, the two methods show different advantages, so the noise density needs to be measured first which is very complex. Considering the advantages of median filtering algorithm and mean filtering algorithm to solve the hybrid noise filtering problem in a medical image, Mengqiang [2] proposed a filtering algorithm for hybrid noise, which saves more detailed information of these images, but this algorithm is mainly for static target images. Liu [3] proposed a weighted norm energy minimization model to suppress the hybrid noise and verified the feasibility of this method by experimental results, but the algorithm is more complex and inefficient. In this paper, a hybrid noise filtering algorithm is proposed to suppress the impulse noise and Gaussian noise. ABDND-based median filtering algorithm is used to suppress the impulse noise, the improved mean filtering algorithm is used to suppress the Gaussian noise. Finally, the experimental results show that the proposed in this paper can effectively suppress the noise, which is superior to the traditional filtering algorithm [4–7].

2 A Brief Introduction to ABDND-Based Filtering Algorithm and Improved Mean Filtering Algorithm

2.1 ABDND-Based Filtering Algorithm

ABDND-based filtering algorithm [8–10] is a kind of improved adaptive switching median filtering (ASMF) algorithm, which has a better detection effect for random impulse noise. ABDND-based algorithm mainly includes the following three detection stages.

2.1.1 The Noise Pixel Detection Stage

At this stage, the gray histogram and its difference value should be got first, and then the maximum gray value and the minimum gray value of the $W \times W$ sliding window is computed to obtain the noise identification matrix of the image to decide whether the pixel is a noise pixel.

2.1.2 The Noise Pixel Verification Stage

At this stage, if the number of non-noise pixels in $w \times w$ sliding window ($w < W$) is less than w , then increase w progressively and traverse all the pixels, otherwise compute difference values between the non-noise pixels and the center temporary noise pixels, and the sum of these values is denoted by d . If d is less than the set

threshold value, this pixel will not be considered as a noise pixel, and noise identification matrix should be modified at the same time, otherwise, this pixel will be considered as a noise pixel.

2.1.3 The Noise Pixel Filtering Stage

Through the above two stages, the noise pixels and non-noise pixels are all found. For non-noise pixels, there is no need to filter, and for noise pixels, the median filtering algorithm should be used for non-noise pixels in the filtering window.

2.2 Improved Mean Filtering Algorithm

For the Gaussian noise, the traditional mean filtering algorithm uses an $M \times N$ sliding window to compute the mean of pixels to instead the target window center pixel value [11, 12]. This algorithm can suppress the Gaussian noise, while the image will be blurred and the edge information will be lost.

Based on the traditional mean filtering algorithm, this paper proposed a local removing extreme value-weighted mean filtering algorithm. First, the maximum and minimum values in the sliding window should be removed, and then the weighted and normalized remaining pixels are used to replace the central pixels.

3 Performance Analysis of Filtering Algorithm

Based on MATLAB platform, this paper verified the effectiveness of the proposed algorithm. Because the image is not an ideal noiseless image, the laboratory image is used as the analysis object. The pantograph image is shown in Fig. 1.

Fig. 1 The laboratory pantograph image



The performance of the filtering algorithm is evaluated by combining the subjective and objective evaluation methods. Subjective evaluation mainly contrasts the filtering performance of different algorithms and the objective evaluation mainly uses the mean square error (MSE) and peak signal-to-noise ratio (PSNR) to evaluate the performance of the image filtering algorithm.

The original image and the filtered image are denoted as X and \hat{X} , respectively, the image size is $I \times J$, and the mean square error (MSE) and the peak signal-to-noise ratio (PSNR) [3] are defined as (1) and (2).

$$X = \{x(i,j) \mid 1 \leq i \leq I, 1 \leq j \leq J\}, \quad (1)$$

$$\hat{X} = \{\hat{x}(i,j) \mid 1 \leq i \leq I, 1 \leq j \leq J\}. \quad (2)$$

The mean square error (MSE)

$$\text{MSE} = \frac{1}{I \times J} \sum_{i=1}^I \sum_{j=1}^J [x(i,j) - \hat{x}(i,j)]^2. \quad (3)$$

The peak signal to noise ratio (PSNR)

$$\text{PSNR} = 10 \lg \frac{\text{MAX}_I^2}{\text{MSE}} \text{dB} \quad (4)$$

Where MAX_I is the maximum value of the pixel color.

3.1 The Single Impulse Noise Filtering

The performance of traditional median filtering algorithm and ABDND-based algorithm are compared through the following images with different impulse noise densities, and the filtering results are shown in Figs. 2 and 3, where (a) is the noise image, (b) is the median filtering algorithm filtered image, (c) is the ABDND-based algorithm filtered image.

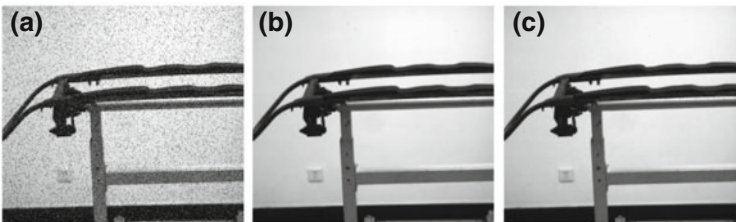


Fig. 2 Variance of 0.1 impulse noise image processing results

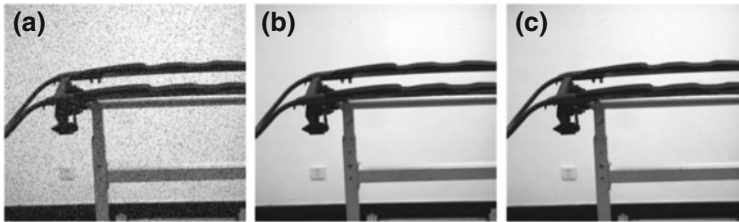


Fig. 3 Variance of 0.3 impulse noise image processing results

The above figures show that when the variance reaches 0.3, the performance of ABDND-based algorithm is better than traditional median filtering algorithm. And when the variance is less than 0.1, the performance of the two algorithms is very good. So in order to compare the advantages of the two algorithms, the MSE and PSNR of filtered images should be calculated and the results are shown in Table 1.

As can be seen from Table 1, MSE values of the ABDND-based algorithm are generally less than that of the traditional median filtering algorithm and PSNR values of the ABDND-based algorithm are generally larger than that of the traditional median filtering algorithm. With the increase of impulse noise density, the MSE values are increasing and the PSNR values are decreasing, which shows that the performance of the two algorithms is decreasing in the high-density impulse noise environment, but the increasing trend of MSE values and the decreasing trend of PSNR values of ABDND-based algorithm are relatively stable.

The result of the experimental analysis shows the performance of the ABDND-based algorithm is superior to the traditional median filtering algorithm.

3.2 The Single Gaussian Noise Filtering

The performance of traditional mean filtering algorithm and improved mean filtering algorithm are compared through the following images with different

Table 1 Single impulse noise image objective indicators

Noise variance	MSE		PSNR (dB)	
	Median filtering	ABDND algorithm	Median filtering	ABDND algorithm
0.02	15.8431	4.0666	60.2148	66.1209
0.05	16.8585	6.5489	59.9450	64.0515
0.1	20.1430	11.0911	59.1720	61.7635
0.15	34.0506	15.5597	56.8919	60.2932
0.2	68.5622	19.9735	53.8524	59.2087
0.3	315.2698	32.4516	47.2264	57.1008

Gaussian noise densities, and the filtering results are shown in Figs. 4 and 5, where (a) is the noise image, (b) is the mean filtering algorithm-filtered image, (c) is the improved mean filtering algorithm-filtered image.

The above figures show the performance of the improved mean filtering algorithm and the traditional mean filtering algorithm is not very different. So in order to compare the advantages of the algorithms, the MSE and PSNR of the filtered images should be calculated and the results are shown in Table 2.

As can be seen from Table 2, MSE values of the improved algorithm are greatly smaller than that of the mean filtering algorithm and the PSNR values of the improved algorithm are generally larger than that of the median filtering algorithm. With the increase of noise density, PSNR values of the two algorithms are

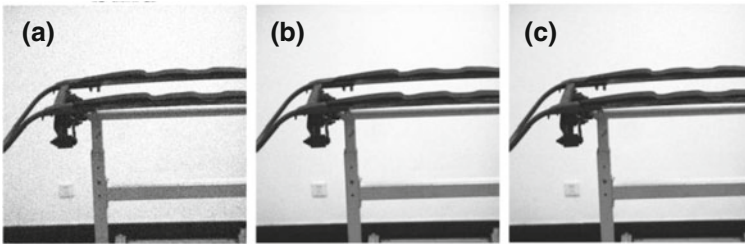


Fig. 4 Variance of 0.1 Gaussian white noise image processing results

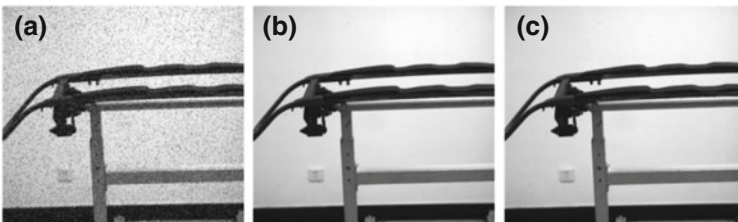


Fig. 5 Variance of 0.5 Gaussian white noise image processing results

Table 2 Objective indicators of single Gaussian noise image

Noise variance	MSE		PSNR (dB)	
	Mean filtering	Improved algorithm	Mean filtering	Improved algorithm
0.02	3.72×10^4	82.7378	26.5017	53.0326
0.05		157.8989	26.5027	50.7387
0.1		428.3777	26.5041	45.8949
0.15		834.7455	26.5051	42.9977
0.2		1.33×10^3	26.5059	40.9606
0.5		5.39×10^3	26.5077	34.8975

decreasing and MSE values are increasing, which shows that the filtering performance of the two algorithms are decreasing in the high-density Gaussian noise environment, but the increasing trend of the MSE values and the decreasing trend of PSNR values of the improved mean filtering algorithm are relatively stable.

The result shows that the performance of the improved mean filtering algorithm is superior to the traditional mean filtering algorithm.

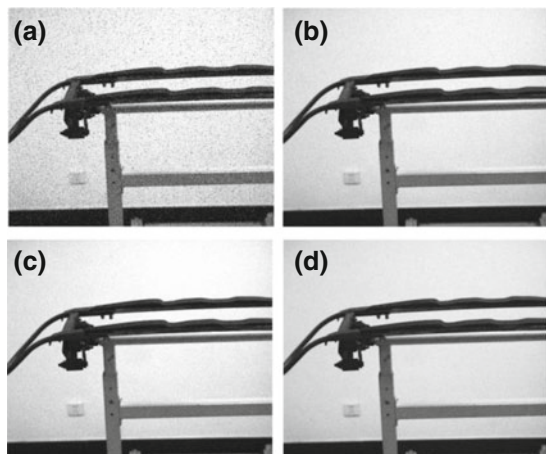
3.3 The Hybrid Noise Filtering

Based on Matlab platform, the traditional mean filtering algorithm, median filtering algorithm, and hybrid filtering algorithm are, respectively, used to suppress the Gaussian-impulse hybrid noise with different densities, and the filtering results are shown in Figs. 6 and 7, where (a) is the noise image, (b) is the median filtering algorithm filtered image, (c) is the mean filtering algorithm filtered image, and (d) is the hybrid filtering algorithm filtered image.

The above figures show the denoising performance of the three filtering algorithms is not very different for the hybrid noise. So in order to compare the advantages of the three filtering algorithms, the filtered images are calculated and the results are shown in Table 3.

As can be seen from Table 3, the MSE values of the hybrid filtering algorithm are smaller than that of the traditional median filtering algorithm and mean filtering algorithm, and the PSNR value of the hybrid filtering algorithm is larger than that of the traditional median filtering algorithm and mean filtering algorithm. With the increase of the hybrid noise density, the PSNR values of the three algorithms are decreasing and the MSE values are increasing, which show that the filtering performance of the three algorithms is decreasing in the high-density noise environment, but the trend of MSE and PSNR values of the hybrid noise filtering algorithm are relatively stable.

Fig. 6 The filtering results of hybrid noise image (variance of impulse noise is 0.02, and variance of Gaussian noise is 0.05)



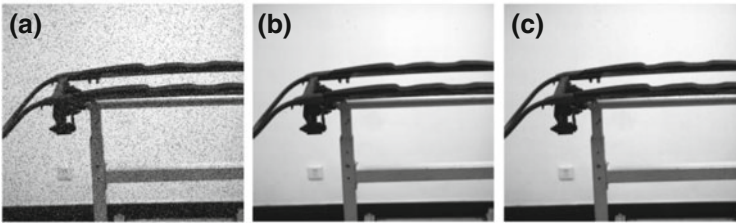


Fig. 7 The filtering results of hybrid noise image (variance of impulse noise is 0.05, and variance of Gaussian noise is 0.02)

Table 3 Hybrid noise image objective indicators

Noise variance		MSE			PSNR (dB)		
Impulse	Gaussian	Median filtering	Mean filtering	Hybrid filtering	Median filtering	Mean filtering	Hybrid filtering
0.02	0.02	140.88	3.73×10^4	95.46	46.72	22.50	48.41
0.05	0.02	147.03	3.7257×10^4	123.03	46.54	22.50	47.31
0.02	0.05	247.06	3.7246×10^4	164.57	44.28	22.50	46.04
0.05	0.05	253.28	3.725×10^4	172.43	44.17	22.50	45.84
0.10	0.10	607.01	3.7244×10^4	437.36	40.99	22.50	41.80

In summary, the above shows that the performance of the hybrid filtering algorithm proposed in this paper is superior to the traditional median filtering algorithm and mean filtering algorithm.

4 Conclusions

This paper proposed a hybrid noise filtering algorithm that the ABDND-based filtering algorithm and improved mean filtering algorithm are progressively used to filter out the hybrid noise. And the experimental results show the algorithm proposed in this paper is more effective than the traditional median filtering algorithm and mean filtering algorithm to filter out the pantograph image noise, which is very important for a pantograph state image detection method.

Acknowledgements This work is supported by national key R&D program of China (2017YFB1201201).

References

1. Xiancheng F (2013) Two new methods of gauss - salt and pepper mixed noise denoising. Jilin University (in Chinese)
2. Mengqiang X (2012) Application research on mixed noise image filtering algorithm in medical image. Lanzhou Jiaotong University (in Chinese)
3. Liu J, Tai XC, Huang H et al (2013) A weighted dictionary learning model for denoising images corrupted by mixed noise. *IEEE Trans Image Process* 22(3):1108–1120
4. Camarena JG, Gregori V, Morillas S et al (2013) A simple fuzzy method to remove mixed Gaussian-impulsive noise from color images. *IEEE Trans Fuzzy Syst* 21(5):971–978
5. Khatri S, Kasturiwale H (2016) Quality assessment of median filtering techniques for impulse noise removal from digital images. *Advanced Computing and Communication Systems (ICACCS)*. In: 2016 3rd International Conference on. IEEE, 1:1–4
6. Tanwar G, Chaudhuri SRB (2016) A novel approach to remove random-valued impulse noise from digital image. *Communication (NCC)*. In: 2016 twenty second national conference on IEEE, 2016:1–6
7. Jie G, Ning L (2012) An improved adaptive threshold canny edge detection algorithm. *Computer science and electronics engineering (ICCSEE)*. In: 2012 international conference on IEEE, 1:164–168
8. Nallaperumal K, Varghese J, Saudia S et al (2007) Salt & pepper impulse noise removal using adaptive switching median filter. In: *OCEANS 2006-Asia pacific*, IEEE pp 1–8
9. Fan AA, Wang GL (2013) A mixed denoising method based on median filter and lifting wavelet technology for sewage sensing signal treatment. In: *Applied mechanics and materials*. Trans Tech Publications, 330:967–972
10. Tripathi AK, Ghanekar U, Mukhopadhyay S (2011) Switching median filter: advanced boundary discriminative noise detection algorithm. *IET Image Proc* 5(7):598–610
11. Zhang S, Karim MA (2002) A new impulse detector for switching median filters. *IEEE Signal Process Lett* 9(11):360–363
12. Hilts M, Jirasek A (2008) Adaptive mean filtering for noise reduction in CT polymer gel dosimetry. *Med Phys* 35(1):344–355

An Effective Railway Dynamic Invader Detection and Tracking Method

Qun Wang, Xiaofeng Li and Limin Jia

Abstract In this paper, an effective railway dynamic invader detection and tracking method is proposed, which can be used to detect and track dynamic invader that appears in the vicinity of the railway track. Generally, this method can be divided into three stages, namely, intrusion region division, object detection, and invader tracking. In the first stage, we propose a novel algorithm, combining morphological processing and progressive probabilistic Hough transform, to extract the railway track, and a set of structure elements are designed to enhance the extraction effect. Based on the result of extraction, the intrusion region, the vicinity of the railway track can be divided to reduce the amount of computation later. In the second stage, we adopt visual background extractor to detect the object. And in the last stage, Kernelized correlation filters is used to track the invader, the object in the intrusion region. The results of the experiment based on actual railway scene show that the method this paper presents has excellent performance.

Keywords Railway dynamic invader · Track extraction · Object detection
Invader tracking · Kernelized correlation filters

1 Introduction

Railway invader is a crucial concern for the safety of train operation and there are many related studies at home and abroad in recent years [1]. The intrusion detection in the railway scene based on the image sequence processing and analysis is studied in [2]. By adopting edge detection to detect the candidate obstacle areas, a Support

Q. Wang (✉) · X. Li
School of Traffic and Transportation, Beijing Jiaotong University,
No. 3 Shangyuancun, Haidian District, Beijing 100044, People's Republic of China
e-mail: 15120773@bjtu.edu.cn

L. Jia
State Key Lab of Rail Traffic Control & Safety, Beijing Jiaotong University,
No. 3 Shangyuancun, Haidian District, Beijing 100044, People's Republic of China

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_42

Vector Machine (SVM) is trained to classify the areas and verify the true obstacles. An embedded algorithm of intrusion obstacle detection is proposed combining obstacle classification with moving behavior analysis in [3], using SVM and Kalman filter [4]. Hough transform to detect the area of interest, where a systematic search is done for finding and delimiting possible obstacles is used. The presented detection in [5] is based on the optical flow estimation and classification of the flow vectors by K-means clustering algorithm [6]. A novel feature based on local shape context descriptor for the object detection on railway tracks is described.

All the mentioned algorithms focus on object detection, but the real-time performance is barely satisfactory and they do not consider the invader tracking. In this paper, we present an effective railway dynamic invader detection and tracking method, the theoretical framework of which is illustrated in Fig. 1. The formulation is generic and has been tested in actual railway scene.

First, we propose a novel algorithm to extract the railway track, which works in two steps using morphological processing and PPHT respectively. Here, we design a set of structure elements to enhance the extraction effect. In order to reduce computation, we divide the intrusion region by calculating the straight line equation of track extraction result.

Second, we adopt ViBe to detect objects, which introduces the random selection mechanism into background modeling and updating for the first time. Compared with Hybrid Gauss and the optical flow, ViBe has higher modeling speed, higher detection precision, and better environmental applicability [7].

Finally, we use KCF to track the invader. The KCF has significant superiority in two aspects. One is that it makes the best of circulant matrices and discrete Fourier transform, reducing both storage and computation by several orders of magnitude, the other is that it adopts kernel trick to achieve fast detection.

2 Intrusion Region Division

The railway track extraction method proposed in this paper consists of two steps: (1) Morphological processing and (2) Progressive probabilistic Hough transform.

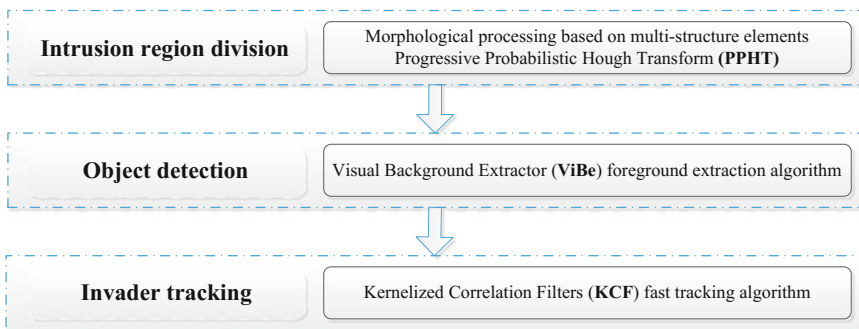


Fig. 1 The theoretical framework of the proposed method

2.1 Morphological Processing

In the process of railway track extraction, it is necessary to distinguish the tracks from sleepers, gravels, and other noise rapidly without losing the edge information of the track. However, Sobel, Prewitt, Robert, Canny, and other traditional edge detection operators are based on the discontinuity of the gray value, easily affected by sleepers and gravels, resulting in nontarget edges [8].

Using large structure elements can remove large-scale noise points, but the details of the image is lost, and using small structure elements can detect and extract the detail of the image edge, but be susceptible to large-scale noise interference [9]. So we design four structure elements of different size as follows:

$$S_1 = \begin{bmatrix} 0001000 \\ 0001000 \\ 0001000 \\ 0001000 \\ 1111111 \\ 0001000 \\ 0001000 \\ 0001000 \\ 0001000 \\ 0001000 \end{bmatrix} \quad S_2 = \begin{bmatrix} 00100 \\ 00100 \\ 00100 \\ 00100 \\ 11111 \\ 00100 \\ 00100 \\ 00100 \\ 00100 \\ 00100 \end{bmatrix} \quad S_3 = \begin{bmatrix} 010 \\ 010 \\ 010 \\ 111 \\ 010 \\ 010 \\ 010 \\ 010 \end{bmatrix} \quad S_4 = \begin{bmatrix} 010 \\ 111 \\ 010 \end{bmatrix}. \quad (1)$$

And we design three extraction operators for morphological processing.

$$\begin{aligned}
 OP_1 &= (f \circ S_2) \oplus S_4 - (f \bullet S_4) \ominus S_4 \\
 OP_2 &= f \circ S_3 \\
 OP_3 &= [(f \bullet S_3) \ominus S_3] \bullet S_1 - (f \circ S_1) \oplus S_1
 \end{aligned} \quad (2)$$

Here, \oplus is expansion operation, \ominus is corrosion operation, \circ is open operation, and \bullet is close operation.

Then based on the morphological result, we use threshold processing to extract the position of the track. The results are shown in Fig. 2.

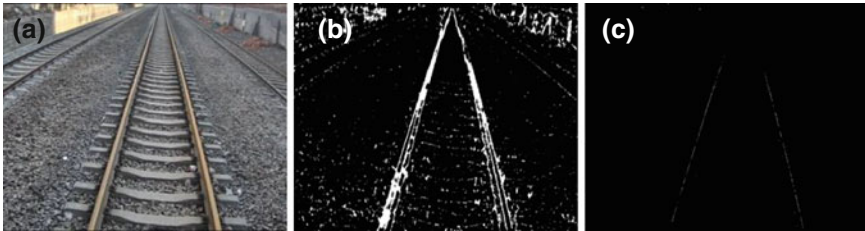


Fig. 2 a The original image, b the result of morphological processing, c the result of threshold processing

2.2 Progressive Probabilistic Hough Transform

Hough transform (HT) is usually used to detect lines and circles, and the core formula of HT is

$$\rho = x \cos(\theta) + y \sin(\theta). \tag{3}$$

Here, ρ is the length between the origin and the pedal of the detected line and θ is the angle of its perpendicular line [10].

Compared to the standard Hough transform, PPHT has two advantages. One is that it only selects part of edge points for processing, greatly reducing the amount of computation and memory required, the other is that the method can detect the endpoints of straight lines, exactly positioning the lines in the image.

Let the parameter T stand for a pair of (λ, θ) , and the highest peak means T with the most voting points. The algorithm PPHT proceeds as shown in Fig. 3.

We perform PPTH on the result of threshold processing, getting the exact location of the track. Then, we calculate the straight line equation of track extraction result and divide the intrusion region successfully. The result is shown in Fig. 4.

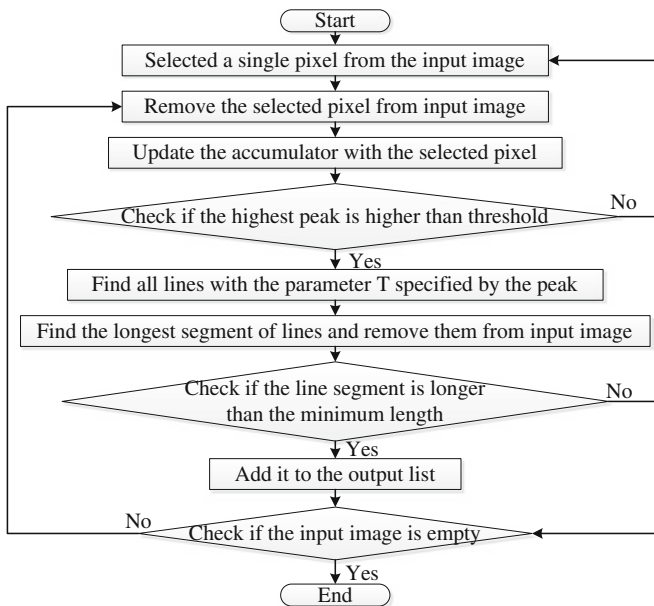


Fig. 3 The flowchart of PPTH

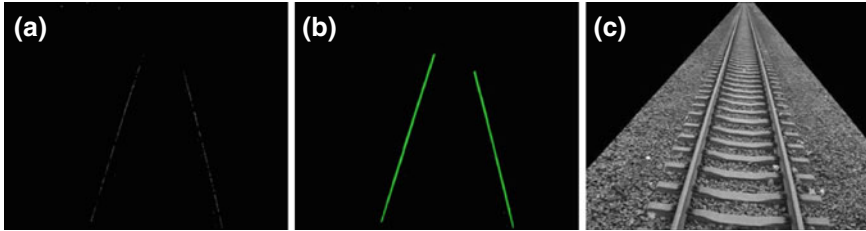


Fig. 4 a The result of threshold, b the result of PPHT, c the divided intrusion region

3 Object Detection

Visual background extractor (ViBe) is a technique for motion detection, which has good performance in both computation speed and detection rate. In this part, we use ViBe to detect an object, which consists of three steps [11].

(1) Pixel model

Each background pixel x is modeled by a collection of N background sample values taken in previous frames, v_i is background sample value with index i .

$$M(x) = \{v_1, v_2, \dots, v_n\} \tag{4}$$

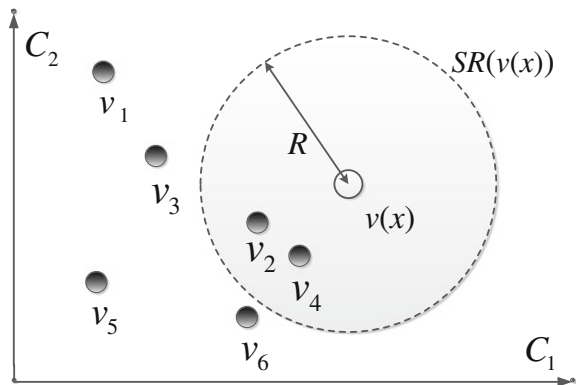
Then, pixel value $v(x)$ will be compared to the closest values within a sphere $SR(v(x))$. $v(x)$ is then classified as background if the collection of model samples $M(x)$ is larger than or equal to a given threshold. As illustrated in Fig. 5.

(2) Background model initialization

Assume that $t = 0$ indexes the first frame and that $N_G(x)$ is a spatial neighborhood of a pixel location x , therefore

$$M^0(x) = \{v^0(y \mid y \in N_G(x))\}. \tag{5}$$

Fig. 5 Comparison of a pixel value with a set of samples in a 2D space (C_1, C_2)



Here, locations y are chosen randomly.

(3) Updating the background model

The probability of a sample present in the model at the time t being preserved after the update of the pixel model is given by $(N - 1)/N$. Assuming time continuity and the absence of memory in the selection procedure, we can derive the following probability.

$$P(t, t + dt) = \left(\frac{N - 1}{N} \right)^{(t+dt)-t}. \quad (6)$$

In our experiments, the ViBe algorithm can accurately detect the person in the field of view. The results are shown in Fig. 6.

4 Invader Tacking

The core idea of KCF tracking algorithm is to use the cyclic shift coding dense sampling process to get the circulant matrix used to train the classifier [12].

Kernel trick is used to calculate the correlation coefficient between the classifier and the search area of the input image, selecting the point of the maximum correlation coefficient to update target position. As the resulting data matrix is circulant, it can be diagonalized with the discrete Fourier transform, reducing both storage and computation by several orders of magnitude [13].

4.1 Nonlinear Classifier Based on Ridge Regression Model

Ridge regression is an improved least squares model, which admits a simple closed-form solution as follows:

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2. \quad (7)$$

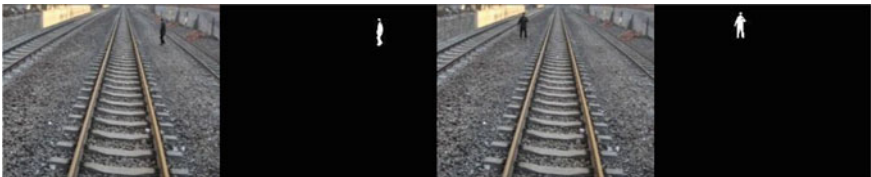


Fig. 6 The input images and the results of object detection using ViBe algorithm

Where λ is a regularization parameter that controls over-fitting, and the minimizer has a closed form, which is given by Rifkin et al. [14].

Map w to high-dimensional space and introduce kernel functions, we get the nonlinear classifier as follows:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda}. \tag{8}$$

Where \hat{k}^{xx} is the first row of the kernel matrix $K = C(k^{xx})$.

4.2 Fast Tracking Using Nonlinear Classifier

Using the nonlinear classifier to track the object, that is, the sample is mapped by the kernel function, the classifier response is as follows:

$$f(z) = w^T z = \sum_{i=1}^n \alpha_i k(z, x_i). \tag{9}$$

Here, z is the area to be detected, and x_i is training samples.

In the video, a person goes from right to left and then returns to the start position. Throughout the process, the algorithm can track the person accurately. The result of invader tracking is shown as in Fig. 7.

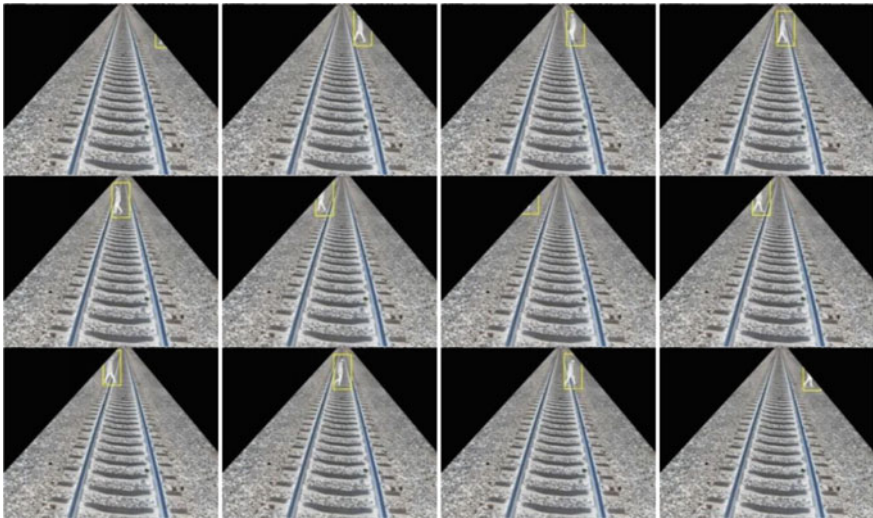


Fig. 7 The results of invader tracking using KCF

5 Conclusion

This paper presents an effective railway dynamic invader detection and tracking method, and we have done experiments on the videos captured by a fixed camera in actual railway scene. The track extraction algorithm we proposed takes full advantage of morphological processing and PPHT, avoiding the low accuracy caused by sleepers, gravels, and other noise. And the use of ViBe and KCF improves the real-time performance of the method.

The experiments show that the track extraction algorithm has excellent performance in railway track extraction. Meanwhile, the ViBe and KCF have achieved very good results in the experiments. The method can guarantee both high accuracy and real-time performance in dynamic invader detection and tracking.

Acknowledgments This work is supported by the National Key Research and Development Program of China (No. 2016YFB1200402).

References

1. Dong H, Dawei GE, Yong Q et al (2010) Research on railway invasion detection technology based on intelligent video analysis. *China railway science* (in Chinese)
2. Tong L, Zhu LQ, Yu ZJ et al (2012) Railway obstacle detection using onboard forward-viewing camera. *J Transp Syst Eng Inf Technol* 12(4):79–83 + 134 (in Chinese)
3. Shi HM, Chai H, Wang Y et al (2015) Study on railway embedded detection algorithm for railway intrusion based on object recognition and tracking. *Tiedao Xuebao/J China Railway Soc* 37(7):58–65 (in Chinese)
4. Rodriguez LAF, Uribe JA, Bonilla JFV (2012) Obstacle detection over rails using hough transform. *Image, signal processing, and artificial vision IEEE*, pp 317–322
5. Silar Z, Dobrovolny M (2013) The obstacle detection on the railway crossing based on optical flow and clustering. In: *International conference on telecommunications and signal processing*, IEEE, pp 755–759
6. Murugan D, Vivek V, Rajalakshmi K et al (2016) Obstacle detection using local shape context descriptor on railway track. *Int J Innovative Res Comput Commun Eng* 4 (2):1209–1214
7. Pan P, Zhao X, Wen C et al (2016) An improved VIBE moving target detection algorithm based on frame difference method. *Electron Technol Softw Eng* 13:115–116 (in Chinese)
8. Zhang L (2012) Automatic extraction of rail based on duda operator. *Ind Sci Tribune* 5:114–116 (in Chinese)
9. Ma H (2014) Research on the technologies of distributed intelligent monitoring for invasion limit of foreign body along the railway based on machine vision. *Lanzhou Jiaotong University* (in Chinese)
10. Mammeri A, Boukerche A, Tang Z (2015) A real-time lane marking localization, tracking and communication system. *Comput Commun* 73(2):229–233
11. Barnich O, Droogenbroeck MV (2011) ViBe: a universal background subtraction algorithm for video sequences. *IEEE Trans Image Process* 20(6):1709–1724 (A publication of the IEEE Signal Processing Society)

12. Tan S, Liu Y, Li Y (2017) Improved kernel correlation filter tracking with Gaussian scale space. *Comput Eng App* 53(1):29–33
13. Henriques JF, Rui C, Martins P et al (2014) High-speed tracking with kernelized correlation filters. *IEEE Trans Pattern Anal Mach Intell* 37(3):583–596
14. Rifkin R, Yeo G, Poggio T (2003) Regularized least-squares classification. *Acta Electronica Sinica* 190(1):93–104

Fault Diagnosis and Countermeasure for the Pantograph Automatic Lowering Under the Condition of Dual-Locomotive Running

Xiao Feng Zhong, De Chang Liu, Ping Luo and Hai Yi Zhou

Abstract A pantograph automatic lowering fault that occurred during dual-locomotive running was discussed, and signal interruption was found to be the direct reason. According to the theory of transmission lines, and combining with the freight environments, particularly, the connecting method of locomotives, this paper points out that the poor contact in the equipment of Wire Train Bus (WTB) was the root cause. Through the theoretical research and field investigation, the major factors that might cause a similar fault were researched, and the fault phenomenon was explained comprehensively. It also provides important reference value for improving WTB's reliability and reducing the possibility of pantograph automatic lowering.

Keywords Wire Train Bus (WTB) · Dual-locomotive running
Pantograph automatic lowering · Failure diagnosis

1 Introduction

Since April 2015, the type HX_D1C locomotives had been put into service in the railway line from Nanning to Kunming. According to the freight environment that includes large slope, lots of tunnel, and small radius, a pair of type HX_D1C locomotive was zero distance connected in running state. The local locomotive controls the connected one by a single transmission line. In practical use, the transmission signal often breakdown and then the pantograph automatically lowers and the

X. F. Zhong (✉)

Department of Mechanical and Automotive Technology,
Guangxi Vocational and Technical College, Nanning, China
e-mail: junxiaofeng0325@126.com

D. C. Liu · P. Luo · H. Y. Zhou

Nanning Locomotive Depot of Nanning Railway Bureau, Nanning, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_43

locomotives were forced to stop urgently. It seriously influenced the normal order of freight traffic. Therefore, it is necessary to do some field investigation and analyze the root causes of the pantograph automatic lowering.

2 Causes Analysis

2.1 Electrical Principle of Dual-Locomotive Running

In electrical principle [1], between the local locomotive and the connected one, the single transmission line was responsible for transmitting control signals. The driver in local locomotive controlled the whole operation and was in charge of monitoring and emergency handling. In network principle, the transmission line works in the upper layer of Train Communication Network (TCN), always called Wire Train Bus (WTB). The wires were redundancy designed, from the local Gateway Module's (GWM's) X1 and X3 interface, connected to another GWM. The wires can be divided into three classes: (1) control line, (2) telephone line, and (3) emergency braking and sanding line. The electrical principle is shown in Fig. 1.

2.2 Control Signal Interrupt

The dual-locomotive control electrical principle diagram shows that the control signal goes from Gateway Model (GWM) to Vehicle Control Module (VCM) and Event Recording Module (ERM) through Multifunction Vehicle Bus (MVB). In every 50 ms, the control parameters and state parameters can be recorded. Therefore, by analyzing the event record, the direct reason of pantograph automatic

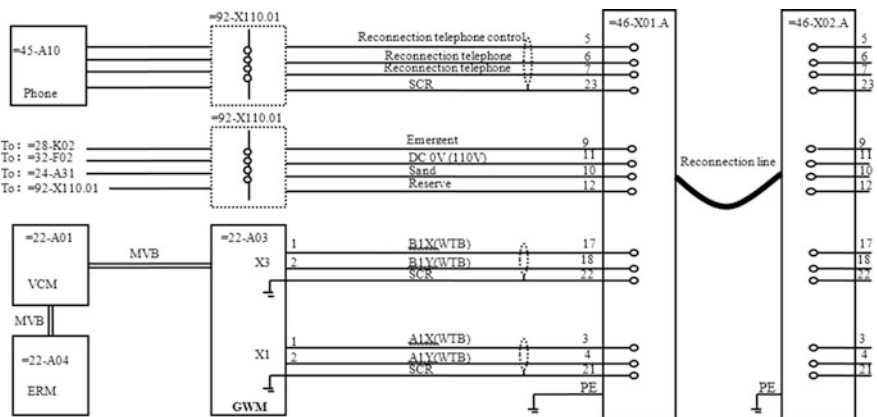


Fig. 1 Electrical principle of dual-locomotive running

lowering may be determined [2–4]. First of all, the connected locomotive does not need to turn on the electric switch in running, that is, there is no manpower occupation. The connected occupation signal was from the local locomotive, such as the GWM’s X1 or X3 interface. Before the pantograph automatic lowering, the main breaker will trigger, and then the main circuit breaker and the pantograph will be in the “blockade” state [5–8]. Based on the network data analysis, we found that before the main circuit breaker trigger, the signal of “driver’s cab occupied” disappear suddenly. So we come to the conclusion that the signal interruption was the direct reason for the pantograph automatic lowering fault. For example, at 9:26 on March 4, 2016, the connected locomotive 0539 (local 0550), the pantograph automatic lowering when the locomotive run through Xitun Station. The network data analysis, shown in Fig. 2, fully verifies the conclusions above.

2.3 Al-locomotive Connector Failure

As analyzed before, the direct cause of pantograph automatic lowering was signal interrupting. However, the fundamental reason lied in the connection components, including the failure of the connection line, connection socket, and internal components, mainly reflected in the following points:

- (1) The tail connection plugs loose, internal wiring of winding extrusion.
- (2) Ratchet buckle failure, slipping.
- (3) Rubber aging or lost.



Fig. 2 Network data analysis

- (4) Plug or socket pin indenting.
- (5) Shortening, oxidation, water, or dust.
- (6) Type or location is not consistent, easily loose.

The specific failure modes are shown in Fig. 3.

3 Countermeasures

3.1 Formulate a Scheme for Connection Parts Overhauling

In accordance with the failure modes of the connection parts, the special overhauling scheme of the relevant components such as the connection socket and the connection line should be formulated as follows:

- (1) Check the rubber no breakage and the inner wire harness is not exposed.
- (2) Rubber both ends of the clamp.
- (3) Check the pins no rust, oxidation, shortening phenomenon.
- (4) Press the contact pin firmly.
- (5) Check the rubber pad no aging phenomenon, or change it.
- (6) Check the locking mechanism is good and the button is flexible.
- (7) Spray the needle and surrounding stains with an insulating detergent.
- (8) Check the ring ratchet lock hermetical or not.
- (9) Check the installation of each pin is firm.
- (10) Check the retaining ring installation in place, no loosening.



a. Tail loosening



b. Ratchet buckle failure



c. Apron aging or loss



d. Shortening



e. Contact inlet oxidation



f. Type or position inconsistency

Fig. 3 Failure modes of connection parts

- (11) Check the wires are correct and the conduction is good.
- (12) Check the pinhole no shrinkage, oxidation, corrosion phenomenon.
- (13) Take a complete test.

3.2 Perform Special Maintenance Process

3.2.1 Overhaul and Maintenance

In the repair process, the connection line should be strictly inspected according to the overhaul scope. In the one-trip inspection, the inspection and maintenance of the connection line should be taken as the key work, especially for the locomotive with the fault of pantograph automatic lowering. The inspection team should check and report the operation, and the repair team should deal with it in a timely manner. If it is unable to be handled, the connection line should be replaced and tested, and the replaced one shall be maintained by the professional team. The requirements of the inspection and repair process are shown in Fig. 4.



a. No sharp part



b. No oxidation, shortening



c. No invalid rubber pad



d. The button must be flexible

Fig. 4 Overhaul and maintenance

3.2.2 Optimizing the Use Method

Both transmission line and connector are extremely vulnerable to breakage, and be used to transmit accurate data without any error. So it is imperative to optimize the use of methods to properly connect and carry out complete tests after maintenance; the process is shown in Fig. 5.

- (1) The transmission line was forbidden to tow or toss.
- (2) The transmission line was forbidden to instruments with hammer, stone, or other percussion connection when it was installed and dismantled.
- (3) The installation of transmission line requires fastening in place.
- (4) After dismantling the transmission line, reset the socket cover and lock it to prevent the inlet water from entering the socket.
- (5) If the line or socket was damaged, it must be recorded in the running diary and reported to professional team when the driver was off duty.
- (6) The transmission line should be placed in a uniform position.
- (7) The transmission line socket in warehouse should be bandaged and protected, kept clean and tidy.

3.3 Optimized and Improved Projects

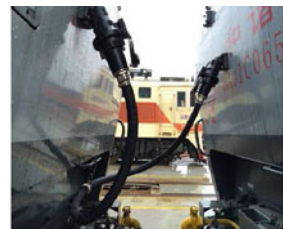
In the renovation process, three projects have been completed. First, the reform of the position of the socket: installation from the high to the low to facilitate the installation and disassembly, and the transmission line can be hung on the front wall of the middle ring, thus avoiding the bad contact caused by the vibration of the connecting wire. Second, unified type of socket: a model for unified connection of sockets and plug connections to prevent pairing errors. Third, transmission line reconstruction: it changes double-line control mode to increase signal redundancy, which is to install a divider, and divide the original two channels into four channels.



a. Check the socket pinhole's state



b. Check the cover plate's state



c. Carry out complete tests

Fig. 5 Improved use method

4 Conclusions

Through the analysis of data and theory, a conclusion was that the fault of dual-locomotive pantograph automatic lowering was directly caused by transmission signal interrupt. By field investigation, combined with the actual production, a series of measures were developed, as standardizing the maintenance method of transmission line and socket. Meanwhile, lots of improved project had been carried out, as optimizing and transforming connection line and socket. By the practice of half year, the fault of pantograph automatic lowering had been obviously reduced. According to incomplete statistics, between July 2015 and December, an average of 15 faults occurred every month. After a series of remediation, as of June 2016, it reduced to less than four faults a month and had achieved remarkable results.

Acknowledgements This work is supported by the Fundamental Research Funds for the Guangxi college young and middle-aged teachers basic ability promotion project (KY2016YB613).

References

1. Yanlin C, Hong L, Zhongjian W (2012) Electrical diagram of high power AC drive 7200KW six axle freight electric locomotive. CSR Zhuzhou Electric Locomotive Ltd., Zhuzhou (in Chinese)
2. (2010) Era of aftermarket. HXD1C locomotive network fault data download and analysis methods introduced [DK]. Zhuzhou CSR era electric Limited by Share Ltd., Zhuzhou, pp 33–35 (in Chinese)
3. Xionghua G (2013) Common faults and treatment of HXD1B/HXD1C electric locomotive. China Railway Publishing House, China, pp 109–111 (in Chinese)
4. Shun X (2016) Research and application of fault analysis algorithm based on EMU MVB network. *Comput Measur Control* (10):64–67 (in Chinese)
5. Ju G, Pengju Z (2016) Analysis of signal attenuation of communication cables connected to metro vehicle signals. *Electric Locomotive Urban Rail Veh* (02):82–83 (in Chinese)
6. Xuexiu Z, Zhongning Z, Zhaoqing L (2016) Test of special cable for three stage repair of type CRH_3 EMUs. *Fail Anal Solution Measures*. *Railway Rolling Stock* (02):76–79 (in Chinese)
7. Qiang W, Jie Y (2016) Study on the fault injection of train MVB bus. *Railway Commun Sig* (01):52–54 (in Chinese)
8. Yongjun L, Guanglin Y et al (2017) Research on MVB communication failure mechanism of Chang-sha maglev train. *Locomotive Electric Drive* (03):98–100 + 109. (in Chinese)

Simulation and Scheme Study on Braking Process of Urban Rail Transit

Xian Zhang, Yanhui Wang and Boyan Hao

Abstract In this paper, through the design of train braking process algorithm based on traction calculation, we compiled the Matlab program to realize the simulation of train braking process, studied the train braking of urban rail system and relevant problems, and set different calculation conditions for line and train, and carried out several simulation examples. The results can be used for analysis of train traction, braking speed, train mass, and other factors on the braking capacity of the train, and can also provide reference for the related engineers in the design and operation of the subway.

Keywords Urban rail transit · Brake · Simulation

1 Introduction

In the course of operation, the outside world will exert external force on urban rail transit, including train braking force, train resistance, and traction. In this paper, based on the summary of the running contents of the train, the computer program is compiled to solve the braking process of the train and simulate the braking condition of the train under different operating conditions.

With the help of braking simulation of urban rail transit vehicles, we can speed up the subway, improve the operation efficiency, and ensure the energy consumption and safe operation. Therefore, under the premise of ensuring the safety of train operation, it is important for train operation calculation and train braking simulation to realize the high-speed operation of the subway train, the energy supply, and the reduction of air pollution.

X. Zhang (✉)
School of Transportation Science and Engineering, Beihang University,
No.37 College Road, Haidian District, Beijing 100191, China
e-mail: 502181657@qq.com

Y. Wang · B. Hao
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

In recent years, with the development and deepening of computer application technology, train brake calculation methods have been greatly improved, which saves the computing time and improves the technical level and calculation efficiency.

2 Model Establishment

2.1 Theoretical Basis of Train Braking Calculation

2.1.1 Train Running Resistance W

The train running resistance is the force that hinders the train's operation and contrary to the direction of the train. It is usually expressed by W . Because of different causes, resistance can be divided into two kinds: basic resistance and additional resistance.

The basic resistance is the force accompanying the movement of the train. The force of the train is complicated, mainly including the resistance of the vehicle to contact the air during travel, the resistance between the wheel and the rail, the resistance of the friction between the vehicle components, and so on. However, many of the factors that affect the resistance of the train are difficult to measure, such as air resistance and friction between bearings. Therefore, in these complex factors, most of the empirical formula is used to calculate the basic resistance of the train. The basic resistance experience formula in China is as follows [1]:

$$w_0 = 2.7551 + 0.014v + 0.0075v^2, \quad (1)$$

where v represents the speed of the train.

Additional resistance is the force that the train receives under special conditions. For example, when the train passes through the tunnel, on the curve, on the lower ramp, and so on, additional resistance will be exerted. Additional resistance generally takes into account additional resistance of tunnel, additional resistance of curve, and additional resistance of ramp. However, there is no definite formula for additional resistance caused by climatic condition.

Therefore, when calculating the total running resistance of a train, the following parts are usually included [2]:

(1) Basic resistance:

$$W_0 = (pw'_0 + Gw''_0)g \times 10^{-3} \text{ (kN)}, \quad (2)$$

where w'_0 and w''_0 are the basic resistances of locomotive and vehicle, p stands for locomotive calculated weight, G is locomotive traction weight, and g is gravity. It comes from empirical formulas.

(2) The total resistance of a train at its operation

The following formulas are derived by mathematical and mechanical theory:

$$W = W_0 + (p + G)i_j g \times 10^{-3} \text{ (kN)}, \quad (3)$$

where i_j represents the starting point plus slope.

(3) The total unit resistance of a train at its operation

The following formulas are derived from the traction technical specifications:

$$w = w_0 + i_j \text{ (N/kN)}. \quad (4)$$

2.1.2 Train Braking Force B

The train braking force is the additional force that can be adjusted by the driver when the train needs to brake or decelerate during the running process.

The friction braking of train includes shoe brake and disc brake. The pressure K of each brake block of the train can be calculated according to the following formula [3]:

$$K = \frac{\pi d_z^2 p_z \eta_z \gamma_z n_z}{4n_k \cdot 10^6}, \quad (5)$$

where d_z is the automatic cylinder diameter of train, p_z is the air pressure, η_z is the transfer efficiency, γ_z is the braking ratio, n_z is the number of brake cylinders, and n_k is the number of brake shoe. When the train starts the power b rake, the motor of the train is equivalent to the generator, which converts kinetic energy into electric energy. At this point, it is obvious that the kinetic energy of the train will decrease, but the mass of the train is a physical attribute of the train and will not change in general [4].

2.2 Resultant Force of the Train

2.2.1 Resultant Force

When the train driving in the downhill, it will be affected by its gravity decline component. We analyze the force condition of the train at this time, if the train

resistance and braking force in the direction of the parallel direction of the slope of the absolute value are less than the gravity component; the train will accelerate, and at this time the resultant force is positive. The resultant force at this time can be expressed in the following form [5]:

$$C = -(W + B) = f(v). \quad (6)$$

We specify a force which equals to the direction of the train as a positive force for the train, such as traction. Conversely, it is called negative force, such as running resistance and braking force. The direction of the force is different, and the impact on the train will be different.

2.2.2 Unit Resultant Force

The unit resultant force can be expressed by the following formula:

$$c = \frac{C \times 1000}{(\sum m_m + \sum m_t)g}. \quad (7)$$

The denominator of the upper formula is the total gravity of the train, which should be calculated according to the actual situation [6].

2.3 Train Motion Principle

The mass of the train in operation can be expressed by pressing type:

$$m = \frac{G}{g}(1 + r) \times 1000, \quad (8)$$

where r represents the ratio of the turning part of the train to the total mass of the train, $(r + 1)$ is the coefficient of train mass, and r usually takes 0.06 in calculations.

Through the above formula, we can obtain the acceleration value of the train as follows:

$$a = \frac{C}{G} \times \frac{g}{1060}. \quad (9)$$

3 Design and Realization of Braking Simulation System for Urban Rail Transit

3.1 System Function Analysis

According to the research requirements, the traction calculation and simulation system designed in this paper have several functions, such as data entry, data visualization, and results presentation [7].

3.2 Simulation System

The system consists of three parts: input, calculation, and output (Fig. 1).

3.2.1 System Input

The input speed, braking speed, traction quality, grouping, adding ramp, and simulation button are inputs by the interface.

3.2.2 Calculation and Simulation

Through the train braking process in Simulink, a simulation model is built (Fig. 2).

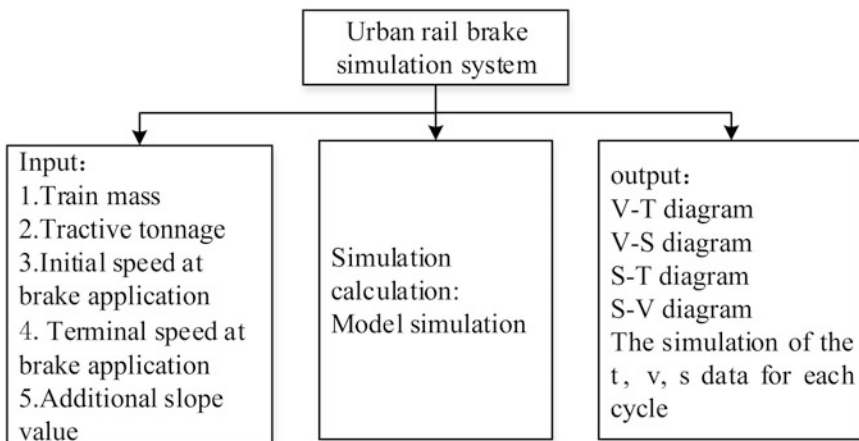


Fig. 1 System structure

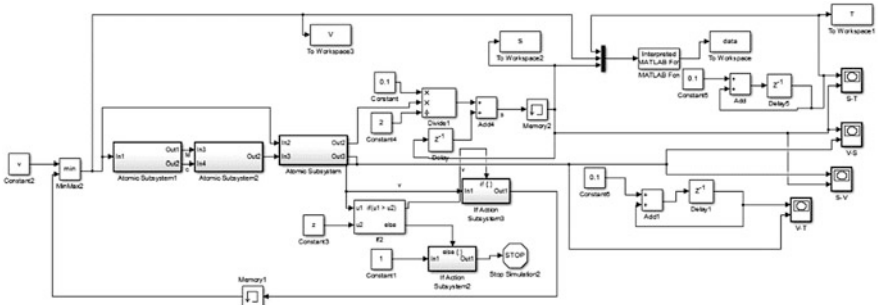


Fig. 2 Simulink simulation model

The equation of resistance is

$$w_0 = 2.7551 + 0.014v + 0.0075v^2. \tag{10}$$

The drag equation is modeled as follows (Fig. 3):

3.2.3 Output

Then, the system output $V-T$ diagrams, $S-T$ diagrams, $S-V$ diagrams, and $V-S$ diagrams containing T , V , and S , where T represents the time/s, V represents the velocity/m/s, and the S represents the distance/m.

According to the simulation model, the $V-T$ diagram is obtained, the horizontal axis is T/s , and the vertical axis is speed $V/km/h$, as shown in Fig. 4.

According to the simulation model, the $S-T$ diagram is obtained. The horizontal axis is T/s , and the longitudinal axis is the braking distance S/m , as shown in Fig. 5.

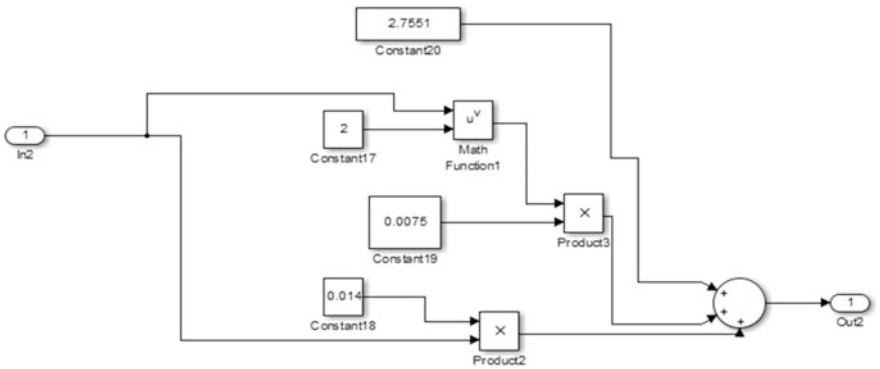


Fig. 3 Resistance equation calculation model

Fig. 4 S-T chart

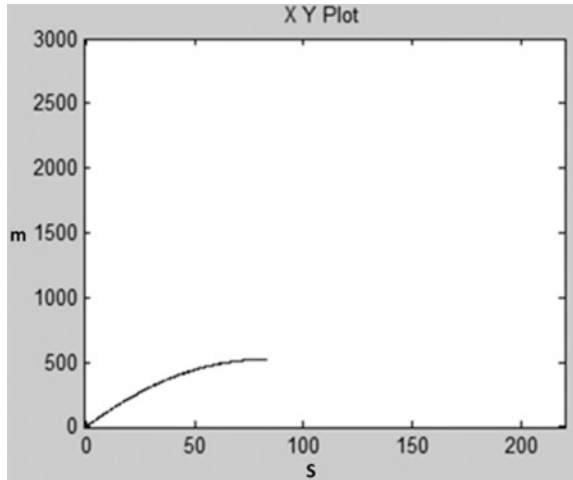
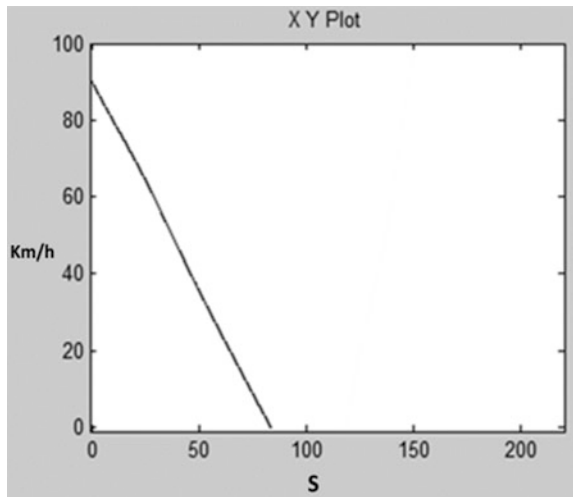


Fig. 5 V-T chart



According to the simulation model, the S-V diagram is obtained, the horizontal axis is speed V/km/h, and the vertical axis is speed S/m, as shown in Fig. 6.

According to the simulation model, the V-S diagram is obtained, the horizontal axis is the braking distance S/m, and the vertical axis is speed V/km/h, as shown in Fig. 7.

Fig. 6 S-V chart

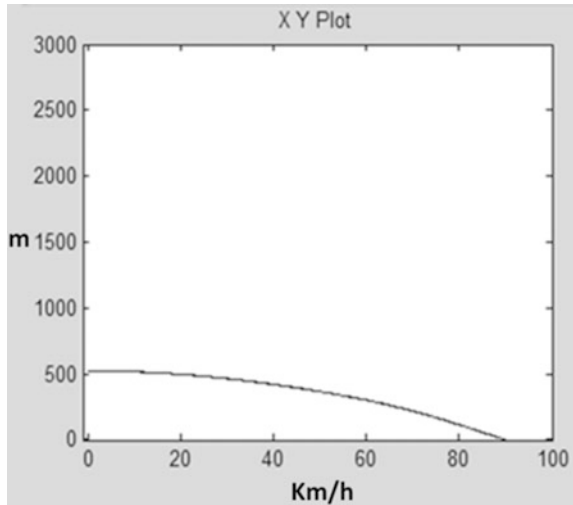
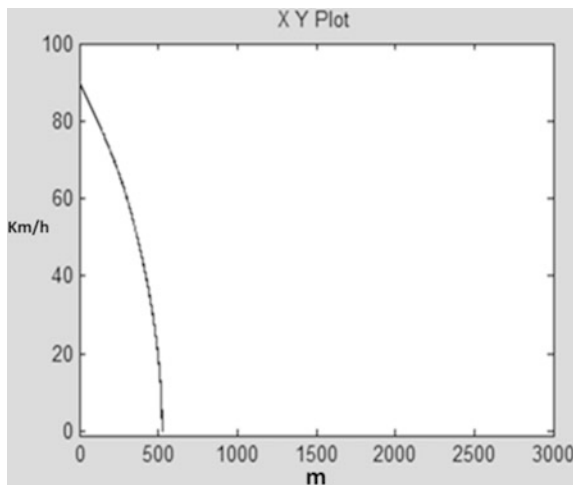


Fig. 7 V-S chart



4 Empirical Study and Result

In the case of emergency braking, the influences of train traction mass, initial braking speed, and ramp on train braking process are analyzed.

- (1) the initial braking speeds of train are 100, 90, 70, and 45 km/h.
- (2) train traction masses are 34×0 , 34×3 , 34×20 , 38×0 , 38×3 , and 38×20 (train weight * load).
- (3) the additional ramps of the train take $-5, 0, 2, 5, 10, 20$, and 35.

4.1 *Influence of Initial Braking Time on Emergency Braking Simulation*

Run the simulation system to get the results shown in Table 1.

As can be seen from the above table, whether it is in the no-load condition or full-load conditions, train braking time and braking distance increase with the initial braking speed. The greater the initial velocity, the faster the braking distance increases, while the braking time increases slowly.

4.2 *Influence of Train Traction Mass on Emergency Braking Simulation*

Run the simulation system to get the results shown in Table 2.

It can be seen that the change of train traction mass has little influence on train braking distance. On the whole, the greater the traction mass is, the greater the braking distance of train is.

Table 1 Simulation results of emergency braking for braking

Brake first speed	Traction mass	Plus ramps	Braking time	Braking distance
100	34*0 (aw0)	0	38.1	285.19
90	34*0 (aw0)	0	33.6	223.32
70	34*0 (aw0)	0	24.3	120.34
45	34*0 (aw0)	0	15.3	48.23
100	34*3	10	41.1	306.23
90	34*3	10	34.3	241.6
70	34*3	10	26.1	130.02
45	34*3	10	16.5	52.24

Table 2 Train traction mass for emergency braking simulation results

Brake first speed	Traction mass	Plus ramps	Braking time	Braking distance
90	34*0 (aw0)	10	29.7	197.9
90	34*3 (aw1)	10	31.8	209.8
90	34*20 (aw2)	10	42.6	275.7
90	38*20 (aw2)	10	44.7	290.6

5 Conclusion

On the basis of train traction calculation, this paper establishes the brake simulation system of urban rail train by GUI and Simulink in Matlab. The system reflects the braking characteristics of the urban rail train [8]. Through the system test, it is confirmed that the system basically meets the design requirements. The simulation data and the simulation curve of this system can reflect the change of the speed of the urban rail train and the change of the braking force at different times. The main achievements of this paper are as follows:

- (1) The dynamic mechanical model of the braking process is built through the related data of the urban rail train.
- (2) The simulation program of braking simulation for urban rail train is realized by Matlab.
- (3) Use the system to simulate the braking process of the urban rail train under the different conditions of trains and lines. Through the actual example, we can get the result that when the train is below 70 km/h, the train can have better emergency braking effect; the greater the traction quality, the greater the braking distance of the train; and so on.

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No. 2016YFB1200402).

References

1. Yang G (2013) Design and implementation of traction calculation and simulation system for urban rail transit. Lanzhou JiaoTong University, Lanzhou, pp 1–19 (in Chinese)
2. He G (2009) Research on train accurate parking control algorithm based on LQR. Beijing Jiaotong University, Beijing (in Chinese)
3. Hu P (2013) Urban rail transit train operation energy optimization and simulation. Southwest Jiaotong University, Chengdu, pp 2–3 (in Chinese)
4. Zhang Z (2010) Based on AMESim on the parallel hydraulic braking energy regeneration system modeling and simulation. Beijing Jiaotong University, Beijing, pp 35–55 (in Chinese)
5. Qi F, Wang L (2014) Simulation of braking process of traction system. *Electr Manufact* (03):54–59
6. Wang W, Zhan Z (2003) Development of traction calculation system based on Visual C ++ 6.0 platform. *J Transp Comput* (05):104–106
7. Guo T, Mu C, Liu R, Shen Z (2000) Design and realization of atp algorithm for a practical high-speed railway. *J Tsinghua Univ (Science and Technology)* (07):51–54 (in Chinese)
8. Ding Y, Liu H, Bai Y, Zhou F (2011) Subway train energy-saving operation of the two-stage optimization model algorithm research. *Transp Syst Eng Inf* (01)

Study of the Destruction Resistance of Urban Rail Network Based on Random and Intentional Attack

Boyan Hao, Yanhui Wang, Xian Zhang and Huiru Zhang

Abstract In complex environments, urban rail transit stations are susceptible to various factors that cause station function loss. Therefore, how to evaluate and quantify the key nodes in the network and solve the fluctuation of the whole network capacity under the intentional and random attacks has become an important issue. In this paper, the method of evaluating important nodes is defined according to the complex network index, and the efficiency of the road network under different attack rules is solved. According to the experimental results, the Beijing railway network was evaluated. It was found that the urban rail network had strong resistance under single-point attack, and it was very vulnerable under coordinated attack.

Keyword Urban rail transit · Resistance research · Random and intentional attack

1 Introduction

Urban rail transit plays an important role in urban transportation operation; however, complex environmental factors such as natural disasters, equipment failure, terrorist attacks, etc. will seriously threaten the operation of the train safety and transport quality [1]. Due to the network transport, the failure of some nodes will have a great impact on the whole network transportation capacity, which makes it necessary and urgent to evaluate the destruction resistance of urban rail network in the case of deliberate or random attack on urban rail transit network.

In complex network theory, scale-free network is a kind of complex network. A scale-free network is a network whose degree distribution follows a power law, at least asymptotically [2]. The existence of this critical node (known as the “hub”) makes the scaleless network have a strong ability to withstand unexpected failures, but it is fragile in face of synergistic attacks [3].

B. Hao (✉) · Y. Wang · X. Zhang · H. Zhang
Traffic and Transportation Planning and Management, School of Transportation,
Beijing Jiaotong University, No. 3, Shangyuan Village, Haidian District, Beijing, China
e-mail: 13251159@bjtu.edu.cn

This paper studies the resistance of urban rail network. Based on Beijing subway data, we first identify the key nodes in the subway network. Then, the attack rule is established to solve the maximum efficiency of network after random attack. Finally, loss of Beijing Metro subway network transport capacity is calculated, which will provide a theoretical basis for subway safety operations.

2 Research Object Analysis

2.1 Network Parameter Calculation

In urban rail transit, the node represents the station. There will be an edge between any two nodes if passengers can transfer from one site to another along the line [4].

Let $G = \langle V, E \rangle$ be a network with n nodes, where V and E are the node set and edge set of the network. v_i ($i = 1, 2 \dots n$) represents the node, and e_{ij} represents the edge linking node v_i and v_j . The adjacency matrix of G is an n th square matrix with the following properties:

$$A_{ij} = \begin{cases} 0, & e_{ij} \notin E(G) \\ 1, & e_{ij} \in E(G) \end{cases} \quad (1)$$

Then, we can describe the reachability of the subway line by the adjacency matrix formed by 0 and 1.

The degree k_i of node v_i describes the number of edges in the network which are connected to node i .

$$k_i = \sum_{j \in N} a_{ij}, \quad (2)$$

where a_{ij} is a nonzero element in adjacency matrix A ; N is the number of nodes in the network (Figs. 1 and 2).

The station with the degree of 1 is the terminus, the station with the degree of 2 is the ordinary station, the degrees are 3, 4, and 5 that are the multiline transfer station (Table 1).

2.2 Key Node Identify

The node betweenness is the shortest path through the node v in the network, which reflects the influence of the node on the whole network [5]. The betweenness expression of node v is

$$C_B(v) = \sum_{i \neq j \in V} \frac{d_{ij}(v)}{d_{ij}}, \quad (3)$$

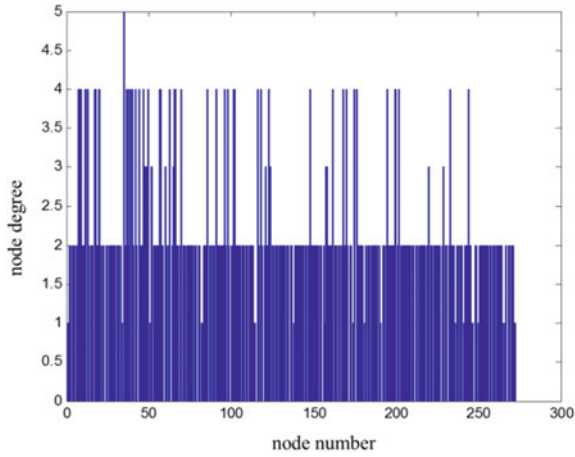


Fig. 1 Distribution of node degree in Beijing Metro network

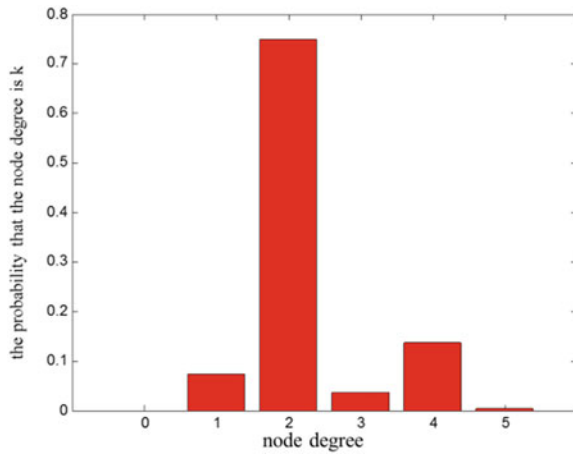


Fig. 2 Probability distribution of node degree in network

Table 1 Beijing subway network parameters

Parameter name	Beijing subway	Parameter name	Beijing subway
Line length (L/km)	574	Transfer node	46
Number of lines	19	Total edge	299
Number of nodes	288	Ordinary edge	148
Ordinary nodes	205	Important edge	96
Important node	61	Maximum number of transfers	4
Tail node	15		

where d_{ij} is the shortest path number between nodes i and j ; $d_{ij}(v)$ is the number of shortest path, which passes through the node v between i and j (Fig. 3).

The figure above shows the shortest distance between Xizhimen subway station and other 272 stations. The figure of Pingguoyuan station is obviously different. The shortest path between Xizhimen and other stations are smaller than Pingguoyuan's, which indicate that the connection of Xizhimen station is better. It has a higher degree of convenience and importance in the whole network.

Find the average number of the shortest path for each station, and draw it using the bar function. As can be seen from Fig. 4, the shortest number of stations is more convenient to transfer to other stations and occupies an important position in the entire network. By calculation, node Chegongzhuang station has the shortest average path.

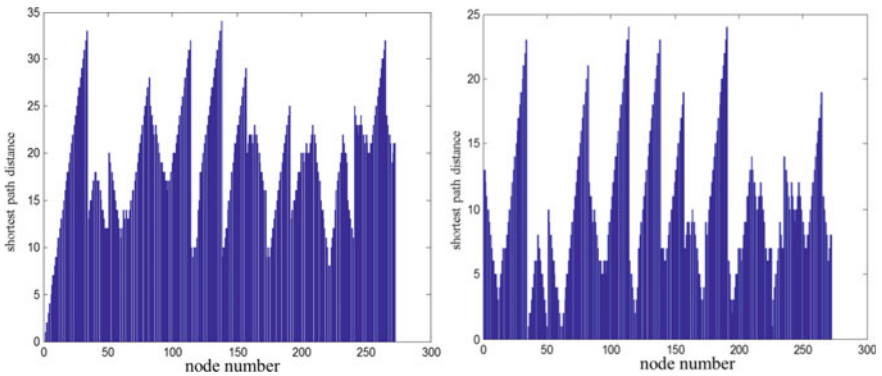
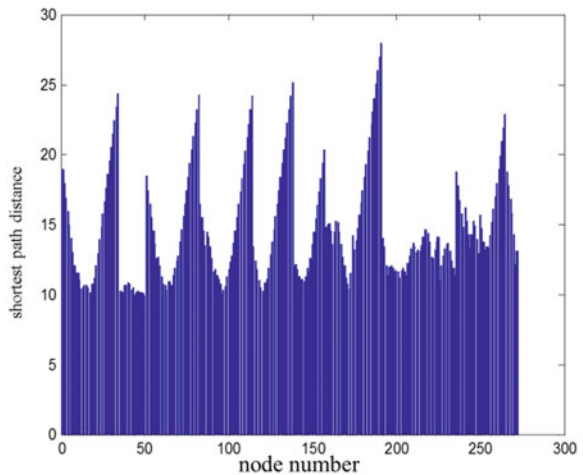


Fig. 3 The shortest distance between Pingguoyuan Station and other nodes of the road network (left), the shortest distance between Xizhimen Station, and other nodes of the road network (right)

Fig. 4 The shortest path average between nodes



3 Analysis of Urban Rail Damage Based on Attack Rules

3.1 The Efficiency of Road Network Under Deliberate Attack Rules

Define the rules of deliberate attack: maximum degree node attack, that is, from the initial state, the largest node will be removed from the network; after deleting the node, all the network features will be recalculated; so continue to delete nodes. And the nodes in the network are deleted from the network one by one.

$$E = \frac{2}{N(N-1)} \sum_{j=1}^J \sum_{i=1}^I \frac{1}{d_{ij}} \tag{4}$$

d_{ij} represents the shortest path between nodes i and j ; N represents the total number of nodes in the network. Suppose when $d_{ij} = \infty$, $\frac{1}{d_{ij}} = 0$.

Remove the largest node from the network: $\text{Max}(P1) = 5$, Xizhimen station, the matrix = $271 * 271$;

Remove the node of the four sites: 36 points removed, the matrix = $235 * 235$;

Remove the node of the three sites: 3 points removed, the matrix = $232 * 232$;

Remove the node of the two sites: 122 points removed, the matrix = $110 * 110$;

Remove the node of the one site: 22 points removed, the matrix = $88 * 88$;

Then, the network efficiency under deliberately attack can be described as shown in Fig. 5 (Table 2).

Fig. 5 Deliberately attacking network efficiency

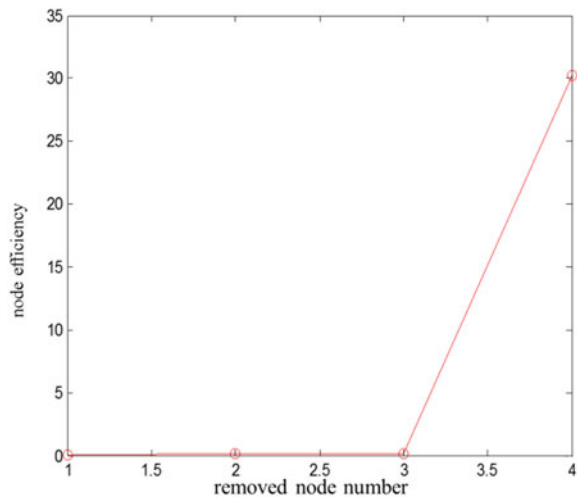


Table 2 Deliberately attacking network efficiency

Removed number	5	4	3	2	1
Network efficiency	30.2202	0.2008	0.1814	0.0018	0

3.2 Maximum Connectivity Graphs Under Random Attack Rules

Random attack rules: from the initial state, randomly delete a node in the network. Afterward, the node is deleted, and the network will be recalculated all the topological performance. Then, we repeat the random deletion of nodes in the network, and the nodes in the network are deleted from the network one by one.

Maximal connected graph: when the network is attacked, a lot of crash nodes are generated in the network, so the area of the network decreases as well, and the network may be divided into two or more unconnected subnets [6]. The region with the most nodes is called the largest connected subgraph. The most common diagram, LCC, is defined as

$$LCC = \frac{S}{S_0}, \tag{5}$$

where S_0 is the initial area of the network, that is, the number of nodes before the network is attacked, and S is the number of nodes in the largest connected subgraph in the network. Then, the network efficiency under random attack can be described by Fig. 6.

Fig. 6 Random attacking network efficiency

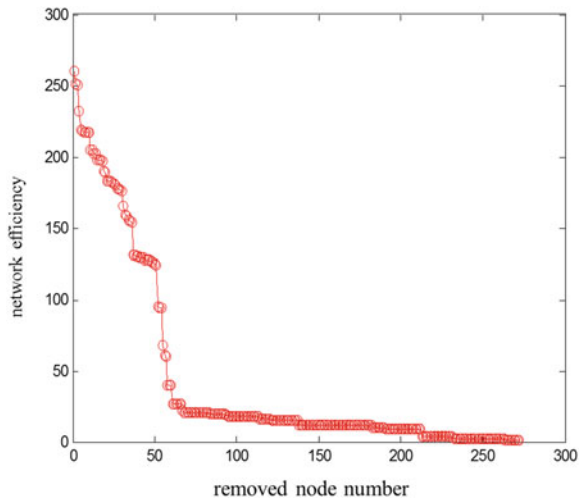
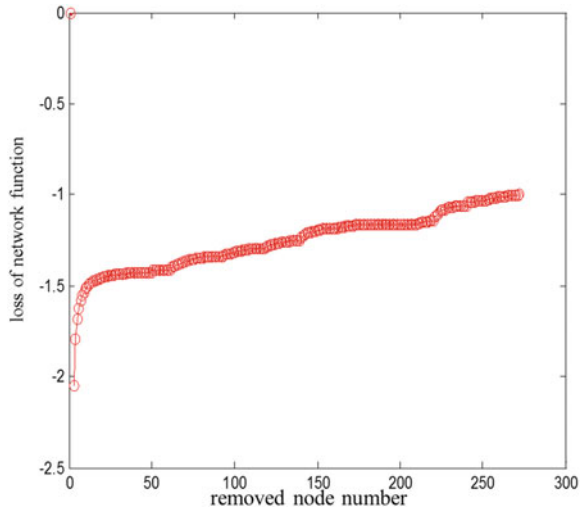


Fig. 7 Loss of network function



3.3 Loss of Urban Rail Transport Function Under Attack Rules

We abstract the complex network of the subway into an undirected graph $G = \langle V, E \rangle$, assuming that each node in the network is operating normally in the initial state, assuming that their initial functions are 1, and assuming that the transport functions of both the isolated node and the crash node are 0. When the node v_i is removed from the network, it is assumed that the deletion is the m th deletion.

$$FLN = \sum_{j=1, j \neq i}^N FL(v_j) \tag{6}$$

Then, the loss of network function under random attack can be described as shown in Fig. 7.

4 Conclusion

According to the results of the above analysis, to effectively maintain the subway transport function, the most important thing is to first ensure the existing subway stations, especially the key operation of the subway station.

The subway station is dense and the train running speed is high. If an accident occurred, it is difficult for passengers to escape, and the rescue is very difficult [7].

In the aspects of transport organization, Metro departments should implement centralized scheduling, unified command, and traffic organization according to the operation diagram. In the aspects of business, the relevant professionals such as lines, tunnels, vehicles, power supply, communication and signal, the station electrical and mechanical equipment, and fire protection system should guarantee a good state and normal operation; in safety assurance, it should rely mainly on the traffic organization and equipment to ensure the normal operation of the traffic between the necessary and correct traffic by road. In order to ensure the safe operation of the subway train punctuality, centralized scheduling, and unified command, in principle, traffic organization, equipment, vehicle maintenance, equipment operation management, and safety guarantee are regulated by a series of rules and regulations [8].

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No. 2016YFB1200402).

References

1. Yang Q, Chen T (2010) Self-organized criticality in a scale-free network with random and intentional attack. *Acta Scientiarum Naturalium Universitatis Nankaiensis* 43(2):27–33
2. Freeman WJ (2006) Origin, structure, and role of background EEG activity. Part 4: neural frame simulation. *Clin Neurophysiol Official J Int Fed Clin Neurophysiol* 116(5):572–589
3. Kurant M, Thiran P, Hagmann P (2007) Error and attack tolerance of layered complex networks. *Phys Rev E: Stat Nonlin Soft Matter Phys* 76(2):026103
4. Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286 (5439):509–512
5. Solé RV, Rosas-Casals M, Corominas-Murtra B et al (2008) Robustness of the European power grids under intentional attack. *Phys Rev E: Stat Nonlin Soft Matter Phys* 77(2):026102
6. Qiu-Ying Y, Ying-Yue Z et al (2008) Nonlinear dynamical behavior in neuron model based on small world network with attack and repair strategy. *Theor Phys* 49(4):965–970
7. Li-Ping C, Chun-Bin Y, Ke M, Xu C (2006) Detrended fluctuation analysis on correlations of complex networks under attack and repair strategy. *Theor Phys* 45(4):765–768
8. Lu ZM, Su YX, Guo SZ (2013) Deterministic scale-free small-world networks of arbitrary order. *Physica A* 392(17):3555–3562

Method of Health State Assessment for Satellite Power System Based on Fuzzy Judgment and Variable Weight Layered Evaluation

Pei Zhen Wang, Lei Wang, Ruichang Qiu, Lijun Diao and Yujia Guo

Abstract In order to improve the reliability and safety of satellite power system, the concept of health management is introduced to the research of health state assessment about satellite power supply. Considering the complex characteristics of satellite power system, a health state assessment method based on fuzzy judgment and variable weight layered evaluation is proposed and applied to the assessment of satellite power system. First, the fuzzy comprehensive evaluation method is used to quantify the influence degree of fault modes on the system and obtain the objective corresponding health degree. The method also combines analytic hierarchy process (AHP) with variable weight synthesis to solve variable weights and calculate health degree of the system step by step, which conducts real-time health state assessment for satellite power system. And then, the calculation results are divided into healthy, good, attentive, deteriorative, and morbid 5 levels to characterize the health states of the evaluation object. Finally, the applicability and validity of the method are verified by the concrete calculation example.

Keywords Satellite power system · Health state assessment · Fuzzy comprehensive judgment · Analytic hierarchy process (AHP) · Variable weight synthesis

1 Introduction

To meet the security requirements of satellite systems and to provide reliability, maintainability, and supportability, satellite power system requires advanced health management techniques. Health state assessment is an important part of the overall health management system architecture and one of the key technologies to ensure

P. Z. Wang (✉)

CRRC Zhuzhou Locomotive Corporation, Ltd., Hunan, China

e-mail: 15126035@bjtu.edu.cn

L. Wang · R. Qiu · L. Diao · Y. Guo

Beijing Engineering Research Center of Electric Rail Transportation,

School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_46

in-orbit reliable and stable operation of the spacecraft [1]. The core of the health state assessment technique is the assessment method, commonly including model method, analytic hierarchy process (AHP), fuzzy comprehensive judgment, artificial neural network (ANN), and Bayesian network (BN) [2]. Due to the characteristics of long life expectancy, high reliability requirement, and less failure samples about satellite power system and the features of complexity and confidentiality for its operating principle [3], the health state assessment method of fuzzy judgment and variable weight layered evaluation is proposed on the basis of fuzzy comprehensive judgment, AHP and variable weight synthesis. Considering the typical fault modes of satellite power supply and the typical attributes of each fault mode, the fault mode health degrees of underlying components are calculated and the superior system health degree is obtained further. The method is applicable to the actual situation of the random fault occurrence and has a good adaptability to the satellite power system.

2 Working Principle and Typical Fault Modes of Satellite Power System

The topological structure of “S3R (Sequential Switching Shunt Regulator) + BCR (Battery Charge Regulator) + BDR (Battery Discharge Regulator)” fully modulated bus circuit of satellite power system is shown in Fig. 1 [4].

Satellite power system is mainly composed of three parts, including power controller, solar array, and storage battery. The power controller modulates the bus

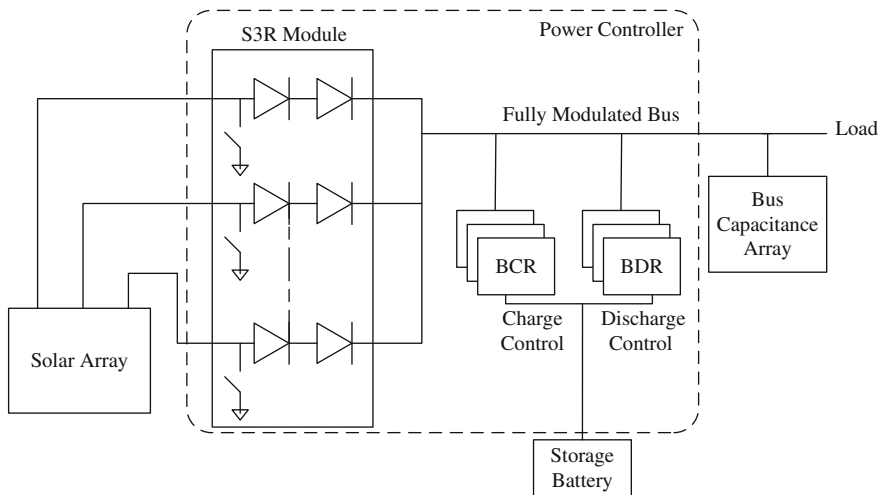


Fig. 1 Topological structure of “S3R + BCR + BDR” fully modulated bus circuit of satellite power system

Table 1 Components and typical fault modes of satellite power system

Subsystem	Subassembly	Serial number	Fault modes
Power controller	S3R	1	S3R shunt circuit shorting
		2	S3R shunt circuit opening
	MEA (main error amplifier)	3	MEA internal failure
	BEA (battery error amplifier)	4	BEA internal failure
	BCR	5	Single/multiple BCR charging fail
	BDR	6	Single/multiple BDR discharging fail
Solar array	Subarray	7	Open circuit of subarray/ substring
Storage battery	Three parallel links	8	Single battery opening
		9	Single battery shorting

voltage to remain stable. The solar array generates electricity. And the storage battery stores energy. In addition, the three parallel links of storage battery are composed of three single batteries in parallel. Above all, the typical fault modes of satellite power system are shown in Table 1 [5].

3 A Health State Assessment Method Based on Fuzzy Judgment and Variable Weight Layered Evaluation

The health state assessment method based on fuzzy judgment and variable weight layered evaluation is proposed and applied to the assessment of satellite power system. There contain nine steps in the evaluation process, which is shown below [6].

Step 1: Building the factor set

The factor set is a collection of health influence factors for the study object, which is denoted by

$$U = \{u_1, u_2, \dots, u_i, \dots, u_n\}, \tag{1}$$

where u_i denotes the i th health influence factor.

Step 2: Establishing the evaluation set

The evaluation set is a collection of health levels for the assessment object, which is shown as

$$V = \{v_1, v_2, \dots, v_j, \dots, v_m\}, \tag{2}$$

where v_j denotes the j th health level.

Step 3: Establishing the fuzzy factor evaluation matrix

In the analysis process of fuzzy comprehensive evaluation for fault mode k , the r_{ij}^k is set to be the membership of health influence factor u_i for health level v_j .

The commonly used evaluation method is to set up an evaluation group of h experts, each member evaluates only one health level v_j for each health influence factor u_i^k . If there are h_{ij}^k people evaluate that u_i^k belongs to v_j in the h members, the membership set of u_i^k is obtained as below:

$$R_i^k = \left\{ \begin{aligned} & \left\{ \frac{h_{i1}^k}{h}, \frac{h_{i2}^k}{h}, \dots, \frac{h_{im}^k}{h} \right\} = \{r_{i1}^k, r_{i2}^k, \dots, r_{im}^k\} \\ & \sum_{j=1}^m r_{ij}^k = 1 \end{aligned} \right. \tag{3}$$

And then, the fuzzy factor evaluation matrix of fault mode k is described as

$$R^k = [R_1^k \quad R_2^k \quad \dots \quad R_n^k]^T = \begin{bmatrix} r_{11}^k & r_{12}^k & \dots & r_{1m}^k \\ r_{21}^k & r_{22}^k & \dots & r_{2m}^k \\ \vdots & \vdots & \dots & \vdots \\ r_{n1}^k & r_{n2}^k & \dots & r_{nm}^k \end{bmatrix}_{n \times m} \tag{4}$$

Step 4: Using AHP to determine the weight of each influence factor

The factor weight set of fault mode k is obtained by using AHP [7, 8], which is shown as

$$W^k = \{w_1^k, w_2^k, \dots, w_n^k\} \tag{5}$$

The W^k meets the conditions of $0 < w_i^k < 1$ and $\sum_{i=1}^n w_i^k = 1$.

Step 5: Carrying out the first-level fuzzy comprehensive evaluation of fault mode k

The fuzzy comprehensive evaluation vector B^k is expressed by the following formula:

$$B^k = W^k R^k = [w_1^k \quad w_2^k \quad \dots \quad w_n^k] \cdot \begin{bmatrix} r_{11}^k & r_{12}^k & \dots & r_{1m}^k \\ r_{21}^k & r_{22}^k & \dots & r_{2m}^k \\ \vdots & \vdots & \dots & \vdots \\ r_{n1}^k & r_{n2}^k & \dots & r_{nm}^k \end{bmatrix} \tag{6}$$

Table 2 Divided standards of health degree

Health state	Health degree
Healthy	0.95–1
Good	0.85–0.95
Attentive	0.70–0.85
Deteriorative	0.60–0.70
Morbid	0–0.60

Step 6: Determining the comprehensive health level

The fuzzy comprehensive evaluation vector B^k is changed to a value C^k using gravity method, which presents the comprehensive health level of fault mode k . The C^k is calculated as below:

$$C^k = \frac{\sum_{j=1}^m (b_j \times v_1)}{\sum_{j=1}^m b_j} = B^k \cdot V^T \tag{7}$$

Step 7: Calculating the subassembly-level health degree based on the comprehensive health level of each fault mode

The comprehensive health level C^k is changed to the health degree HM^k , which is as below:

$$HM^k = \frac{\sum_{j=1}^m (b_j \times v_1)}{\sum_{j=1}^m b_j} \Big/ V_{\max} = B^k \cdot V^T / V_{\max} \tag{8}$$

$$V_{\max} = \max (v_1, v_2, \dots, v_m) \tag{9}$$

After the training of algorithm and the experience of satellite power system experts, the calculation results are divided into healthy, good, attentive, deteriorative, and morbid 5 levels. The divided standards of health degree are shown in Table 2.

Due to some subassemblies influenced by more than one fault mode, the AHP is used to determine the weight ratio of each fault mode for the corresponding subassembly.

Assuming that the subassembly l is affected by m fault modes, the weight set of the subassembly l is obtained as $W^l = \{w_1^l, w_2^l, \dots, w_m^l\}$ by AHP, which meets the conditions of $0 < m < n$ and $0 < w_m^l < 1$. As a result of the $w_1^l, w_2^l, \dots, w_m^l$ are static values, they cannot dynamically change with the influence degree of each fault mode in the actual status. In order to solve the problem, the variable weight synthesis is introduced to the static weight of AHP to transform constant weights into variable weights.

The variable weight synthesis makes use of balanced functions to achieve the goal of changing weights [9]. The variable weight formula is as below:

$$W_i^l(HM^1, \dots, HM^m, w_1^l, \dots, w_m^l) = w_i^l \cdot (HM^i)^{a-1} / \sum_{k=1}^m w_k^l \cdot (HM^k)^{a-1}, \quad (10)$$

where W_i^l is one of the variable weights of subassembly l and the HM^1, \dots, HM^m correspond to the health degrees of fault modes. The $a(0 < a \leq 1)$ is variable weight coefficient [10].

From the above, the health degree of subassembly l is shown as

$$HM^l = W^l \cdot [HM^1 \quad HM^2 \quad \dots \quad HM^m]^T \quad (11)$$

The W^l is the variable weight set of subassembly l , which is expressed as $W^l = \{W_1^l, W_2^l, \dots, W_m^l\}$.

Step 8: Calculating the subsystem-level health degree based on multilevel fuzzy comprehensive evaluation

Repeating steps 1 through 6, the second-level fuzzy comprehensive evaluation of subsystem z carries out on the basis of first-level fuzzy comprehensive evaluation of subassemblies.

The factor set of subsystem z is represented as Eq. (12):

$$U^z = \{\text{subassembly 1, subassembly 2, } \dots, \text{subassembly } l\} \quad (12)$$

The evaluation set V is invariable.

And then, the fuzzy factor evaluation matrix of subsystem z is obtained as $R^z = [B^1 \quad B^2 \quad \dots \quad B^l]^T$. The variable weight set $W^{z'}$ of subsystem z is got using AHP and variable weight synthesis. It means the weight ratio of each subassembly-level health degree for the subsystem z health degree. From the above, the health degree of subsystem z is shown as

$$HM^z = W^{z'} \cdot [HM^1 \quad HM^2 \quad \dots \quad HM^l]^T \quad (13)$$

In this formula, HM^1, HM^2, \dots, HM^l are the health degrees of subassemblies that are obtained in step 7.

Step 9: Getting the system-level health degree

According to the second-level fuzzy comprehensive evaluation of step 8, the importance of subsystems is classified on the basis of the comprehensive health level of each subsystem. The variable weight set $W^{s'}$ of system s is got using AHP and variable weight synthesis. It means the weight ratio of each subsystem-level health degree for the system s health degree. In conclusion, the health degree of system s is

$$HM^s = W^{s'} \cdot [HM^1 \quad HM^2 \quad \dots \quad HM^z]^T \quad (14)$$

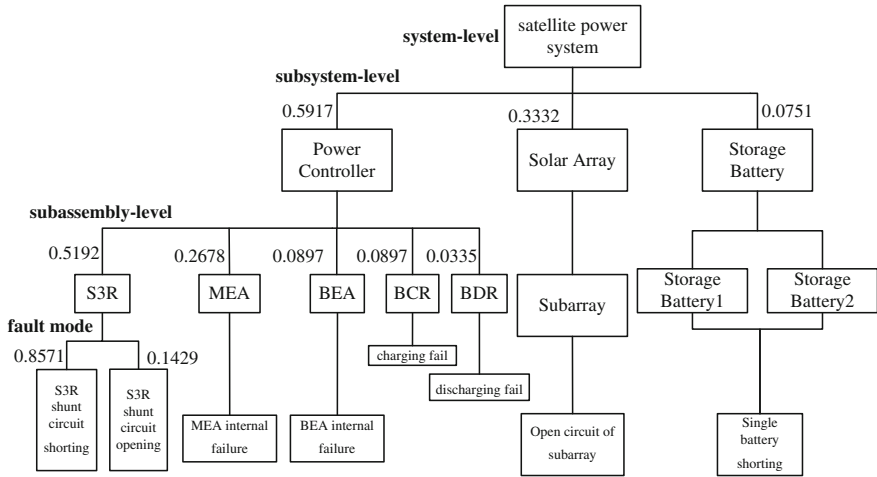


Fig. 2 Satellite power supply health assessment system hierarchical structure

In this formula, HM^1, HM^2, \dots, HM^z are the health degrees of subsystems that are obtained in step 8.

Through the above steps, combining with the working principle and structure characteristics of the satellite power system, its health assessment system hierarchical structure is established as shown in Fig. 2. In the figure, the data are constant weights of influence factors at all levels. The method realizes the calculation of fault mode, subassembly-level, subsystem-level, and system-level health degrees.

4 Concrete Assessment Example

In order to verify the validity and applicability of the abovementioned method, the concrete example that makes use of these steps is calculated. According to the example, the 3 three-parallel links of storage battery 1 occur in short-circuit fault and the other components are normally operating; the calculation results of subassembly-level, subsystem-level, and system-level health degrees are acquired to evaluate current system health states.

The health influence factors set of fault mode 9 is

$$U = \{\text{fault probability, severity, ease or complexity of detection}\}$$

And the evaluation set of fault mode 9 is $V = \{5, 4, 3, 2, 1\}$.

To classify the levels of each influence factor, the standards of factor levels classification are shown in Table 3.

Table 3 Factor levels' classification table

Influence factors	Levels				
	5	4	3	2	1
Fault probability (u_1)	Impossibly occurring	Rarely occurring	Occasionally occurring	Possibly occurring	Frequently occurring
Severity (u_2)	Almost no risk	Slight risk	Moderate risk	Serious risk	Disastrous risk
Ease or complexity of detection (u_3)	Directly detecting	Preceding detecting	Easily detecting	Comparatively difficultly detecting	Incapably detecting

Table 4 Health degrees of each level in satellite power system

System	Health degree	Subsystems	Health degrees	Subassemblies	Health degrees
Satellite power system	0.8522	Power controller	1	S3R	1
				MEA	1
				BEA	1
				BCR	1
				BDR	1
		Solar array	1	Subarray	1
Storage battery	0.7695	Three parallel links	0.7695		

In the condition of the 3 three-parallel links of storage battery 1 occur in short-circuit fault, the fault probability is calculated by the handbook GJBZ 299C-2006 [11]. Referring to Table 3, the fuzzy factor evaluation matrix R^9 is established as

$$R^9 = \begin{bmatrix} 0.86 & 0.14 & 0 & 0 & 0 \\ 0 & 0.50 & 0.50 & 0 & 0 \\ 0 & 0.70 & 0.30 & 0 & 0 \end{bmatrix}$$

Using AHP to determine the weight of each influence factor, the factor weight set of fault mode 9 is obtained as

$$W^9 = [0.2426 \quad 0.6694 \quad 0.0880]$$

Afterward, the fuzzy comprehensive evaluation vector B^9 is acquired as

$$B^9 = [0.2086 \quad 0.4303 \quad 0.3611 \quad 0 \quad 0]$$

Further, the comprehensive health level C^9 is got. Its value is 3.8475.

In the end, the calculation results of subassembly-level, subsystem-level, and system-level health degrees are acquired as in Table 4.

According to Tables 3 and 4, the health state of satellite power system belongs to “Good”, and the health state of subsystem storage battery belongs to “Attentive”. Although the system has not been seriously affected by storage battery 1 three-parallel links short-circuit fault which has not resulted in a serious decline in system power supply capacity, the output voltage of storage battery 1 decreases and the output current increases in the process of constant power output, which leads to accelerating the aging process of the whole storage battery. The capacity of storage battery 1 reduces, which results in that the depth of discharge of storage battery 2 increases. And then, the reliability of system has declined. Therefore, the ground controller of satellite power system has to enhance the monitoring of storage battery, preferentially maintain it and adjust the spacecraft operation, and load scheduling to the reduction of capacity.

5 Conclusion

In this paper, a method of health state assessment for satellite power system based on fuzzy judgment and variable weight layered evaluation is proposed to obtain the health degree to quantify the health status of the system. Combining with the working principle and structural characteristics of the satellite power system, the hierarchical structure of the health assessment system is constructed and the health degree calculation algorithm which adapts to the characteristics of satellite power supply is obtained in the thesis. What is more, the validity and applicability of the method are verified by the concrete example.

The method objectively realizes the health state assessment for satellite power supply. And the calculation results of health degrees can dynamically adapt to the random state changes for satellite power system in orbit. It is of great significance to the health management of satellite power system.

Acknowledgements This work is supported by The National Key Research and Development Program of China (2016YFB1200504-C-02), the Fundamental Research Funds for the Central Universities of China (2016JBM062), and Satellite Power Supply Failure Feature Extraction and Health Degree Calculation Modeling Project (WY - YY/D022011324JY015).

References

1. Feng W, Pan Y, Bai D, Guo J, Zhang G (2014) Application of analytical hierarchy process in spacecraft health evaluation. *Spacecraft Eng*, 4, 23(2):12–18 (in Chinese)
2. Zhou L, Zhao J, Feng G (2015) Equipment fault prediction and health management technology. National Defense Industry Press, pp 184–186 (in Chinese)
3. Liu L (2015) Faults simulation and diagnosis of satellite power system. Harbin Institute of Technology (in Chinese)
4. Wu W (2016) Research on simulation and control technology of S4R based satellite power. Nanjing University of Aeronautics and Astronautics (in Chinese)

5. Cui Z (2007) Research on satellite power system diagnosis based on qualitative model. Harbin Institute of Technology (in Chinese)
6. Xia J, Ren J, Ji H (2014) Reliability analysis of metro door system based on fuzzy FMECA. *Mach Build Autom* 43(2):184–187 (in Chinese)
7. Sun Y (2010) Research on reliability of UAV landing system based on FMECA and FTA. University of Electronic Science and Technology of China (in Chinese)
8. Zhou Z, Ma D, Yu X, Niu B (2010) Application of fuzzy FMECA in analysis of product reliability. *Electr Mach Control*, 14 (10) (in Chinese)
9. Liu W (1997) Balanced function and its application for variable weight synthesizing. *Syst Eng Theor Pract* 27(4):58–64 (in Chinese)
10. Li X, Liu Y, Li G, Liu R, Zhou J (2014) Assessment of satellite health state based on fuzzy variable weight theory. *Syst Eng Electron*, 3, 36(3):476–480 (in Chinese)
11. GJB/Z 299C-2006 (2006) Reliability prediction handbook for electronic equipment. Standards Press of China, pp 164 (in Chinese)

Intrusion Detection for High-Speed Railway Perimeter Obstacle

Qinghong Liu, Yong Qin, Zhengyu Xie, Tangwen Yang
and Gaoyun An

Abstract Perimeter intrusion detection is one of the most important conditions for high-speed railway safety. The research and application of perimeter intrusion detection are first introduced, including an infrared detector, pulse electronic fence, vibration cable/fiber optic cable, intelligent video analysis, etc. Then, we analyze the application of video surveillance in perimeter intrusion detection, point out the difficulties of video surveillance for the day- and nighttime, and summarize the popular methods of the intruding object recognition and the intruding behavior analysis. Finally, we put forward the future research direction of perimeter intrusion detection for high-speed railway.

Keywords High-speed railway · Perimeter intrusion detection · Video surveillance

1 Introduction

With the fast advancements of the high-speed railway technology, the perimeter security is becoming the most important thing to the operating safety of bullet train. The operation safety of high-speed railway has attracted much attention, and the intrusion of obstacle seriously threatens the safety of the railway. Such as people, animals, rocks, and other obstacles which invaded into the boundaries of high-speed railway will bring a huge threat to high-speed railway operation order and safety, and even cause huge loss of human life and property. At present,

Q. Liu (✉) · Y. Qin · Z. Xie

State Key Laboratory of Rail Traffic Control and Safety, School of Traffic and Transportation,
Beijing Jiaotong University, Shangyuan Village No. 3, Haidian District, Beijing 100044,
China

e-mail: 15114218@bjtu.edu.cn

T. Yang · G. An

School of Computer and Information Technology, Beijing Jiaotong University,
Shangyuan Village No. 3, Haidian District, Beijing 100044, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_47

countries in the world study various kinds of railway perimeter intrusion detection methods including fiber optic cable, ultrasonic, laser, infrared, and video surveillance, and which play very important roles in ensuring the safety of railway operation. In order to reduce and eliminate the high-speed railway accidents caused by obstacle intrusion, the perimeter intrusion detection is still the hotspot among the worldwide researchers.

2 Research Status of Intrusion Detection in High-Speed Railway

At present, the intrusion behaviors that threaten the safety of high-speed/ordinary-speed railway mainly include unauthorized entry limits and destruction of important facilities behavior as shown in Fig. 1a, b [1], and Fig. 1c, d from the Internet.

In addition to the traditional manual monitoring and physical monitoring, the common perimeter intrusion detection includes infrared detector, pulse electronic fence, vibration cable/fiber optic cable, intelligent video analysis, etc.

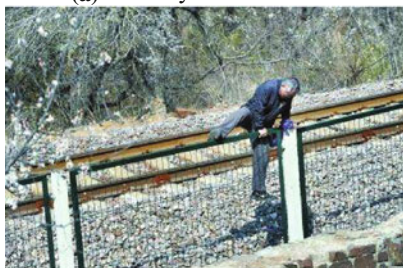
The train operation management automation system, which is composed of fiber optic cable sensors used in Japan Shinkansen railway implements the monitoring of obstacle intrusion by railway [2, 3]. Spain Juan Jesús García using infrared sensor



(a) Destroy the metal fence



(b) Destroy the cement fence



(c) Climb over the fence



(d) Perimeter Intrusion stones

Fig. 1 The different intrusion of the railway

to achieve the size of more than $0.5 \times 0.5 \times 0.5$ m obstacle detection which is installed on both sides of the railway [4]. Czech Zdenek Šilar proposed an obstacle detection using [5] optical flow method and clustering algorithm for the railroad crossing. Catalano [6] proposed an optical fiber intrusion detection system for railway security. The sensing system is composed of a rubber mat with FBG strain sensors integrated on the lower surface of the mat. It is installed and tested by a large square of approximately 20 m^2 at the Enter Autonomo Volturno (EAV) railway in Naples, Italy. Uribe [7] put forward a method of railway obstacle detection to check for obstacle intrusion with the Hough transform. In China, Tong Wang of China Academy of Railway Sciences [8] proposed the obstacle detection program combined with dual cable sensor and video surveillance to achieve passenger line obstacle intrusion detection. Wuhan University of Technology, Li [9] proposed a railway slope protective netting detection technology based on fiber grating sensor, and it determined the problem of obstacle intrusion according to the analysis of sensor response and the relationship between obstacle and other issues. Beijing Jiaotong University has studied the machine vision detection technology of obstacle intrusion. According to the geometric features of intruding objects, SVM is used to classify the intruding objects, and the Kalman filter is used to track the target [10–12]. Lei [10] extracted the size parameter information for classifying railroad turnouts, transponders, and obstacle, and used color index parameters for classification shadows. Shi [11] designed the automatic detection system, which is based on machine vision and embedded technology. It used FPGA and ARM chip to complete hardware platform of the image acquisition, and for early warning the trend of obstacle intrusion. Wang [12] proposed a hardware implementation of the railway obstacle detection algorithm based on FPGA. Central South University, Wang [13] proposed an intrusion detection to railway lines based on machine vision inspection technology and affine geometry principle, which provides technical support for railway system. China Jiliang University, Hou [14] used principal component analysis to classify video objects in induction thermal imaging.

The intelligent detection of video surveillance has the characteristics of a high degree of intelligence, flexible application, wide application range, and high recognition accuracy. Therefore, the intrusion detection based on video surveillance image intelligent identification technology has become the main development direction of high-speed railway intrusion detection.

3 Research Status of Railway Perimeter Intrusion Detection Based on Video Surveillance

The existing video monitoring system for high-speed railway lines in China can provide storage, viewing, with preliminary video analysis, and identification ability. But it has low identification accuracy and easy to be disturbed by the complex environment of the existence of large number of false alarms. It also cannot be used

in complex scenes and night. Therefore, the focus of video surveillance is to break through the video target detection methods under these extreme conditions and improve the image quality. It has great research value in the field of intelligent identification of high-speed railway perimeter video surveillance.

3.1 Daytime High-Speed Railway Perimeter Video Surveillance

In the daytime, the difficulty of high-speed video surveillance is that the poor natural environment (wind, rain, snow, fog, haze) leads to degradation of image quality, low accuracy of video object detection and low efficiency. Such as haze reduces the brightness and contrast of the image resulting in the distant objects being blurred. Rain and snow directly obscure objects in the foreground resulting in the near object to be indistinct.

(1) There are two classes of research methods on the haze/fog removal for video image: the restoration method based on the physical model and the image enhancement method [15]. He [16] proposed a fog removal algorithm based on the dark channel priority, and calculated the propagation function by minimum filtering estimate, and finally obtained the fog-free image. In the image enhancement based on nonphysical model, Wang [17] proposed an image fog removal algorithm based on bilateral filtering, and the resulting image is clear and natural. Zhang [18] used Gaussian filtering and corrosion to optimize the transmission image and enhance processing for the removal fog of image by the Gamma transformation and contrast enhancement method. The effect of local image processing is not ideal. Ting [19] proposed an image fog removal algorithm to process the thin fog image based on dark prior and the retinex theory. (2) Research on rain removal for video images, Krishnan [20] proposed an algorithm for detecting the gradient of the brightness of the raindrops pixel. Fu [21] detects and removes the high-frequency part of the rain image from the perspective of the frequency domain, with using background information as low-frequency information and raindrops as high-frequency information characteristics. Kang [22] also utilized image frequency domain characteristics to decompose the high-frequency part based on morphological component analysis, and finally realize image rain removal. Jia [23] proposed a near-field enhancement algorithm to identify rain or snow pixels in the near field view for snow, fog, heavy rain, and used the far-field enhancement algorithm to restore images. Barnum [24] detects and removes image raindrops in the image frequency domain by the fuzzy Gauss model. (3) Research on snow removal for video images, Shen [25] identifies the characteristics of snow through a series of filters, and achieve snowflake pixels extraction. Hase [26] proposed a snowfield degradation image enhancement algorithm in the time domain, and selected the midpoint of the luminance of the multi-frame image instead of the pixel in the processing frame. And Xu [27] detects and removes snow images based on the kinetic models and

photometric model in the video image. The difference is that Pei [28] proposed a method of removing rain and snow in images based on saturation and visibility characteristics, which not only removes most of the rain and snow but also preserves the details of the image. Sun [29] proposed the rain and snow removal algorithm based on the Snake model, by defining the external energy function, which can accurate positioning of the boundaries of rain and snow.

3.2 *Nighttime High-Speed Railway Perimeter Video Surveillance*

At present, the main nighttime video surveillance equipment used in railway includes infrared LED camera [30], laser night vision instrument [31], and thermal imaging camera [32], as shown in Table 1. These three kinds of equipment have their own characteristics. The monitoring distance of infrared LED camera is short, and many factors may lead to the false alarms, such as jitter, occlusion, virtual coke, light and shadow changes, mosquitoes, rain and snow, haze, etc. The monitoring distance of laser night vision is long. It was usually used to patrol but not conducive to automatic identification. Thermal imaging cameras can monitor a wide range. The target can be found in a timely manner, but cannot identify the clothing, facial features of intruding object, while the resolution is low and easy to be affected by extreme environmental temperature.

Infrared target detection is mainly based on the background modeling method and the feature extraction method. Because the detection method based on the target feature is close to the human visual recognition, Infrared target detection becomes the hotspot in the field of infrared image processing. The target information in the infrared image is mainly concentrated in the contour, shape, brightness. But it is extremely lack of color, texture, and other characteristics. Dalal [33] used a local

Table 1 Nighttime video surveillance equipment comparison

Technology type	Laser night vision	Thermal imaging camera	Infrared LED camera
Imaging principle	Active near-infrared beam reflection	Passive environmental thermal radiation	Active near-infrared light reflection
Monitoring distance (m)	1500	0–1500	150
Automatic identification distance (m)	–	0–400	0–100
Rainy days, fog, and haze use	General	Not affected	General
Key device life (h)	10,000	50,000	10,000

histogram of oriented gradient (HOG) as the image feature classification. Ulhaq [34] presented a color video fusion method based on automated color morphing to give natural day-like color appearance for nighttime video surveillance applications. Chen [35] put forward the improved shape context (ISC) method based on the traditional shape context descriptor (SCD). Most of the objects in nature will occur nonrigid deformation, because they won the ability to exercise. Felzenszwalb [36] proposed the target multiscale deformable parts model (DPM), which inherited the advantages of HOG features and SVM classifiers. Ren [37] believes that HOG features were artificially designed in previous HOG feature detection methods. For this purpose, a histogram of sparse code (HSC) based on sparse representation learning theory was proposed. In addition, some researchers have developed specifically method for pedestrian characteristics of infrared images, such as intensity self-specificity (ISS) [38], and histogram of local intensity differences (HLID) [39]. As a single feature description of the target information is limited, and combined with multiple features description of the image target detection is the development trend. HeFei University of Technology, Hu [40] proposed infrared pedestrian detection algorithm based on multiple features extraction, and combined HOG with ISS classification, by using the trained SVM intrusion detection. On the basis of that [40], Hu [41] added the integral channel feature (ICF) of multi-feature fusion infrared image pedestrian detection method.

3.3 Object Identification and Behavior Analysis

The traditional behavior identification methods mainly include hidden Markov model (HMM), dynamic Bayesian network (DBN) [42, 43], artificial neural network (ANN) [44] and SVM. In recent years, the object identification and behavior analysis based on Deep Learning have become the focus of research in this field. The behavior identification based on the deep learning methods includes unsupervised learning, convolutional neural network(CNN), recurrent neural network (RNN), and other extension models: (1) Based on the two-stream depth feature fusion method, Simonyan [45] proposed the concept of two stream, using frame image and optical flow image as CNN input and get very good output. Because the two streams cannot be of very good use for long time information, Wang [46] put forward the idea of segmentation, the video is divided into the former, middle and back three segments, each pass through two streams and then fusion results. Wu [47] at first used CNN extraction features, including RGB optical flow and spectrum image, and then through LSTM (long short-term memory) fusion. Moreover, Sharma [48] further introduced the attention mechanism into the LSTM based on behavior recognition algorithm, which has achieved the key moving target recognition in the video. (2) Based on the CNN, Tran [49] divides the video into multiple fragment containing 16 frames as input to the network. CNN input is generally pictures, for the video input. Bilen [50] tries to use pictures to represent video information, and the RGB video processes rank pooling as a CNN input.

4 Conclusion and Outlook

Video surveillance is “eyes” to discover the perimeter intrusion effectively and intuitively. The perimeter intrusion behavior can be identified automatically with pattern automatic identification technology. With China technical norms of railway video surveillance released such as “[2013] No. 71”, “[2015] No. 233” and “[2016] No. 18”, so it marks that the video surveillance is become more important in railway.

Traditional methods for high-speed railway perimeter intrusion detection have a high requirement on the environment or light conditions of the video, and it is difficult to adapt to the behavior analysis under the open non-constraint condition. And, the feature extraction method is designed by artificial prior design. The perimeter intrusion behavior analysis method based on the deep learning does not need to manually design the feature extraction method, and can be trained and studied on the video data to get the most effective representation method. However, the speed of the deep learning method is slow and the amount of sample data is huge. How to extract key information accurately and quickly from the massive video data is the trend of the future research and development about algorithms and hardware. How to overcome the environmental interference factors to reduce false alarm is an important and difficult point for the future research of intrusion detection.

Acknowledgements This work is supported by the National Key research and Development Program of China (No. 2016YFB1200203-02-02).

References

1. Institute of computing technologies of China academy of railway sciences, “High-speed railway perimeter intrusion alarm system technology implementation plan (2016)”
2. Ma H (2014) Research on the technologies of distributed intelligent monitoring for invasion limit of foreign body along the railway based on machine vision. Lanzhou Jiaotong University, Lan Zhou (in Chinese)
3. Lou R (2006) Study on video surveillance technology for railway roadblock. Southwest Jiaotong University, Chengdu (in Chinese)
4. García JJ, Ureña J, Hernández Á et al (2010) Efficient multisensory barrier for obstacle detection on railways. *IEEE Trans Intell Transp Syst* 11(3):702–713
5. Šilar Z, Dobrovolný M (2013) The obstacle detection on the railway crossing based on optical flow and clustering. In: Telecommunications and signal processing, 2013 36th international conference, Rome, IEEE, 755–759
6. Catalano A, Bruno FA, Galliano C, Pisco M, Persiano GV, Cutolo A, Cusano A (2017) An optical fiber intrusion detection system for railway security. *Sens Actuators A: Phys* 253:91–100
7. Uribe JA, Fonseca L, Vargas JF (2012) Video based system for railroad collision warning. In: Security technology (ICCST), 2012 IEEE international Carnahan conference on, IEEE, pp 280–285

8. Wang T, Shi H et al (2009) Research on intrusion monitoring system for PDL. *Railway Comput App* 18(7):8–10 (in Chinese)
9. Li L (2013) Study on fiber bragg grating sensors for slop network safety monitoring. Wuhan University of Technology, Wuan, pp 7–14 (in Chinese)
10. Lei T, Zhu L et al (2012) Railway obstacle detection using onboard forward-viewing camera. *J Transp Syst Eng Inf Technol* 12(04):79–83 + 134 (in Chinese)
11. Shi H, Chai H et al (2015) Study on railway embedded detection algorithm for railway intrusion based on object recognition and tracking. *J China Railway Society*. 37(07):58–65 (in Chinese)
12. Wang Y, Yu Z et al (2016) Hardware implementation of detection algorithm for railway clearance based on FPGA. *J China Railway Soc* 38(03):84–91 (in Chinese)
13. Wang Q, Liang X et al (2014) Visual detection method for the invasion of slowly changing foreign matters to railway lines. *China Railway Sci* 35(3):137–143 (in Chinese)
14. Hou D, Sun X et al (2014) Stimulation non-uniformity in induction thermography and its separating method. *Chin J Sci Instrum* 35(07):1466–1475 (in Chinese)
15. Yu J, Xu D et al (2011) Image defogging: a survey. *J Image Graph*, 16(9):1561–1576 (in Chinese)
16. He K, Sun J, Tang X (2011) Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell* 33(12):2341–2353
17. Wang Y, Yin C et al (2014) Image haze removal using a bilateral filter. *J Image Graph* 19(3):387–392 (in Chinese)
18. Zhang J, Wu X et al (2013) Improved algorithm for image dehazing using dark channel prior. *Video Eng* 37(23):192–195 (in Chinese)
19. Ting S (2015) The image defogging algorithm based on physics and nonphysics model. Jishou University (in Chinese)
20. Krishnan S, Venkataraman D (2012) Restoration of video by removing rain. *Int J Comput Sci Eng App* 2(2):19–28
21. Fu YH, Kang LW, Lin CW et al (2011) Single-frame-based rain removal via image decomposition. In: *Proceeding of 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE Press, Prague, Czech, pp 1453–1456
22. Kang LW, Lin CW, Fu YH (2012) Automatic single-image-based rain streaks removal via image decomposition. *IEEE Trans Image Process* 21(4):1742–1755
23. Jia Z, Wang H, Caballero RE et al (2012) A two-step approach to see-through bad weather for surveillance video quality enhancement. *Mach Vis Appl* 23(6):1059–1082
24. Barnum PC, Narasimhan S, Kanade T (2010) Analysis of rain and snow in frequency space. *Int J Comput Vision* 86(2–3):256–274
25. Shen Y, Ma LZ, Bao Y et al (2010) Detecting and extracting natural snow from videos. *Inf Process Lett* 110(24):1124–1130
26. Hase H, Miyake K, Yoneda M (2008) Real-time snowfall noise elimination. In: *Proceedings of the 1999 IEEE international conference on image processing*, Springer, Berlin, Heidelberg, pp 451–458
27. Xu L (2007) Detection and removal of snow from videos. Xinjiang University (in Chinese)
28. Pei SC, Tsai YT, Lee CY (2014) Removing rain and snow in a single image using saturation and visibility features. In: *IEEE international conference on multimedia and expo workshops*. IEEE, pp 1–6
29. Sun Y, Duan X et al (2011) Research on removal algorithm of rain and snow from images based on improved snake model. *Appl Res Comput* 28(5):1991–1993 (in Chinese)
30. Nan X (2016) Railway integrated video monitoring system technology and its linkage scheme. Jiangxi University of Finance and Economics (in Chinese)
31. Li K, Feng J (2015) Discussion on the application of new technology of video monitoring in railway. *Chin Railways* 04:99–102 (in Chinese)
32. Han L, Yu Y (2014) Application of thermal imaging technology in railway monitoring. *Chin Railways* 05:112–114 (in Chinese)

33. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. *Computer vision and pattern recognition, 2005. CVPR 2005*. In: *IEEE computer society conference on IEEE*, 2005, 1:886–893. 2005, 1:886–893
34. Ulhaq A, Yin X, He J, Zhang Y (2016) FACE: fully automated context enhancement for night-time video sequences. *J Vis Commun Image Represent, Part B*, October 2016, 40:682–693
35. Chen L, Li W, Xu Z et al (2012) Pedestrian detection based on ISC in infrared images. *Networking and distributed computing (ICNDC)*. In: *2012 third international conference on, IEEE*, pp 166–169
36. Felzenszwalb PF, Girshick RB, Mc Allester D et al (2010) Object detection with discriminatively trained part-based models. *IEEE Trans Pattern Anal Mach Intell* 32(9): 1627–1645
37. Ren X, Ramanan D (2013) Histograms of sparse codes for object detection. *Computer vision and pattern recognition (CVPR)*. In: *2013 IEEE conference on IEEE*, pp 3246–3253
38. Miron A, Basbes B, Rogozan A et al (2012) Intensity self-similarity features for pedestrian detection in far-infrared images. In: *Intelligent vehicles symposium (IV)*, 2012 IEEE, pp 1120–1125
39. Kim DS, Kim M, Kim BS et al (2013) Histograms of local intensity differences for pedestrian classification in far-infrared images. *Electron Lett* 49(4):258–260
40. Hu Q, Wang L (2016) Pedestrian detection in infrared images based on multi-features. *Electronic Des Eng* 24(04):182–185 + 189 (in Chinese)
41. Hu Q, Lv P (2016) Pedestrian detection in infrared images based on multi-features fusion. *J Comput App* 36(S1):157–160 + 195 (in Chinese)
42. Luo Y, Wu TW, Hwang JN (2003) Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks. *Comput Vis Image Underst* 92(23):196–216
43. Park S, Aggarwal JK (2004) A hierarchical Bayesian network for event recognition of human actions and interactions. *ACM J Multimedia Syst Spec Issue Video Surveill* 10(2):164–179
44. Buccolieri F, Distanto C, Leone A (2005) Human posture recognition using active contours and radial basis function neural network. In: *Proc of Conference on advanced video and signal based surveillance*
45. Simonyan K, Zisserman A (2014) Two-stream convolutional networks for action recognition in videos. *Adv Neural Inf Process Syst* 1(4):568–576
46. Wang L, Xiong Y, Wang Z et al (2016) Temporal segment networks: towards good practices for deep action recognition. *Acm Trans Info Syst* 22(1):20–36
47. Wu Z, Jiang YG, Wang X et al (2015) Fusing multi-stream deep networks for video classification. *Comput Sci*
48. Sharma S, Kiros R, Salakhutdinov R (2015) Action recognition using visual attention. *Comput Sci*
49. Tran D, Bourdev L, Fergus R et al (2014) Learning spatiotemporal features with 3D convolutional networks, pp 4489–4497
50. Bilen H, Fernando B, Gavves E et al (2016) Dynamic image networks for action recognition. In: *IEEE conference on computer vision and pattern recognition, IEEE Computer Society*, pp 3034–3042

Research into Thermal Response Time Fault Diagnosis of Temperature Sensor in High-Speed EMU

Xuguang Lin, Wenjun Wang, Bin Li, Jianwei Miao and Yantong Liu

Abstract High-speed EMU temperature sensors are used to detect important parameters such as temperature on the train. The thermal response time of the temperature sensor is the main factor affecting the temperature measurement. The failure of the thermal response time will lead to the failure or false positives of the vehicle temperature sensor, which will cause the speed limit and stop, which will seriously affect the railway operation order. This paper focuses on the thermal response time of the temperature sensor of the high-speed EMU and analyzes the dissection and simulation of the temperature sensor. The results show that the reason for the failure of the thermal response time is that the sensor probe into the air leads to the thermal resistance becomes larger, reduces the heat transfer speed, and finally makes the thermal response time longer. Finally, according to the diagnostic analysis, results put forward the corresponding measures.

Keywords High-speed EMU · Temperature sensor · Thermal response time
Fault diagnosis

1 Introduction

High-speed EMU temperature sensor is used to detect and monitor the train temperature and other important parameters to ensure that the train is running in the normal state or the best condition. Therefore, to ensure the reliability of vehicle temperature sensor, the stability of its track train operation and safety has a very important significance.

X. Lin (✉) · W. Wang · B. Li · J. Miao · Y. Liu
CRRC Changchun Railway Vehicles Co., Ltd., No. 2001 Changke Road,
Changchun, Jilin Province, China
e-mail: linxuguang@ccar.com.cn

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_48

Common vehicle temperature sensor failure, including platinum thermal resistance performance damage caused by temperature jump, electrical connector pin, and wire connection failure, thermal response time failure, sensor insulation, or breakdown occurs. Commonly used sensor fault detection methods include temperature tolerance test, thermal response time test, high-temperature vibration test, and protection level test. The thermal response time of the temperature sensor is the main factor affecting the temperature measurement speed point, and it is also an important parameter to measure the performance of vehicle temperature sensor [1–3]. Thermal response time failure will lead to vehicle temperature sensor that cannot keep up with high-speed train temperature change speed lag, resulting in failure or false positives, resulting in speed and parking, seriously affecting China’s railway operation order. Therefore, it is very meaningful to study the thermal response time of the temperature sensor. This paper focuses on the failure of the thermal sensor time test, the temperature sensor for anatomy, and this numerical simulation calculation, analysis of the causes of failure, and follow-up measures [4].

2 Temperature Response Failure of Temperature Sensor

2.1 Thermal Response Time Fault Phenomenon

From a domestic joint venture plant products sampling a number of temperature sensors for testing, in which a temperature sensor through the data acquisition system to test the thermal response time 3 times to take the average. The thermostat tank is filled with water, the temperature sensor temperature part into the thermostat, while opening the data acquisition system began to record the thermal response time curve. Set the initial temperature of 20 °C, the temperature step value of not more than 10 °C. Record the sensor resistance change to 90% of the thermal response time. Each test result for the average deviation should be within ±10% [5], requiring that thermal response time is not more than 12 s. It is found that the thermal response time recorded by the data acquisition system is smoothed by digital filter and can be used to calculate the thermal response time directly. It is found that the test result is unqualified and the average thermal response time is more than 12 s. The data are shown in Table 1.

Table 1 Test results of the fault sensor thermal response time

Number of times	First time		Second time		Third time		Average value	
	1	2	1	2	1	2	1	2
Thermal response time(s)	14.45	14.36	14.20	14.05	14.22	14.21	14.29	12.26
Result	Unqualified							

2.2 Analysis of Thermal Response Time

In high-speed EMU temperature sensor using platinum thermal resistance for temperature sensing, platinum resistance of the resistance value increases with the temperature rise. Platinum resistance connection wire is welded together, and high-temperature heat shrink tubing on the resistance pin shrinkage protection is used. It will be sealed after the resistance of the sensor into the forefront of the probe, in order to achieve the role of sealing protection and the application of potting plastic potting resistance potting treatment [6]. And then, it uses the seizure machine for sleeve and clamp hose crimping. And then according to the wiring diagram, into the insulator. The sensor probe material is made of stainless steel, and the internal potting is used in silicone. After the thermal response time test results failed, the fault pieces were dissected. Analysis of the fault sensor will be worn off the end of the protective tube, visible protection tube at the bottom of filling full. Remove the protective tube from the outside of the platinum thermal resistor, and the gap between the thermally and platinum resistors is filled with the sealant. Remove the protective tube from the outside of the platinum thermal resistor, and the gap between the thermally and platinum resistors is filled with the sealant.

Continue to grind out the rest of the protective pipe, and the clearance between the heat shrinkable tube and the inner wall of the protective tube is also filled with sealant. However, there are obvious black oxides on the inner wall of the protective tube and the sealing surface. Through the micrometer, it can be measured that the distance between the cavity and the inner wall of the sensor probe is about 0.3 mm, suspected to be in the air. The oxide is derived from the inner wall of the protective tube, and it is initially determined that the oxide is a ferrous oxide, which is caused by the oxidation of the stainless steel and air in the inner wall of the protective tube. The black oxide on the surface of the potting surface is presented in Fig. 1.

In order to verify the above judgment, reproduction of test pieces for testing and verification. The replacement of the new protective tube (no black oxide on the wall) temperature sensor to re-thermal response time test, the test results are qualified, the specific data in Table 2.

After the temperature sensor was dissected and tested again, the temperature sensor test results of the replacement protection tube were found to meet the thermal

Fig. 1 Black oxide on the surface of the potting surface



Table 2 Test results of the normal sensor thermal response time

Number of times	First time		Second time		Third time		Average value	
	1	2	1	2	1	2	1	2
Channel								
Thermal response time(s)	11.40	11.20	11.46	11.10	10.93	11.21	11.26	11.17
Result	Qualified							

response time requirements. The analysis suggests that the black oxide in the inner wall of the protective tube is ferroferric oxide, which is the protection of the inner wall of the stainless steel and air generated by the oxidation reaction, and the heat response time failure is due to protection of the tube wall in the air thermal conductivity that is very low, reducing the heat transfer speed, and ultimately lead to longer thermal response time.

3 Fault Analysis Based on Time Distribution Characteristic of Temperature Sensor

3.1 Normal Temperature Sensor Time Distribution Characteristics

In order to verify that the thermal response time is not qualified, the reason is that the air in the inner wall of the protective tube is caused. The simulation model is calculated. Because the liquid in the thermostat is flowing, the general liquid is water, so the heat transfer coefficient of the external surface and the liquid is the convective heat transfer coefficient. To analyze the cause of temperature sensor thermal response time failure, convective heat transfer coefficient is required. The convection heat transfer coefficient [7] is determined by the simulation analysis and the normal temperature sensor.

The simulation takes the normal temperature sensor as the research object, and the main structural parameters of the sensor are shown in Table 3.

From Table 2, it can be obtained that without failure of the vehicle temperature sensor thermal response time is about 11 s or so. Through the thermal response time, test can be seen that the initial temperature is 20 °C, the temperature step is set to 9 °C, according to the thermal response time calculated and temperature rise is

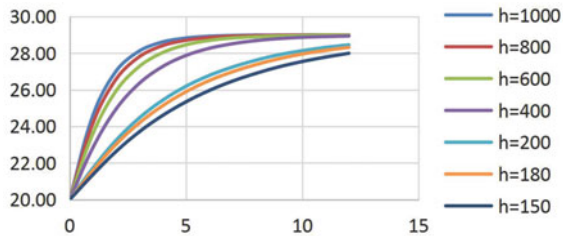
Table 3 Specific technical parameters of vehicle temperature sensing

Units	Parameter	Stainless steel	Silica gel
kg/m ³	Density	7930	1900
J/(kgK)	Specific heat capacity	500	270
W/(mK)	Thermal conductivity	16	1.5

Table 4 Relationship between temperature and time of normal temperature sensor

Time (s)	h = 1000		h = 800		h = 600		h = 400		h = 200		h = 180		h = 150	
	293.15	20.00	293.15	20.00	293.15	20.00	293.15	20.00	293.15	20.00	293.15	20.00	293.15	20.00
0	297.76	24.61	297.33	24.18	296.76	23.61	296.00	22.85	294.88	21.73	294.74	21.59	294.52	21.37
1	300.18	27.03	299.74	26.59	299.10	25.95	298.11	24.96	296.40	23.25	296.18	23.03	295.79	22.64
2	301.27	28.12	300.96	27.81	300.43	27.28	299.51	26.36	297.62	24.47	297.34	24.19	296.86	23.71
3	301.76	28.61	301.56	28.41	301.18	28.03	300.42	27.27	298.58	25.43	298.28	25.13	297.75	24.60
4	301.97	28.82	301.85	28.70	301.60	28.45	301.02	27.87	299.34	26.19	299.04	25.89	298.49	25.34
5	302.07	28.92	302.00	28.85	301.84	28.69	301.41	28.26	299.93	26.78	299.65	26.50	299.11	25.96
6	302.11	28.96	302.07	28.92	301.97	28.82	301.66	28.51	300.39	27.24	300.13	26.98	299.62	26.47
7	302.13	28.98	302.10	28.95	302.04	28.89	301.83	28.68	300.76	27.61	300.52	27.37	300.04	26.89
8	302.13	28.98	302.12	28.97	302.08	28.93	301.94	28.79	301.06	27.91	300.84	27.69	300.40	27.25
9	302.14	28.99	302.13	28.98	302.10	28.95	302.01	28.86	301.28	28.13	301.10	27.95	300.69	27.54
10	302.14	28.99	302.13	28.98	302.12	28.97	302.05	28.90	301.46	28.31	301.30	28.15	300.93	27.78
11	302.14	28.99	302.14	28.99	302.13	28.98	302.08	28.93	301.60	28.45	301.46	28.31	301.14	27.99
12	302.14	28.99	302.14	28.99	302.13	28.98	302.08	28.93	301.60	28.45	301.46	28.31	301.14	27.99

Fig. 2 Normal temperature sensor temperature changes over time



reached at 28.1 °C. Vehicle temperature sensor from 20 °C up to 28.1 °C, through the temperature sensor structure and the main parameters and simulation calculations [8], the convective heat transfer coefficient h is shown in Table 4.

From the data obtained in the simulation, when the convective heat transfer coefficient is $h = 180$, a temperature sensor for vehicles is at about 11S, and the temperature rise from 20 to 28.15 °C. The test results are in accordance with the test data of normal temperature sensor, so the heat transfer coefficient is $h = 180$ convection between the temperature sensor and the liquid. The specific temperature sensor with time changes in the simulation figure, shown in Fig. 2.

3.2 Time Distribution Characteristic of Fault Temperature Sensor

Through the previous dissection of the fault temperature sensor, the initial judgment is that the black oxide is the tri-iron oxide, which is generated by the oxidation of the inner wall of the protective tube and the air. Therefore, the parameters of the probe location of the fault temperature sensor are not only silica gel and stainless steel [9] but also air. The specific parameters are shown in Table 5.

As can be seen from Table 5, the thermal conductivity of iron oxide is very large, because the greater the thermal conductivity, the faster the heat transfer rate, the shorter the thermal response time. So it is possible to eliminate the cause of the failure of the thermal response time due to the iron oxide. However, since the black oxide is iron oxide, it is possible to determine that the air is mixed in the potting liquid. Since the thermal conductivity of the air is low from Table 4, the presence of air causes the thermal resistance to become large, so that the thermal response time lengthens, in order to verify that the thermal response time becomes longer due to

Table 5 Faulty vehicle temperature sensing specific technical parameters

Units	Parameter	Stainless steel	Silica gel	Fe_3O_4	Air
kg/m ³	Density	7930	1900	5180	1.225
J/(kgK)	Specific heat capacity	500	270	618	1006.43
W/(mK)	Thermal conductivity	16	1.5	5	0.0242

air mixing. According to the specific technical parameters of the faulty vehicle temperature and the convective heat transfer coefficient h obtained before, when the temperature rises from 20 to 28.1 °C, the simulation is carried out to obtain the fault temperature sensor when the temperature rises to about 28.1 °C thermal response time, as shown in Table 6.

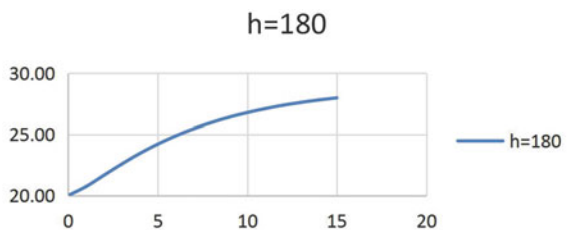
As can be seen from Table 6, the thermal response time of the fault temperature sensor is calculated from about 20 °C to about 28.1 °C by simulation. The thermal response time is about 14–15 s. Specific fault temperature sensor temperature changes with time curve shown in Fig. 3.

By the simulation calculation results, the temperature sensor whether thermal response time for about 14–15 s, compared with before test data consistent, so to determine the cause of temperature sensors thermal response time breakdown because of heat shrinkable tube between the sensor and potting glue into the air, thermal resistance due to the presence of air, reduced the heat transfer rate, eventually led to the thermal response time.

Table 6 Normal temperature sensor temperature versus time

Time (s)	h = 180	
0	293.15	20.00
1	293.89	20.74
2	294.82	21.67
3	295.74	2.59
4	296.60	23.45
5	297.36	24.21
6	298.03	24.88
7	298.61	25.46
8	299.11	25.96
9	299.55	26.40
10	299.92	26.77
11	300.24	27.09
12	300.52	27.37
13	300.75	27.60
14	300.95	27.80
15	301.12	27.97

Fig. 3 Fault temperature sensor temperature changes with time



4 Fault Temperature Sensor Handling Measures

It is recommended that the manufacturer inspects the protective tube parts of the follow-up vehicle temperature sensor to check whether the inner wall of the protective tube has a foreign matter inspection project, strictly controls the quality of the parts, and performs 100% thermal response time test on all products and products. In the process of sensor potting, to avoid improper operation of the staff lead to air into the sensor probe wall potting glue, each process is completed by the staff after the confirmation and then enters the next process.

5 Conclusion

In this paper, it is found that the thermal response time of a certain temperature sensor is unqualified. The reason for the failure of the thermal response time is determined by the dissection and simulation of the sensor. The reason that the thermal response time is not qualified is that the presence of air causes the heat resistance to increase and the heat transfer rate is reduced, resulting in a longer thermal response time. The fault of thermal response time is proposed to the manufacturer, and the reliability, availability, and safety of the vehicle temperature sensor are improved to ensure the safe operation of the high-speed train group.

References

1. Tangawa M, Kato K, Kanada K et al (2003) Response compensation of temperature sensors. *Trans Jpn Soc Mech Eng, Part B* 69(678):414–421
2. IEC 60751 (2007) Industrial platinum resistance thermometers and platinum temperature sensors
3. Lu WK, Lu XZ, Zhu CC, Liu QH, Zhang HX (2012) Solving three key problems of the SAW yarn tension sensor. *IEEE Trans Electron Devices* 59(10):2853–2855
4. Zhu J, Guo T. The dynamic thermal response model of Pt100 temperature sensor
5. Hoshino K, Mori D Ieee. (2008) Three-dimensional tactile sensor with thin and soft elastic body. In: *IEEE workshop on advanced robotics and its social impacts*. Taipei, TAIWAN, pp 125–130
6. Zhang XW, Wang FY, Li L (2007) Optimal selection of piezoelectric substrates and Crystal cuts for SAW-based pressure and temperature sensors. *IEEE Trans Ultrason Ferroelectr Freq Control* 54(6):1207–1216
7. Ochoa J (1989) A practical determination of CTD platinum resistance thermometer response time, and its use to correct salinity bias and spikes. *Deep Sea Res Part A* 36(1):139–148
8. Kimoto A, Sugitani N, Fujisaki S (2010) A Multifunctional Tactile Sensor Based on PVDF Films for Identification of Materials. *IEEE Sens J* 10(9):1508–1513
9. Yong Y, Patel M, Masako Tanaka M. Estimation of quartz resonator Q and other figures of merit by an energy sink method. *IEEE Trans Ultrason Ferroelectr Freq control* 54(7):1386–1398

Super-Gaussian Random Vibration Test Inductive Method Based on Measured Data

Peng Wang, Hua Deng, Liuqing Xiao and Guotao Liu

Abstract Toward the situation that the traditional inductive method is not suitable for super-Gaussian data, this paper presents an inductive method of super-Gaussian vibration data of locomotive equipment according to the actual measured data of Lanxin line and the actual operation situation. By verifying the measured vibration data, it is found that the data shows a steady random supergrid characteristic. The test conditions of body installation equipment of locomotive under super-Gaussian random vibration are put forward by analyzing the inductive values and the kurtosis value, which solves the problem of using the measured vibration data to compute the simulation conditions of the super-Gaussian random vibration environment.

Keywords Measured data · Data induction · Rolling stock equipment
Super-Gaussian vibration · Vibration test

1 Introduction

Vibration environment is a significant factor that affected the operation, RAMS (reliability, availability, maintainability, and safety) and life expectancy of locomotive equipment. Therefore, the vibration simulation of the equipment according to the actual situation has an irreplaceable role in inspecting and testing product's quality and ensuring the RAMS of the vehicle.

At present, Chinese locomotive equipment vibration test standard GB /T 21563 [1] (idt IEC 61373) gives the Gaussian random vibration conditions, but vibration environment data of some special line shows significant super-high characteristics.

P. Wang (✉) · H. Deng
School of Mechanical and Electrical Engineering, Central South University,
No. 932 Lu Shan, Yue Lu District, Changsha, China
e-mail: 3bfz@163.com

P. Wang · L. Xiao · G. Liu
CRRC Zhuzhou Electric Locomotive Research Institute Co., Ltd.,
No. 169 Shi Dai, Shi Feng District, Zhuzhou, China

Since the Gauss vibration (kurtosis is 3) exceeds 3 times of RMS value that is only 0.27% of the time, the super-Gauss vibration with kurtosis of 7 exceeds 3 times of RMS value that is up to 1.5%, and most of the product's fatigue damage is caused by a peak acceleration of 2–4 times the root mean square (RMS). In this station, Gauss vibration will significantly reduce the actual use of fatigue damage of the equipment. So it is necessary to measure the vibration in specific line and then construct a more realistic super-Gaussian random vibration environment. At present, the commonly used vibration data induction methods include the extreme envelope method and the statistical tolerance method [2–4]. The extreme envelope method is too harsh. The statistical tolerance method proposed by GJB /Z 126 [4] is a leap in the induction of vibration data from the extreme envelope to the statistical method. But this method is only applicable to Gaussian vibration. Therefore, it is necessary to redevelop the suitable induction method of super-Gaussian vibration data for the locomotive equipment.

2 Vibration Environment Measurement

2.1 Test Line

Lanzhou West–Urumqi West is selected as the measured line on which the similar locomotive equipment has a high failure rate. Test line contains 10 stations (9 line intervals). The running speed is 60–120 km/h, which is tested for three cycles, and the total test mileage is about 5800 km.

2.2 Measuring Point Arrangement

In GB/T 21563, the test levels are divided into 1 class body mounted, 2 bogie installation, and 3 axle installation, according to the installation location area of the locomotive equipment. Referring to the rule of area division of GB/T 21563, the test sets 12 measuring points in the 1 class body mounting area of locomotive. The measuring points are distributed in the typical body mounting area such as the driver's cab, the mechanical room, the roof, the car bottom, and the side wall. Meanwhile, this article defines the locomotive running direction as X (longitudinal direction), the left and right directions of the running direction as Y (horizontal direction), and the up and down directions of the running direction as Z (vertical direction).

3 Test of the Measured Data

The 100G raw data is obtained by testing, and its stationarity, periodicity, ergodic, and normality are tested before processing the measured data.

3.1 Stationary Test

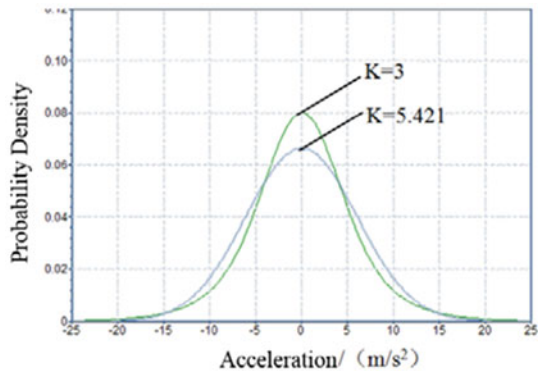
First, the measured vibration data of 8000 s is randomly divided into 40 segments, and the standard deviation of each segment is calculated and arranged in chronological order: 0.61838, 0.30188, 0.66049, 0.54905, 0.67923, and 0.62726. The average value of the 40 standard deviations is 0.639578, greater than the average value counts “+”, and less than the average counts “-”. The number of statistics rounds is 19. By checking the round distribution table, we know that when a is equal to 0.05, its confidence interval is (14, 27). So the measured data meets the stationarity assumption and passes through the stationary test.

3.2 Periodic Test

The purpose of the periodic test is to determine whether the measured vibration data contains the periodic components in order to separate them. Physical test, visual inspection, probability density analysis or autocorrelation analysis, self-spectral density analysis, and variance test are the commonly used test method [2]; this paper adopts the probability density analysis method.

The probability density functions of random, sinusoidal, and random superimposed sinusoidal data are bell-shaped, bowl-shaped, and saddle [5–7]. Since the probability density function of the measured vibration data (shown in Fig. 1, $k = 5.421$) is bell-shaped, it can be seen that it has randomness.

Fig. 1 Probability density function of measured vibration data (bilinear coordinates)



3.3 Normality Test

The probability density function analysis method, the chi-square goodness-of-fit test method, and the partial peak state test method are normality (Gaussian) test methods, and this paper adopts the last method. The kurtosis value of the random variable x can be obtained from Eq. (1):

$$k = \frac{E[x - E(x)]^4}{\{E[x - E(x)^2]\}^2} \tag{1}$$

The kurtosis value k is 3 when the random variable x follows the Gaussian distribution. When the k is greater than 3, it is super-Gaussian distribution.

The random vibration data is taken randomly, and the kurtosis is calculated by formula (1), and the value k is equal to 5.421 (the probability density function is shown in Fig. 1). It can be seen that the measured vibration data is obviously hyper-Gauss distribution.

4 Power Spectral Density Calculation

The time-domain data should be converted into frequency-domain data, and the power spectral density (PSD) of each set of data should be calculated. In this paper, the PSD is calculated using the efficient Welch method [5, 8] (shown in Fig. 2).

In Fig. 2, $x_i(n)$ is a section of data of length of i , here $i = n/k$.

By calculating the average PSD of each segment, the estimation value $P(w)$ of PSD is obtained, as shown in Eq. (2)

$$P(\omega) = \frac{1}{KU} \sum_{k=1}^K P_k(\omega) \tag{2}$$

Here, $P_k(\omega) = \frac{1}{i} \left| \sum_{n=1}^i x_k^N(n) \cdot W(n) \cdot e^{-j\omega n} \right|^2$, k presents the overlapping segments, $U = \frac{1}{i} \sum_{n=0}^{i-1} W^2(n)$

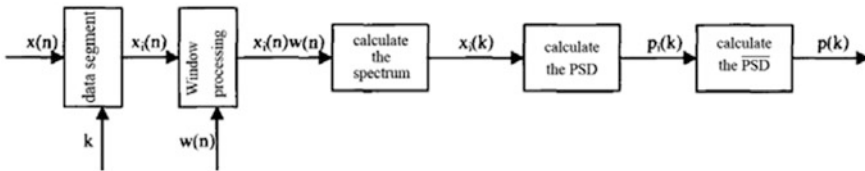


Fig. 2 Welch method to estimate the PSD

5 Induction of Measured Data

A large number of vibration measurement data and analysis results can reflect the comprehensive true vibration characteristics of a certain type of equipment only after proper inductive processing [9, 10]. However, the current extreme value envelope method and the statistical tolerance method do not suit for super-Gaussian vibration data [2–4]. Therefore, a method for data induction of super-Gaussian vibration for locomotive equipment is proposed as follows.

5.1 The Induction of the Measured Data at the Same Point on the Same Road Section

According to the Welch method, the PSD of the multiple measured of the same point is calculated, and then the PSD is summarized by the linear average method according to formula (3):

$$\bar{G}(f) = \frac{1}{n} \sum_{i=1}^n G_i(f) \quad (3)$$

Here, n presents the number of measurements on the same road section, $G_i(f)$ presents the i -th measured PSD at the frequency f , and $\bar{G}_i(f)$ presents the average value of the measured PSD at the frequency f .

5.2 Data Induction of Several Measured Points on the Same Installation Area on the Same Road Section

The inductive method is defined as follows: r measuring points are arranged in the same installation area at the same road section, and then the r PSDs are arranged on the same frequency axis. At the same frequency point, the corresponding PSD values are G_1, G_2, \dots, G_r . \bar{G} and \bar{S} are defined as follows:

$$\bar{G} = \frac{1}{r} \sum_{i=1}^r G_i \quad (4)$$

$$\bar{S} = \sqrt{\frac{1}{r-1} \sum_{i=1}^r (G_i - \bar{G})^2} \quad (5)$$

Here, G_i is the measured PSD of the i -th point at the same installation area on a certain frequency, \overline{G} is the average value of r measurement points' PSD at the same installation area on a certain frequency, \overline{S} is the standard deviation of r measurement points' PSD at the same installation area on a certain frequency, and r is the number of measured point at the same installation area.

According to the confidence coefficient a , the upper confidence limit of the average value of measured PSD is obtained as follows:

Where $1 - \alpha$ is the specified confidence coefficient, here $\alpha = 0.1$, $t_{r-1, \alpha/2}$ is the value found by the degree of freedom r and b from the t distribution.

The values of $G_{1-\alpha}^u$ at each frequency point can be obtained by Eq. (6). Arranging all $G_{1-\alpha}^u$ on the frequency axis yields an inductive PSD that contains information about the r measurement data in the same installation area:

$$G_{1-\alpha}^u = \overline{G} + \frac{\overline{S} \times t_{r-1, \alpha/2}}{\sqrt{r}} \tag{6}$$

5.3 Measurement Data Induction at the Same Installation Area on Different Road Sections

The operating mileage of the locomotive equipment corresponds to its service life. Therefore, when the measured data are inducted, the proportion of the measured data should be determined according to the mileage of the section, that is, the longer the mileage of the section, the greater the proportion in the induction data.

For this reason, the weighted square method can be used to induct the measured data of multiple road sections in the same installation area. The total test line (such as Lanzhou West–Urumqi West) is separated into m segments by station (such as Lanzhou West–Wuwei South for one). M measured PSD can be obtained on the total test line according to the methods put forward in Sects. 5.1 and 5.2. Then, the measured PSD is arranged on the same frequency axis, and the corresponding PSD value $G_1, G_2 \dots G_m$ at a certain frequency point can be obtained. $G_1, G_2 \dots G_m$ can be inducted into Eq. (7)

$$G_L = \sum_{i=1}^m \gamma_i G_i \tag{7}$$

Here, γ_i is the weighted coefficient of the road section I, $\gamma_i = l_i/L$, $\sum_{i=1}^m \gamma_i = 1$, l_i is the mileage of the i -th section of the test line (such as Lanzhou West–Wuwei South) (km), L is the total mileage of test line (such as Lanzhou West–Urumqi

West) (km), G_L is the PSD induced by measured data of M sections at a certain frequency point, G_i is the PSD induced by measured data of a certain section at a certain frequency point, and m is the number of segments of the test lines.

The G_L value at each frequency point can be obtained by Eq. (7). An inducted PSD can be obtained by arranging all the values on the frequency axis. The PSD contains the data of the m sections in the same installation area, and the longer the mileage of a certain section, the larger the proportion of the measured data in the summarized PSD.

5.4 Kurtosis Statistics of Measured Data

In engineering, kurtosis is used to describe the super-Gauss random process, and kurtosis is defined by Eq. (1). It can be seen that kurtosis is a parameter that describes the characteristics of amplitude probability density curve of random process [11, 12].

In kurtosis statistics, the typical measured data of X , Y , and Z directions can be extracted, respectively, and kurtosis values are obtained by formula (1), respectively, and then the average value of kurtosis of X , Y , and Z directions is calculated.

5.5 Construction of Measured Super-Gauss Random Vibration PSD

According to the method in Sects. 5.1–5.3, a number of measured data is induced into one PSD, and the average kurtosis value is calculated by the method of Sect. 5.4. Then, the super-Gauss random vibration test condition of the locomotive equipment with X , Y , and Z directions is built.

6 Engineering Examples

The measured Lanzhou West–Urumqi West line can be divided into nine sections, a total of three tests. 12 measuring points are arranged in the installation area of 1 class body mounted. Before the induction of the measured data, the measured data is tested with smoothness, randomness, the state of history, and super-Gaussian properties according to the method in Chap. 4, and the PSD of the measured data is calculated by Welch method.

6.1 Induction of the Measured Data at the Same Point on the Same Road Section

According to the linear average method in Sect. 5.1, the PSD of the X, Y, and Z directions of each measurement point is induced, and the PSD of each measurement at a certain point is shown in Fig. 3.

6.2 Induction of Several Measured Points on the Same Installation Area on the Same Road Section

According to the method in Sect. 5.2, the measured data of X, Y, and Z directions of 12 measuring points in 1 class body mounted are induced. The PSD of the X-direction of the measured data from the Lanzhou West to Wuwei South is shown in Fig. 4.

6.3 Induction at the Same Installation Area on Different Road Sections

The measured line is divided into nine sections, and the mileage of each section and weighted coefficients calculated by Eq. (7) are shown in Table 1.

Using the weighted coefficients in Table 1, the X, Y, and Z data of 1 class body mounted in nine sections are induced according to the weighted average method in Sect. 5.3, the Z-direction PSD of the nine sections of the measured data is shown in Fig. 5.

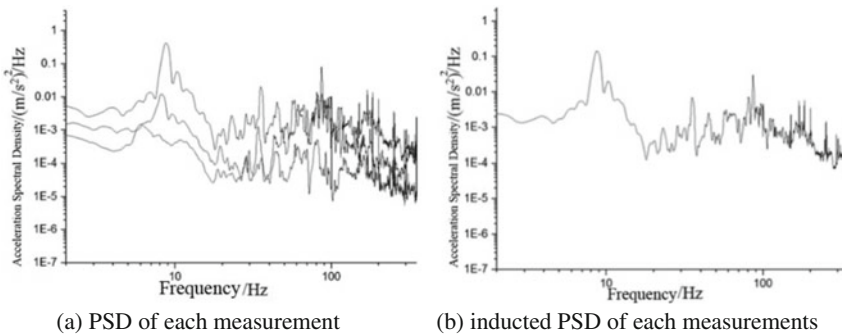
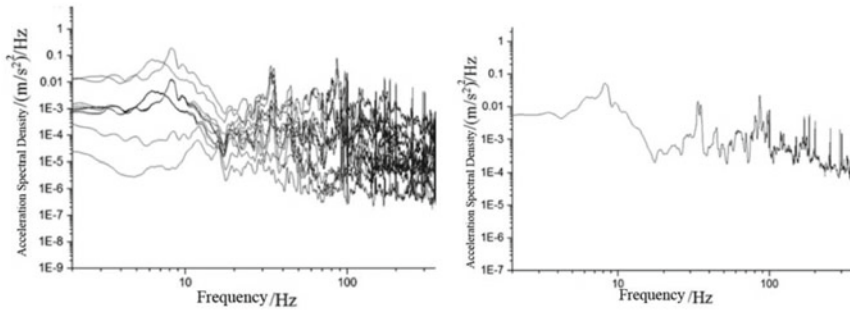


Fig. 3 PSD of the data of each measurement at a certain point

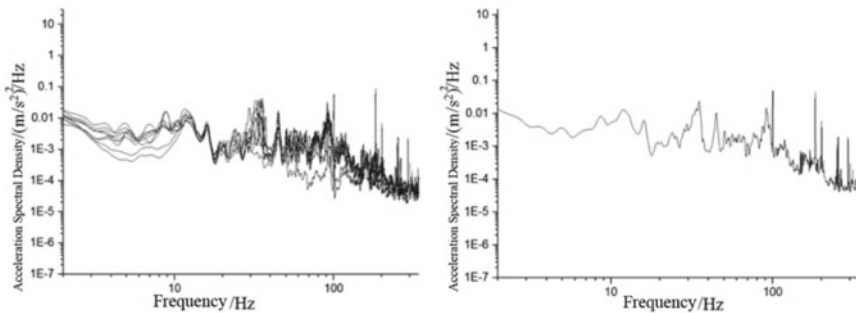


(a) PSD of each measuring point data (b) induced PSD of each measuring point data

Fig. 4 Induced PSD of X-direction of 1 class body mounted in Lanzhou West–Wuwei South section

Table 1 Mileage and weighting factors for each section

Number	Station		Mileage/km	Weighted coefficients (γ)
1	Lanzhou West	Wuwei South	277	0.142
2	Wuwei South	Zhangye	257	0.132
3	Zhangye	Jiayuguan	223	0.114
4	Jiayuguan	Yumen	132	0.068
5	Yumen	Liu Yuan	165	0.085
6	Liu Yuan	Hami	321	0.165
7	Hami	Shanshan	263	0.135
8	Shanshan	Turpan	148	0.076
9	Turpan	Urumqi West	162	0.083
10	Total	/	1948	1.000



(a) PSD of the measured data of 9 sections (b) induced PSD of the measured data of 9 sections

Fig. 5 The Z-direction PSD of the nine sections of the measured data

Table 2 Kurtosis statistics

Data number	Direction	Kurtosis value k	Average kurtosis \bar{k}
01	X	6.46459	6.48
02		6.26224	
03		6.38436	
04		6.46459	
05		6.43592	
06		6.86638	
07	Y	7.85564	6.58
08		6.33004	
09		7.07417	
10		7.24693	
11		5.54652	
12		5.4207	
13	Z	6.29286	6.81
14		6.95777	
15		5.70755	
16		6.56908	
17		7.41236	
18		7.92011	

6.4 Statistics of Kurtosis Values of Measured Data

Many groups of measured time-domain data are extracted from nine sections, and the kurtosis is calculated using LMS Test Lab software according to Eq. 1. The result is shown in Table 2, the average kurtosis values of X, Y, and Z directions are 6.48, 6.58, and 6.81, respectively.

6.5 Construction of Measured Super-Gauss Random Vibration PSD

The PSD of X, Y, and Z directions is deduced according to the method in Sects. 6.1–6.3, then the average kurtosis value \bar{k} is obtained, as shown in Table 2; finally, PSD of the measured super-Gauss random vibration of locomotive equipment is constructed, as shown in Figs. 6, 7, and 8.

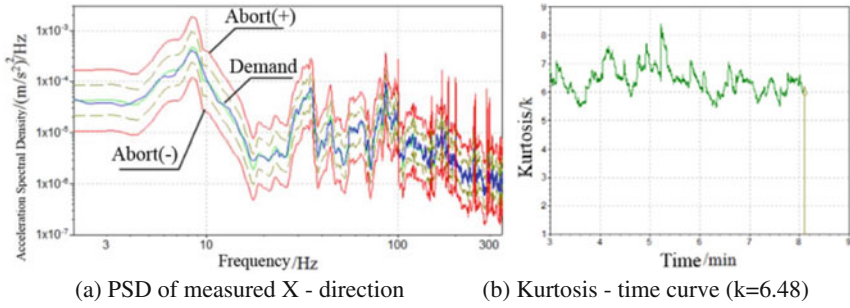


Fig. 6 PSD of measured X-direction of super-Gaussian vibration

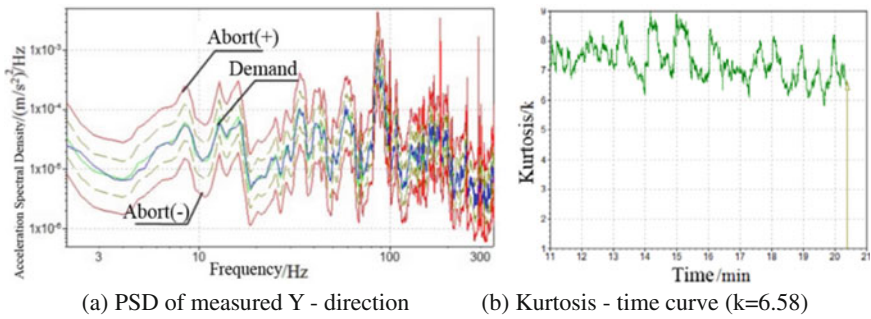


Fig. 7 PSD of measured Y-direction of super-Gaussian vibration

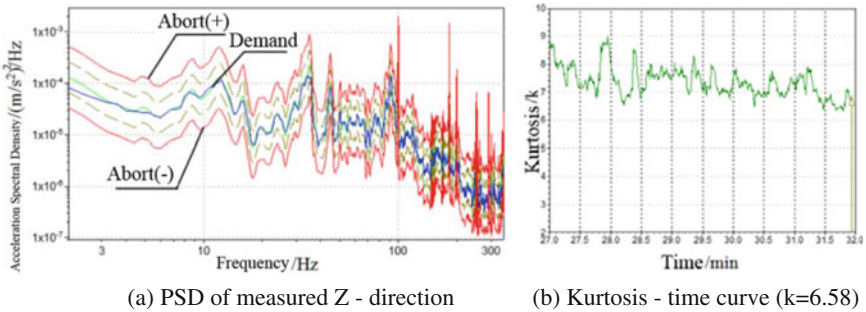


Fig. 8 PSD of measured Z-direction of super-Gaussian vibration

7 Conclusion

There are three commonly used methods to determine the test conditions of vibration environment, such as standard regulation, the data estimation of the similar products, and the induction of the measured data. It is recommended to use the last method to induct the test conditions [13, 14]; therefore, the induction method plays a vital role in the formulation of test conditions. At present, the locomotive equipment is usually tested by GB/T 21563, and few measured data is used to induce the vibration test conditions, and the vibration data of super-Gauss is seldom induced. Based on the measured vibration data of locomotive and the actual running situation of locomotive equipment on Lanxin line, this paper puts forward the induction method of super-Gaussian random vibration data of locomotive equipment, sets up the conditions of super-Gaussian vibration test, and puts forward an engineering solution for setting up the simulation environment of super-Gauss random vibration using the measured data of locomotive equipment, and it has a good engineering application value.

Acknowledgements This work was supported in part by the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2015BAG14B01).

References

1. Standards Press of China (2008) GB/T 21563-2008 Railway applications–Rolling stock equipment–Shock and vibration tests. Beijing (in Chinese)
2. Standards Press of China (1990) GB 10593.3-90, Measurement method of environmental parameters for electric and electronic Products, vibration data processing and the induction method. Beijing (in Chinese)
3. China Aero Palytechnology (1984) HB/Z 87-84 The inductive methods for environmental measured data of aircraft flying. Beijing (in Chinese)
4. Standards Press of China (1999) GJB/126-99, The inductive method for environment measured data of vibration and shock. Beijing (in Chinese)
5. Feng Z (2005) Vibration/Temperature data processing and program design of environmental spectrum for aircraft airborne equipment. Chinese academy of agricultural mechanization sciences, Beijing (in Chinese)
6. Zhu Y, Xiao Y, Su Y (2013) Application research of vibration data processing technology for armed vehicle platform induced environment. Equip Environ Eng 10(3):72–76 (in Chinese)
7. Sun Z (2011) Study on environmental stress induction technique for equipment. Command control & simulation 33(05):108–113 (in Chinese)
8. Feng Z, Yan C, Zhang S (2006) Date processing and compilation of vibration environmental spectrum for aircraft equipment. Acta Aeronaut Astronaut Sin 27(02):289–293 (in Chinese)
9. Liu K (2014) Study on the method of making free flight vibration profile for air-to-air missile based on measured data. Equip Environ Eng 11(05):114–118 (in Chinese)
10. Mu L, Huang H, Zhang J (2007) Vibration test and data processing of truck compartment. Equip Environ Eng (04):27–30 (in Chinese)

11. Van Baren J, Van Baren P (2009) The fatigue damage spectrum and Kurtosis control. Vibration Research Corporation, Jenison, MI
12. Vibration Research Corporation. Random Control. KURTOSION™—Getting The Kurtosis Into The Resonances. Vibration Research Corporation
13. Standards Press of China (1999) GJB 150.16A-2009, Laboratory environmental test methods for military materiel part 16: vibration test. Beijing (in Chinese)
14. Standards Press of China (1999) GJB 899A-2009, Reliability testing for qualification and production acceptance. Beijing (in Chinese)

RAM Assessment of CTCS-1 ATP Based on Enhanced Bayesian Network

Yue Xu and Zhongtian Liu

Abstract China Railway Corporation proposes the upgrading and updating plan of general speed train (train running speed under 160 km/h) to improve the outdated status of general speed railway. The new CTCS-1 Automatic Train Protection (ATP) system will be safer, more reliable and advanced than the existing train control system. In this paper, enhanced Bayesian Network (BN) is first used in CTCS-1 ATP RAM assessment, which is mainly used to analyze the structure of two different ATPs and establish the corresponding enhanced BN models. Finally, the Reliability, Availability, and Maintainability (RAM) of these two different ATPs are obtained and analyzed through enhanced BN simulation model. The assessment results show that the RAM indicators of CTCS-1 modular redundancy ATP and system redundancy ATP can satisfy the requirements of Chinese train control system requirement specifications, but the RAM indicators of CTCS-1 modular redundancy ATP are better than CTCS-1 system redundancy ATP.

Keywords Enhanced Bayesian network · CTCS-1 ATP · RAM

1 Introduction

In order to realize the full autonomy of train control system key technology, China Railway Corporation puts forward a major issue of upgrading general speed train control equipment: CTCS-1 Train Control System Key Technology Research [1]. To make sure the reliability and safety of ATP, the participants of CTCS-1 research project, such as China Academy of Railway Sciences, CRSC company, Hollysys company, use different redundancy design schemes to meet its reliability and security requirements. CTCS-1 ATP equipment will adopt modular, distributed design, and its Safety Integrity Level (SIL) is required to achieve SIL4 [2, 3].

Y. Xu (✉) · Z. Liu

School of Electronic and Information Engineering,

Beijing Jiaotong University, Beijing 100044, People's Republic of China

e-mail: xuyueted@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_50

497

The existing literature on reliability of CTCS is mainly focused on CTCS-2 and CTCS-3 train control equipment. However, reliability research of CTCS-1 train control system is still blank in China, and this paper takes CTCS-1 ATP as the research object, using enhanced Bayesian Network (BN) to model CTCS-1 modular redundancy ATP and system redundancy ATP, applying enhanced BN to establish ATP models and evaluating these two different kinds of ATP RAM indicators, assessing whether the RAM indicators of these two redundancy schemes can meet current train control system requirement specification, and which scheme is better.

In this paper, brief introduction of enhanced BN and CTCS-1 ATP system is introduced first. Second, establishing CTCS-1 ATP system redundancy reliability block diagram and CTCS-1 ATP modular redundancy reliability block diagram. Third, the RAM indicators of these two different ATPs are acquired, and the results are compared with current train control system requirement specification, and finally estimating whether these two schemes can meet the specification requirements, and assessing which schemes is better.

2 Enhanced Bayesian Network Modeling Theory

There are two modeling methods of enhanced BN: direct establishment method and model transformation method. Direct establishment method can be used to build enhanced BN model when system structure and relationship between modules are known, reliability block diagrams can be established according to system structure, and then system-enhanced BN model can be built on the base of reliability block diagrams. However, for some large complex systems, the structure has many levels, and dependency relationship between modules is not clear, so the model built through direct establishment method may be not correct, finally resulting in subsequent analysis error. Therefore, for some complex large-scale systems, it is necessary to establish its enhanced BN model through model transformation method according to the existing model when the internal structure and modules dependency relationship are not known. As CTCS-1 ATP structure diagram can be obtained through Integral Technical Scheme of CTCS-1 Train Control System, therefore, its enhanced BN model is built through direct establishment method.

3 Structure of CTCS-1 ATP

CTCS-1 ATP is a new-generation ATP equipment of general speed train, and its components are shown in Fig. 1. According to Integral Technical Scheme of CTCS-1 Train Control System [4], Track Circuit Reader (TCR) & TCR-Antenna (TCR-Ant), Balise Transition Module (BTM) & BTM-Antenna (BTM-Ant), Speed-Distance Unit (SDU), GSM for Railway (GSM-R) & GSM-R Antenna (GSM-R Ant), CTCS-1 Control Module (C1-CM) and CTCS-0 Control Module

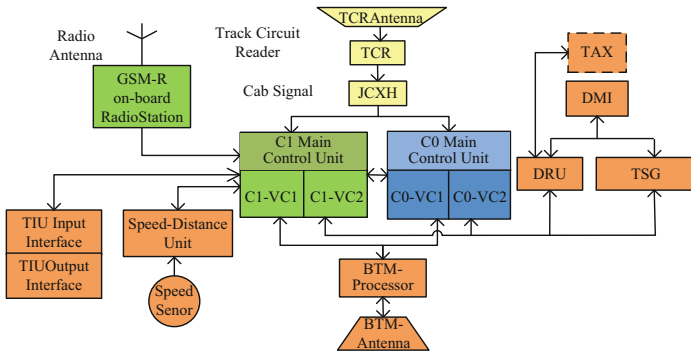


Fig. 1 Structure diagram of CTCS-1 ATP

(C0-CM) are all associated with train safety operation, and they are all safety-critical units. Driver Machine Interface (DMI) and SDU do not provide information for generating monitoring and control curve, they are non-safety-critical units. For safety-critical units, they need hot standby configuration; for non-safety-critical units, they are configured through cold standby pattern.

4 Establishment of CTCS-1 ATP-Enhanced BN Model

4.1 Establishment of CTCS-1 Enhanced BN Model

4.1.1 Establishment of CTCS-1 ATP Reliability Block Diagram

At present, CTCS-1 ATP mainly adopts modular redundancy and system redundancy schemes to ensure the reliability and safety of ATP. In order to assess the reliability of these two ATP schemes, corresponding enhanced BN model is established on the basis of the reliability block diagrams below, and the reliability evaluation is carried out upon ATP-enhanced BN model.

Building CTCS-1 Modular Redundancy ATP Reliability Block Diagram

In modular redundancy system, modules are the basic units to configure ATP, and for safety-critical modules, they are configured through hot standby pattern; finally, different modules are connected in series and parallel. The reliability block diagram is built according to CTCS-1 modular redundancy ATP, and it is shown in Fig. 2.

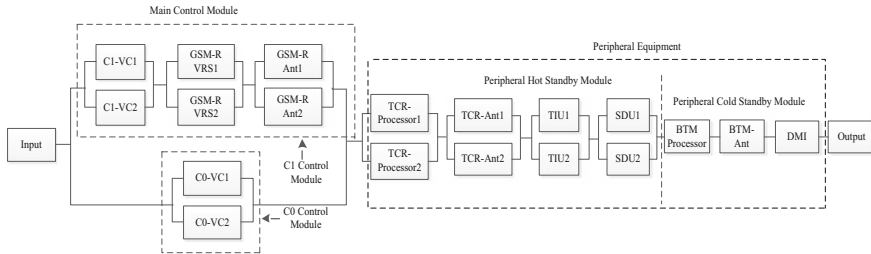


Fig. 2 Enhance BN model of CTCS-1 modular redundancy ATP

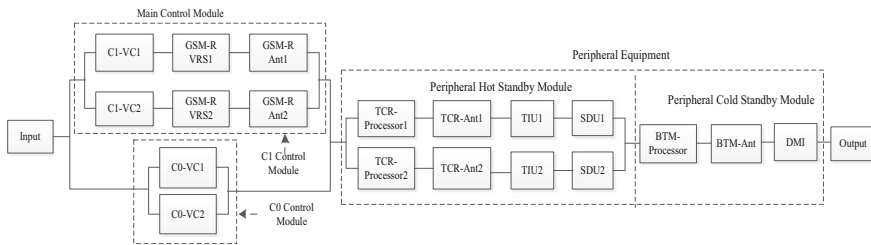


Fig. 3 Reliability diagram of CTCS-1 system redundancy ATP

Building CTCS-1 System Redundancy ATP Reliability Block Diagram

CTCS-1 system redundancy ATP also includes two parts: Main control module and peripheral equipment module. Differently, basic units of main control module are connected in series first, and then the series structure is configured through hot standby pattern. Peripheral equipment module is configured in the same way, and finally, main control module and peripheral equipment module are connected in series. CTCS-1 system redundancy ATP reliability diagram is shown in Fig. 3.

4.1.2 Establishment of CTCS-1 ATP-Enhanced BN Model

CTCS-1 modular redundancy ATP-enhanced BN model and system redundancy ATP-enhanced BN models are established according to reliability block diagrams shown in Figs. 2 and 3, and system redundancy ATP-enhanced BN models are shown in Figs. 4 and 5.

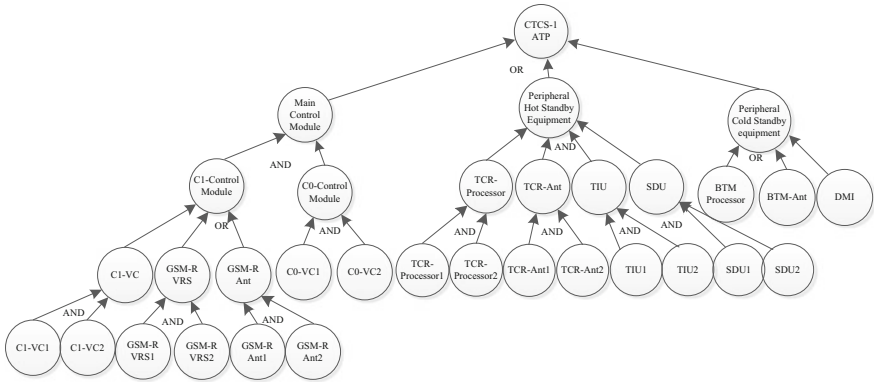


Fig. 4 Enhance BN model of CTCS-1 modular redundancy ATP

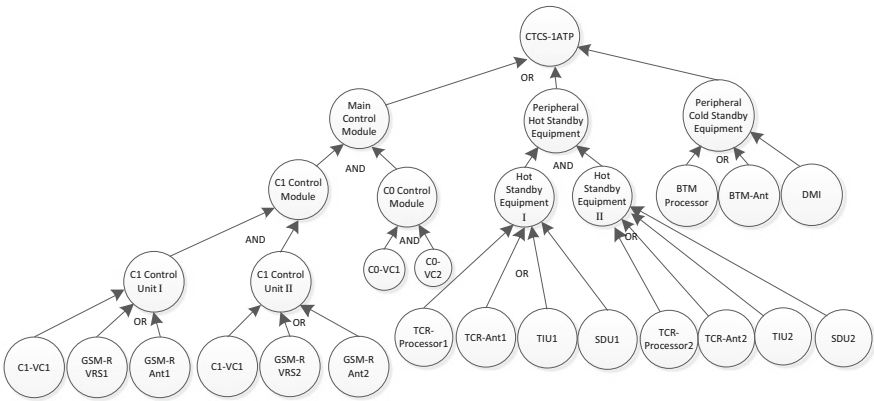


Fig. 5 Enhance BN model of CTCS-1 system redundancy ATP

4.2 RAM Evaluation of CTCS-1 ATP

CTCS-1 train control system requirement specification is still in development; therefore, RAM indicators of CTCS-1 ATP cannot be assessed according to CTCS-1 system requirement specification now. As electronic components of CTCS-3 and CTCS-1 ATP are basically the same, and their reliability indexes and software architecture are similar, Hollysys company and CRSC company are developing CTCS-1 ATP based on CTCS-3 ATP hardware platform; therefore, CTCS-3 train control System Requirement Specification (SRS) can be used as CTCS-1 ATP RAM indicators assessment criteria. The RAM indicators discussed here are Mean Time between Failure (MTBF), Mean Time to Repair (MTTR), availability, reliability, failure rate λ , and repair rate μ . ATP RAM requirements of Chinese train control system requirement specification [5, 6] are shown as follows:

1. $MTBF \geq 10^5$ h;
2. $A \geq 99.99\%$
3. $MTTR \leq 10$ h

The average lifetime of train control equipment is in accordance with exponential distribution, so reliability of basic unit is $R = e^{-\lambda t}$, $MTBF = 1/\lambda$, and $MTTR = 1/\mu$. Basic unit RAM indicators are calculated using the following formulas:

$$\text{Reliability } R = e^{-\lambda t} \tag{1}$$

$$\text{Availability } A = \frac{MTBF}{MTBF + MTTR} = \frac{\lambda}{\lambda + \mu} \tag{2}$$

$$\text{Unavailability } U = 1 - A = 1 - \frac{\lambda}{\lambda + \mu} = \frac{\mu}{\lambda + \mu} \tag{3}$$

$$\text{MTTR } MTTR = \frac{MTBF}{A} = MTBF \tag{4}$$

Basic unit reliability indexes are shown in Table 1. Reliability parameters of basic units used in this section (λ and μ) are from DI “Research on the Evaluation Method for the RAM Goals of CTCS-3” [5]. Reliability, unavailability, and MTTR of basic units are calculated according to the existing reliability parameters λ and μ . Cold standby units of basic units marked by “*” cannot work immediately, when the working unit fails, so it is considered as mono-system, while the hot standby units can be enabled immediately when the working part breakdown; it is considered as dual system. Basic units of reliability parameters are shown in Table 1.

Table 1 Basic units reliability index ($t = 2 \times 104$ h)

Basic unit ($x = 1,2$)	λ/h^{-1}	μ/h^{-1}	R	U	MTTR (h)
DMI*	5.00×10^{-6}	2	0.9048	2.50×10^{-8}	0.5
SDU _x	2.50×10^{-9}	0.25	0.9999	1.00×10^{-8}	4
BTM-CU	2.00×10^{-6}	2	0.9607	1.00×10^{-6}	0.5
BTM-Ant*	7.00×10^{-8}	0.25	0.9986	2.80×10^{-7}	4
TIU _x	2.10×10^{-5}	2	0.7993	1.05×10^{-5}	0.5
TCR _x	2.30×10^{-6}	2	0.9550	1.15×10^{-6}	0.5
TCR-Antx	7.50×10^{-8}	0.25	0.9985	3.00×10^{-7}	4
C1-CU _x	1.49×10^{-5}	2	0.7423	7.45×10^{-6}	0.5
VRS _x	1.20×10^{-5}	2	0.7866	6.00×10^{-6}	0.5
GSM-RAntx	1.45×10^{-8}	2	0.9997	7.25×10^{-6}	0.5
C0-CU _x	1.20×10^{-5}	2	0.7866	6.00×10^{-6}	0.5

* Cold standby unit

Mathematical model of system-level reliability:

$$P(A = 0) = \sum_{i=1}^n P(E_i = 0) \cdot P(A = 0|E_i = 0) \tag{5}$$

In the formula above, “0” stands for the normal state, “1” stands for the failure state, $P(E_i = 0)$ represents the availability of basic unit in normal state, $P(A = 0|E_i = 0)$ represents the availability of CTCS-1 ATP when E_i in normal state, and $P(A = 0)$ is the overall availability of CTCS-1 ATP. Formula 5 is the Bayesian probability formula, and it is a bottom-up causal reasoning mathematical model, from which the overall availability of CTCS-1 ATP can be acquired.

MSBNX is a simulation software based on BN theory; it supports node state definition and definition of logical relationship between adjacent nodes and definition of root node prior probability; and it has bidirectional reasoning function. In this paper, Microsoft MSBNX software is used to simulate CTCS-1 ATP-enhanced BN model, and RAM assessment is also conducted by MSBNX.

4.2.1 RAM Indicators Calculation of CTCS-1 Modular Redundancy ATP

The simulation result in Fig. 6 shows the availability of CTCS-1 modular redundancy ATP:

Availability : $A = 99.9996\%$

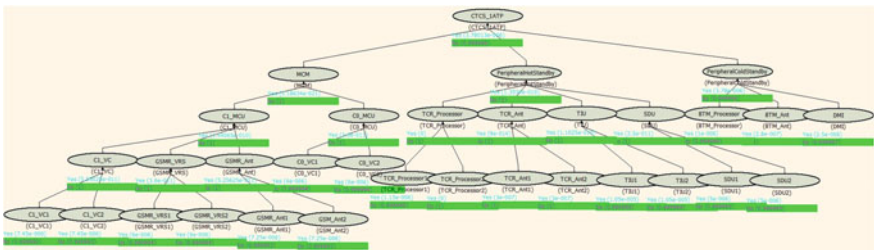


Fig. 6 Simulation and availability calculation of CTCS-1 modular redundancy ATP model in MSBNX

Similarly, input reliability R of basic units in ATP simulation model and the reliability of modular redundancy ATP can be obtained:

$$\text{Reliability : } R = 0.8275$$

Modular redundancy ATP failure rate λ can be obtained through formula (1):
Failure rate:

$$\lambda = -\frac{\ln R}{t} = -\frac{\ln 0.8189}{2 \times 10^4} = 9.98 \times 10^{-6}$$

$$\text{MTBF} = 1/\lambda = 1.06 \times 10^5 (\text{h})$$

As $A = 99.9996\%$, $\text{MTBF} = 1.06 \times 10^5$, and

$$A = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$

$$\text{MTTR} = \frac{\text{MTBF}}{A} - \text{MTBF}$$

So MTTR of CTCS-1 ATP can be obtained: $\text{MTTR} = 0.42 \text{ h}$.

4.2.2 RAM Indicators Calculation of CTCS-1 System Redundancy ATP

CTCS-1 system redundancy ATP model can also be simulated in MSBNX based on Fig. 5. The simulation and calculation diagram is not shown here due to the length of the article.

Availability of CTCS-1 system redundancy ATP is

$$\text{Availability : } A = 99.9996\%$$

Similarly, input reliability of basic units in ATP simulation model and the reliability of system redundancy ATP can be obtained:

$$\text{Reliability : } R = 0.8189$$

Modular redundancy ATP failure rate is obtained through formula (1):

$$\text{Failure rate : } \lambda = \frac{\ln R}{t} = -\frac{\ln 0.8189}{2 \times 10^4} = 9.98 \times 10^{-6}$$

$$\text{MTBF} = 1/\lambda = 1.06 \times 10^5 (\text{h})$$

Table 2 RAM indicators comparison between CTSCS-1 modular and system redundancy ATP

SRS RAM requirements	Modular redundancy ATP RAM indicators	System redundancy ATP RAM indicators	Meet specification requirements	Which is better
$MTBF \geq 10^5$ h	1.06×10^5 h	1.00×10^5 h	Yes	Modular redundancy ATP
$A \geq 99.99\%$	99.9996%	99.9996%	Yes	Identical
$MTTR \leq 10$ h	0.424 h	0.4	Yes	Basically same
R	0.8275	0.8189	—	Modular redundancy ATP
λ	9.49×10^{-6}	9.98×10^{-6}	—	Modular redundancy ATP

As values of A and $MTBF$ have been acquired, and $A = MTBF/(MTBF + MTTR)$, so we can get $MTTR$ of CTCS-1 system redundancy ATP:

$$MTTR = 0.4 \text{ h.}$$

4.2.3 RAM Indicators Comparison Between Two Different Redundancy ATP

RAM indicators comparison result among CTCS-1 modular redundancy ATP, system redundancy ATP, and CTCS system requirement specification are displayed in Table 2; we can see that RAM indicators of CTCS-1 modular ATP and system redundancy ATP can meet the requirements of current CTCS requirement specification, and some indicators are much better than the required values. From Table 2, $MTBF$ and reliability values of modular redundancy ATP are greater than that of system redundancy ATP, but their availability values are equal. It is known that maintainability of CTCS-1 two different ATPs is basically the same, and they are much less than the request 10 h. Overall, CTCS-1 modular redundancy ATP design scheme is more excellent in RAM aspect.

5 Conclusion

Fault logic gates are introduced to enhanced BN on basis of traditional BN, which enable enhanced BN keep the property of Fault Tree; therefore, CTCS-1 ATP-enhanced BN model can be established according to the corresponding reliability block diagrams. CTCS-1 modular redundancy ATP and system redundancy

ATP models are simulated with the help of Microsoft Bayesian simulation tool MSBNX; meanwhile, RAM indicators of these two different ATPs are calculated. Main conclusions of this paper are as follows:

1. Evaluation results show that the availability of CTCS-1 modular redundancy ATP and system redundancy ATP is identical, and MTTR values of these two ATPs are much smaller than the request 10 h; their maintainability is basically the same, but CTCS-1 modular redundancy ATP reliability indicators such as MTBF and failure rate are better than that of system redundancy ATP. Although these two different ATPs RAM indicators can satisfy the requirement of CTCS train control system requirement specification, in comparison, CTCS-1 modular redundancy ATP RAM indicators are much better.
2. Inverse inference of CTCS-1 ATP-enhanced BN model is conducted with the help of MSBNX, and it is found that conditional failure rate of peripheral cold standby module is much higher than that of other modules, which is a potential cause of CTCS-1 ATP failure. Specifically, DMI and BTM-Processor conditional failure rates are higher; they are the weak link of CTCS-1 modular redundancy ATP and system redundancy ATP. It is suggested that designers should configure DMI through hot standby pattern and configure BTM-Processor via chip-level hot standby pattern so that the RAM indicators of CTCS-1 ATP can be optimized.

Acknowledgements This Research is supported by Scientific and Technology development plan of China Railway Corporation (Grant No.2015X002-A).

References

1. China Railway Corporation (2015) Research on the key technology for CTCS self-dependence —Contract of CTCS-1 train control system key technology research (in Chinese)
2. CENELEC (1998) EN-50126-1998 railway applications: the specification and demonstration of Reliability, Availability, Maintainability Related and Safety (RAMS), October[S]. CENELEC, Brussels
3. CENELEC (2003) EN-50129-2003 Railway applications: safety related electronics systems for signaling, December[S]. CENELEC, Brussels
4. China Academy of Railway Sciences et al (2015) Integral technical scheme of CTCS-1 Train Control System (V0.1): Beijing, [P/OL] (in Chinese)
5. Di L, Yuan X, Wang Y (2010) Research on the Evaluation Method for the RAM Goals of CTCS-3. China Railway Science 31(6):91–97 (in Chinese)
6. Department of Science and Technology, Ministry of Railways. No. 127 of Science and Technology Transport File (2008). CTCS-3 Train Control System Requirements Specification SRS (V1.0). Department of Science and Technology, Ministry of Railways, Beijing (in Chinese)

A Novel Online Measuring System for Wheel Size Based on Laser Displacement Method

Yong Zhang and Changgeng Wang

Abstract A new wheel size online detection system for subway vehicle based on laser displacement method is introduced to solve problems such as large workload and low accuracy of manual measurement for metro wheelset. This paper describes the system structure, proposes the design of hardware and software, and expounds the principle of using laser displacement sensor to realize the measurement of the thickness of wheel flange, height of wheel flange, and wheel diameter parameters. In order to verify the usability of the system, the field round test and the vehicle test are carried out, and the measurement results between the online system and workman are compared. The results show that the detection system has the advantages of accurate measurement and good consistency, and can replace the existing manual measurement to meet the requirements of wheel size measurement.

Keywords Metro · Wheel size · Laser displacement · Online measurement

1 Introduction

As the coupling component of train and the track, urban rail train wheels carry the entire weight of vehicle and ensure the safe train operation [1]. The wheels sustain the impact and pressure of wheelset-track contact when the vehicle is running. Wheel flange wear increases seriously especially when the vehicle is crossing the connection part of track as well as turning and braking. It will pose a threat to the safe operation and ride comfort of vehicle when the wear exceeds a certain limit [2]. So it is of great significance to obtain the vehicle wheel parameters accurately and quickly.

The detection method of wheel size mainly includes contact measurement and noncontact measurement. Contact measurement apparatus include the fourth

Y. Zhang (✉) · C. Wang

Guangdong Provincial Key Laboratory of Modern Geometric and Mechanical Metrology Technology, 510405 Guangzhou, China
e-mail: 13503035170@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_51

507

checker [3, 4] and electronic type laser measuring instrument [5]. It mainly depends on manual operation of workers. With the disadvantage of high labor intensity, long maintenance time, and low efficiency, the method is easy to bring in hidden trouble [6]. To overcome the blemish of contact measurement, many domestic and abroad institutions develop various un-contact measure systems. The institutions in America Japan, Australia, etc. [7–9], have developed an online measuring system based on light section image method. The system contains laser source, CCD camera, wheel detector, auto train identification system, and air cleaner. When the system works, the CCD camera shots the line laser projecting on the wheel at high speed, the wheel size will obtain after filtering, lining and extract the centreline of image. Some institution domestic (Zhengzhou University of Light Industry and Beijing Jiaotong University) [10–12] also adopted similar principle to develop online detection system for wheel size which has been put into application in some bullet train and metro corporation. However, the system is hard to meet the high reliability of wheelset measurement, because of complex structure and poor anti-interference performance. Some domestic research institutes have carried out laser displacement sensors wheelset measurement, but it is limited to laboratory studies, and has not yet formed the actual product [13, 14]. An Italian company achieves accurate measurement of wheelset by using eight laser vision sensors and error compensation technology. But the cost is quite expensive. A Swiss company [15] achieves the auto measurement of tread parameter and diameter. But the cost is too high, besides, the system needs to replace the track into nonbearing guide rail which does not meet domestic safety requirements.

With the rapid development of society and computer technology, the manual measurement of high labor intensity and low measurement precision will be eliminated, light screenshots of complex structure and poor anti-interference ability as the method will gradually withdraw from the stage of history. Online measuring system based on laser displacement sensor of new generation with high precision and strong anti-interference ability certainly will be a trend in the future.

2 Design of Online Measurement System for Wheels

The online measuring system consists of three parts. As shown in Fig. 1. Detection sensor located in the trackside, data acquisition device located in equipment room, and data processing device located in the control room. First trackside sensor converts wheel size and field control signals into electrical signals by data acquisition device. Then, the digital signal is transferred to server in control room by optical fiber transmission. Finally, the results including wheel size obtained by algorithm processing and vehicle number will be published by Web server.

The trackside detection sensor includes laser displacement sensor, axial position sensor, and auto equipment identification (AEI). The installation layout is shown in Fig. 2. When the vehicle passes through the axial position sensor P1, the auto equipment identification is activated. At the same time, the laser displacement

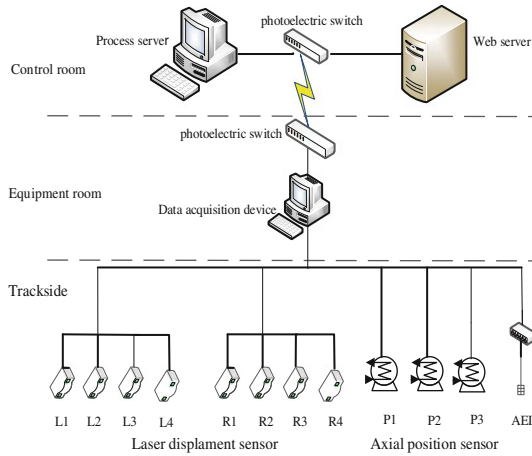


Fig. 1 Schematic diagram of system structure

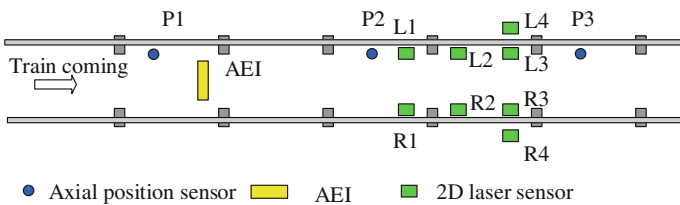
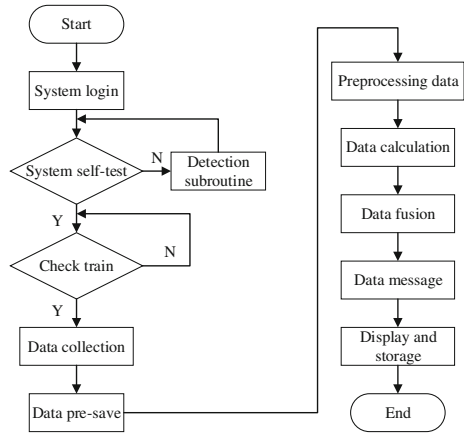


Fig. 2 Equipment installation layout

sensor starts preheating. When the train passes through AEI, the radio frequency identification tag system automatically reads the vehicle, thus completing the identification function. When the train passes through the axial sensor P2, the shield window of the laser sensor is opened, the laser displacement sensor is ready to collect wheel size. When the wheels passes through the detection area which the laser displacement sensor were placed one by one. The laser displacement sensor start the measurement of wheel dimension. L1–L4 and R1–R4 displacement’s laser sensor finish wheel dimension data acquisition. When the last wheel leaves axial position sensor P3, the system shuts AEI, shield window and laser displacement sensor down, waiting for the next measurement.

The development of system software uses C++ development tools. To achieve the function of online system, the system contains data collection, data storage, data processing, data display, and data transmission. (1) The data acquisition module achieve the data collection of various kind by controlling and communicating between data acquisition card and AEI. (2) The data storage module stores the data acquired by data acquisition module and processed wheel size in the form of documents. (3) Data processing module fuses and processes sensor data and AEI signals to get the wheel parameters. (4) The data display the thickness of wheel flange, height of wheel flange, and wheel flange thickness by man–machine

Fig. 3 System software flow chart



interface (5) Data transmission module sends the processed results to the Web server (Fig. 3).

3 Function Principle of Online Measuring System for Wheelset

Wheel size online detection system adopts the laser triangulation method, mainly to realize the measurement of the thickness of wheel flange, height of wheel flange, and wheel diameter.

3.1 Measurement Principle of Wheel Flange Parameters

The height of wheel flange and the thickness of wheel flange are obtained primarily by the laser displacement sensors L_3 , L_4 . As shown in Fig. 4, the two laser displacement sensors are mounted on both sides of the track at an angle and distance, and the two laser displacement sensors emit the laser surface coplanar. Due to the occlusion of the wheel flange, the individual laser displacement sensor can only obtain partial information of the tread contours. The outer laser displacement sensor L_4 cannot acquire the information on the inner side of the wheel, and the inner laser displacement sensor L_3 cannot acquire the inner side information of the rim. Therefore, it is necessary to combine the measurement results of the two laser displacement sensors L_3 and L_4 on the inner and outer sides to obtain complete tread contour line information. The fusion diagram of the two laser displacement sensors is shown in Fig. 5.

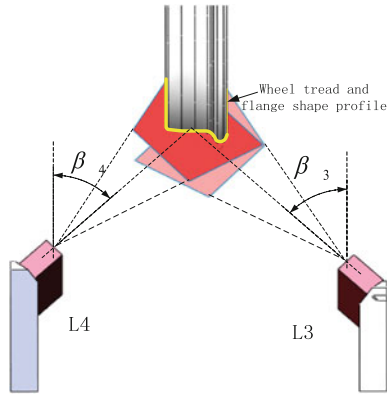
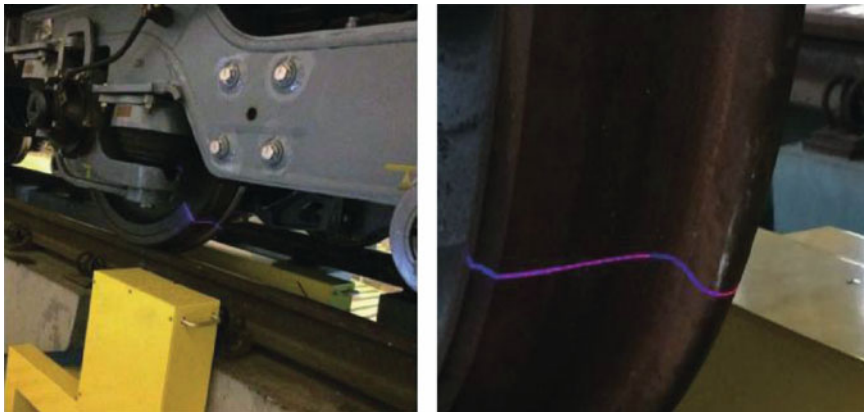


Fig. 4 Wheel flange size measurement schematic diagram

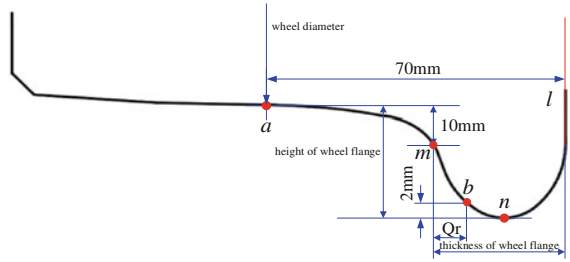


(a) Overall schematic diagram (b) Partial enlarged schematic diagram

Fig. 5 Laser displacement sensor fusion schematic diagram

The tread data is obtained by the two laser displacement sensors L_3 , L_4 (take left wheel as example), then the tread contour is converged by coordinate transformation which is shown in Fig. 6. Select the key measurement points such as the vertex of the wheel flange n , the reference point of the thickness of wheel flange m , and the tread base point a , so that the height of wheel flange and the thickness of wheel flange are calculated according to the wheel flange parameter definition.

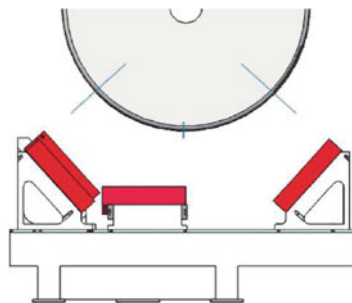
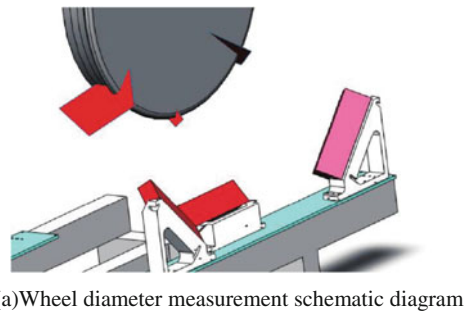
Fig. 6 Tread and key point schematic diagram



3.2 Measurement Principle of Wheel Diameter Parameters

The wheel diameter parameters are mainly obtained by laser displacement sensors L_1, L_2, L_3 . As shown in Fig. 7, three laser displacement sensors are respectively installed in orbit inside at a certain angle and distance, L_1 and L_3 are mirror symmetric. When the wheel enters the valid detection interval of the laser displacement sensor, three laser displacement sensors separately measure the three vertices of the outer edge of the wheel D_1, D_2, D_3 . According to the coordinate transformation and three-point co-circle method, it is easy to determine the diameter of the outer edge of the wheel, and then subtract twice the height of wheel flange to get the wheel diameter value.

Fig. 7 Wheel diameter measurement schematic diagram



(b) Wheel flange three vertex measurement schematic diagram

4 Test and Verification

Wheel size online detection system is installed in the red sand depot of Guangzhou Metro Corporation. In order to verify the availability of the system, wheel test and vehicle passing test have been carried out.

4.1 Wheel Test

The wheel test is to place a single wheel on the track, uniform speed push the wheel through the detection system, and carry on artificial measurement. In this experiment, A, B two rounds were measured five times. Test results are shown in Table 1, this table comes from the test report, which is issued by the Guangdong Institute of Metrology.

The maximum and minimum values of the height of wheel flange in the system five measurements are 0.11 and 0.09 mm, the difference between the thickness of wheel flange are 0.04 and 0.09 mm, the difference between the wheel diameter are 0.36 and 0.27 mm. The system can meet the requirement of consistency. Five times average measurement data is with manual measurement average values, the difference between the height of wheel flange are 0.03 and 0.03 mm, the difference between the thickness of wheel flange are 0.05 and 0.04 mm, the difference between the wheel diameter are 0.13 and 0.20 mm. The measurement accuracy of the system can meet the requirements of the scene.

Table 1 The total entropy calculation results of each grouping experiment

轨轮 Wheel		轮缘高 Height of wheel flange	轮缘厚 Thickness of wheel flange	轮径 Wheel diameter
A	标准值/Standard Value	28.04	31.70	844.49
	示值/Indicating Value	28.07	31.65	844.62
	均值误差/Mean Error	0.03	0.05	0.13
	最大误差/Maximum Error	0.11	0.04	0.36
	标准差/Standard Deviation	0.03	0.03	0.10
	95%置信度偏差/Confidence Coefficient Deviation	0.05	0.06	0.17
B	标准值/Standard Value	28.04	31.69	844.71
	示值/Indicating Value	28.07	31.73	844.91
	均值误差/Mean Error	0.04	0.04	0.19
	最大误差/Maximum Error	0.09	0.16	0.27
	标准差/Standard Deviation	0.03	0.06	0.05
	95%置信度偏差/Confidence Coefficient Deviation	0.06	0.07	0.21

Table 2 The average of vehicle passing test system and manual measurement

Wheel number	Average value of the thickness of wheel flange		Average value of the height of wheel flange		Average value of the wheel diameter	
	System	Manual work	System	Manual work	System	Manual work
1	30.52	30.40	28.21	28.27	937.35	837.63
2	30.48	30.46	28.27	28.40	837.12	837.33
3	30.60	30.47	28.27	28.39	837.59	837.59
4	30.78	30.62	28.27	28.43	837.45	837.57
5	28.67	28.69	28.06	28.22	817.22	817.36
6	29.05	29.15	28.10	28.28	817.83	817.93
7	29.16	29.01	28.04	28.18	818.70	818.96
8	29.08	29.17	28.12	28.33	818.45	818.56

4.2 Vehicle Passing Testing

Vehicle passing test refers to the actual train at low speed through the wheeled size online detection system. In this experiment, eight rounds of the same train were tested six times, test results are shown in Table 2.

The experimental results show that the system meets the requirements of the actual field detection accuracy and replace the traditional manual detection.

5 Conclusion

This paper introduces a new wheel size online measurement system, gives the design of the system, expounds the principle of system measurement, and analyzes in detail the field test and test. The experimental results show that the system has high consistency and accuracy, and can meet the needs of the wheel size measurement. It can be used as a new technology and system to replace traditional manual measurement and light screen image measurement. The wheel size online detection system is installed in the library. The train speed is limited, the speed is only 0–5 km/h. The further propose is that the system is installed in the entry section line and improves the speed to 30 km/h.

References

1. Chen G, Ren G (2012) Online detecting system of wheelset state. Urban Rail transit research (10):79–81,124 (in Chinese)

2. He Y (2015) The design of wheel sets abrasion detection technology based on machine vision. Beijing Jiaotong University, Beijing (in Chinese)
3. Deli F, Wang X (1998) Development of fourth types of detector for LLJ-4 wheel. *Railway Vehicle* 2:42–44 (in Chinese)
4. Liang J (2013) Design and application of simplified wheel flange checker for rail vehicles. *Urban Rail Transit Res* 10:127–128 (in Chinese)
5. He Z (2011) Development of portable intelligent measuring instrument for wheel tread parameters. *Railw Oper Technol* 17(3):56–57, 60 (in Chinese)
6. Ceng Xiaodong (2014) Research on application of Performance Detecting Sys-tem for passenger train in overhaul. *Res Dev* 23(3):10–12 (in Chinese)
7. Ben D, Peng X (1999) Development of Automatic Measuring Device for Wheel Tread Shape. *Foreign Locomot Roll Stock Technol* 2:43–46 (in Chinese)
8. Huang X (2008) Detection technology of vehicle wheel tread based on image processing and machine vision. Central South University, Changsa (in Chinese)
9. Zhang Y, Wang L, Gao X (2002) A review of wheel tread damage detection technologies in and out of China. *Locomot Veh Technol* (1):1–4, 8 (in Chinese)
10. Xiong H (2013) Successful application and improvement of passenger wheel inspection system in riverside locomotive depot. *Electr Lo-comotives Mass Transit Veh* 36(6):72–74, 77 (in Chinese)
11. Rong J, Maozi Z (2010) Application of online inspection system for wheel size and its technical standard. *Modern Urban Rail Transit* 3:30–32 (in Chinese)
12. Wu T, Fangcheng W (2009) Automatic measuring system for wheel set geometric parameters based on laser triangulation method. *Mech Eng* 7:92–94 (in Chinese)
13. Gao Y (2014) Research on methods for automatically and dynamically meas-uring parameters of train wheels. Beijing Jiaotong University, Beijing (in Chinese)
14. Zhifeng Z, Lu C, Fangzhen Z (2012) A Novel Method for non-contact measuring diameter parameters of wheelset based on wavelet analysis. *Optik* 123:433–438
15. Ministry of Railways Standard Measurement Research Institute (2003) TB/T 449-2003 Locomotive vehicle wheel tread shape. China Railway Press, Beijing (in Chinese)

Sub-problems Interaction Analysis-Based Three-Level Decomposition Algorithm for Real-Time Train Scheduling and Routing Problems in Railway Stations

Lijie Bai, Thomas Bourdeaud'huy, Emmanuel Castelain, Qi Zhang and Ziyuan Liu

Abstract The management of rail traffic in stations requires careful scheduling to fit to the existing infrastructure, while avoiding conflicts between large numbers of trains and satisfying safety or business policy and objectives. The train scheduling and routing problem studied includes four tasks: scheduling, routing, platforming and conflict resolution. We propose a three-level decomposition method based on the interactional relationship among four tasks mentioned above to enhance the computational efficiency. This method is tested on full-day timetable obtained from a real-world station.

Keywords Train routing · Train scheduling · Conflict resolution
Decomposition method

1 Introduction

A railway station with complex entering and leaving routes network is a busy system with increasing patterns of train services that require accurate scheduling and routing to adapt to the limited infrastructures. At peak hours, the infrastructure

L. Bai · Q. Zhang (✉)

National Research Center of Railway Intelligence Transportation System
Engineering Technology, China Academy of Railway Sciences, Beijing, China
e-mail: zhangqi@rails.cn

L. Bai

Beijing Hua-Tie Information Technology Development Co, Beijing, China
e-mail: lijie.bai@hotmail.com

T. Bourdeaud'huy · E. Castelain

LAGIS - Laboratoire d'Automatique, Génie Informatique et Signal,
École Centrale de Lille, Cité Scientifique, Villeneuve d'Ascq, France

L. Bai · Z. Liu

Signal & Communication Research Institute, China Academy
of Railway Sciences, Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_52

is operated nearly at capacity margins. When unexpected disturbances and disruptions affect the normal course of daily operations, dispatchers need to regenerate a feasible timetable by rerouting, delaying or even cancelling some trains to reduce the propagation of delay.

D'Ariano [1] has developed a real-time dispatching system, called ROMA, to automatically recover disturbances. ROMA is able to automatically control traffic, evaluating the detailed effects of train reordering [2] and local rerouting [3] actions, while taking into account minimum distance headway between consecutive trains and the corresponding variability of train dynamics [4–6]. Corman et al. [7] focus on the real-time CDR (conflict detection and resolution) problem through junctions and proposes a constraint programming formulation for the combined routing and sequencing problem. Caprara [8] develops two new routing neighbourhoods of different size in order to search for more effective routings and studies their structural properties in a tabu search scheme. Caprara [9] and Cacchiani et al. [10] treat train platforming problem considered for the case of multiple routes where platform times can vary in a discrete interval. Bai et al. [11] propose an integer linear programming model (ILP) for the TTP in railway node.

This paper studies the scheduling and routing problem at the operational level faced by railway station managers to generate a conflict-free non-cyclic timetable which consists of two sets of circulations. One set is made of commercial circulations given by regional levels. The other set corresponds to technical circulations (shunting) added by the railway station managers to prepare or repair the trains.

The problem that we address in this paper is NP-hard and can be stated as follows. Given the layout of a railway station, arrival and departure times, as well as the destination and origin of trains, we aim to schedule the technical circulations within the allowable deviations and to route as many as possible trains through the station. The blocking trains should be pointed out if a solution for all trains cannot be found. The solution must ensure that no pair of trains is conflicting over routes and tracks, while allowing the coupling and uncoupling of trains at tracks and respecting their preferences of platforms.

To solve our problem, we update the timetable under rolling-horizon framework to adapt to the dynamic railway traffic, as stated in our previews article. To complete the update calculation in short time, we propose sub-problems interaction analysis-based three-level decomposition algorithm.

2 Problem Formalization

A railway station contains a set L^e of lines located outside the home signals at the entrance of the station, a set L^i of lines next to the platforms and a set of switches. A route $p \in P^{(i,e)}$ contains a set of ordered switches $p = [s_1^p, s_2^p, \dots, s_{|p|}^p]$ which connects a platform line l^i with a section line l^e . The traffic in the railway station is defined by a set of trains T . Every train t contains a set of ordered movements

$M^t = [m_1^t, m_2^t, \dots, m_{|M^t|}^t] \subset M$. The index of a movement represents its chronological order, for example m_1^t occurs before m_2^t .

Four types of movements are defined depending on their commercial or technical nature and their direction as shown in Table 1. The set of movements M is divided thus into four subsets $M = M^{CL} \cup M^{TL} \cup M^{CE} \cup M^{TE}$.

A train t contains at least two movements $[m_1^t, m_2^t]$: one entering movement (from section line to platform line) and one leaving movement (from platform line to section line). The reference arrival and departure times of trains, given by administrative levels and station manager, are represented by reference starting and ending times of their movements $[\alpha_m^{ref}, \beta_m^{ref}]$. While ensuring the minimum operational time on tracks implicitly given by the ideal timetable, a deviation interval L^{dev} is permitted for the technical circulations depending on the direction. The relationship between the reference time and flexible adjustment time interval $[\alpha_m^{flex}, \beta_m^{flex}]$ is resumed in Table 2.

Speed variation dynamics of trains running in stations is not taken into account. Occupation of routing resources is detailed in Fig. 1. The route occupied is locked until the train leaves the last switch of the route $[\alpha_m, \beta_m]$. Route locking time is fixed as $S_{m,p}$ according to the train length, route length and speed limitation on routes. The track reserved is locked from the train entering the station until the track totally released $[A_t, B_t]$.

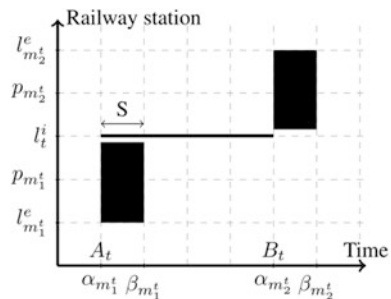
Table 1 Movements classification

	Commercial	Technical
Leaving	M^{CL}	M^{TL}
Entering	M^{CE}	M^{TE}

Table 2 Flexible adjustment time interval of movements

Type of movements	α_m^{flex}	β_m^{flex}
$M^{CL} \cup M^{CE}$	α_m^{ref}	β_m^{ref}
M^{TL}	α_m^{ref}	$\beta_m^{ref} + L^{dev}$
M^{TE}	$\alpha_m^{ref} - L^{dev}$	β_m^{ref}

Fig. 1 Occupation of routing resources



In our problem, the coupling and uncoupling operations of trains at tracks are taken into account. The choice of tracks for trains depends on the trains' characters, passenger service and usual practices in stations. We provide a set of tracks L_t^{Pref} in priority order for each train which combines the length compatibility and the direction preference.

The ideal timetables, including reference times and section lines, are generally conflicting, while we struggle to route all trains through the station. In order to resolve these conflicts, we are allowed to modify (anticipate or delay) the technical movements within flexible adjustment time interval $[\alpha_m^{flex}, \beta_m^{flex}]$, to alleviate the conflicts on paths, but it causes the increase of stopping time at tracks. In some complex and busy stations, a conflict-free timetable is not guaranteed. So trains cancellation is permitted, ensuring that a feasible solution is always obtained. The objective function is to minimize the number of trains cancelled. The detailed mathematical model can be found in our previous paper. In this paper, we try to improve the computational efficiency by implementing three-level decomposition algorithm.

3 Three-Level Decomposition Algorithm

The problem described in our paper consists of three parts of decision: scheduling, routing and cancellation. The cancellation decision is supported by both scheduling and routing decisions. A wise combination of scheduling and routing decisions can efficaciously reduce the number of trains cancelled. However, the combination of scheduling and routing decisions represents a huge solutions search space. In order to reduce the calculation complexity, we divide the problem into two parts: scheduling-cancellation sub-problem and routing-cancellation sub-problem, so as to avoid the huge amount of combinations between scheduling and routing variables. At the same time, the correlation between the two sub-problems cannot be ignored. In this case, a three-level decomposition method shown in Fig. 2 is designed as a circle mechanism to satisfy the demand of independence and cooperation between scheduling and routing in the resolution process.

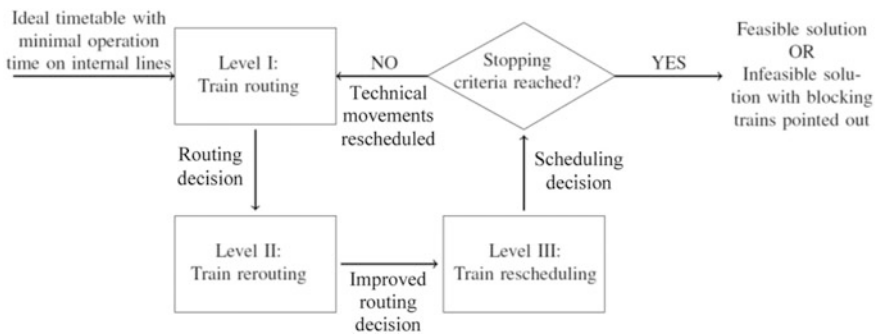


Fig. 2 Architecture of closed-loop iterative three-level framework

In each level, we propose a mixed integer linear programming (MILP) model or an integer linear programming (ILP) model to describe the sub-problems. Level I, modelled as ILP, solves the routing-cancellation sub-problem with the given scheduling decision. Level II, modelled as MILP, is responsible for the smooth cooperation between the two sub-problems. Succeeding the given scheduling decision and the cancellation decision solved in Level I, Level II reroute the trains to maximize the margin of scheduling decisions for trains cancelled in Level I. Level III, modelled as MILP, succeeds the routing decision solved in Level II and reschedule the trains to minimize the number of trains cancelled.

Three-level decomposition method simplifies the whole problem in two ways. On one hand, the huge amount of combinations between scheduling and routing decisions are avoided. In each sub-problem, the solution search space is narrowed. On the other hand, the cooperation relationship between routing and scheduling is recovered by Level II which tries to reroute trains and movements with the cancellation decision solved in Level I, so as to maximize the scheduling margin for Level III.

4 Experiments and Results

To observe the performance of three-level decomposition algorithm, we solve a full-day timetable containing 247 trains more than 500 movements in a real-world station with 15 platform lines, 10 section lines and 310 routes as shown in Fig. 3. The same instance is tested in our previous paper [11]. The feasible timetable is generated with 9 trains cancelled by cumulative sliding window algorithm. The solve time is 2496 s. In this paper, we use three-level method to solve the full-day problem. Maximum iteration of three-level method is set to 10.

The computational results are obtained by using CPLEX version 12.6 on a 64 bits computer under Linux with Intel i5-2520 M CPU at 2.5 GHz and 8 GB memory RAM. The calculation time is limited to 500 s.

Fig. 3 Railway station topology

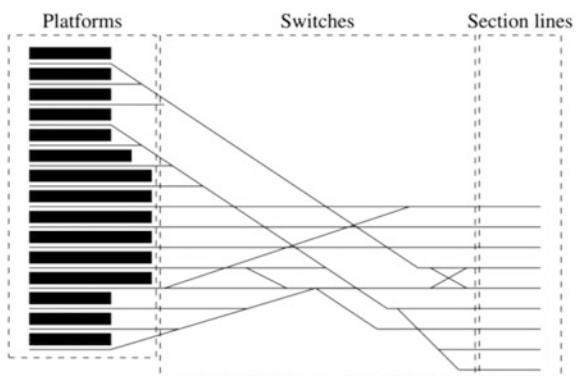
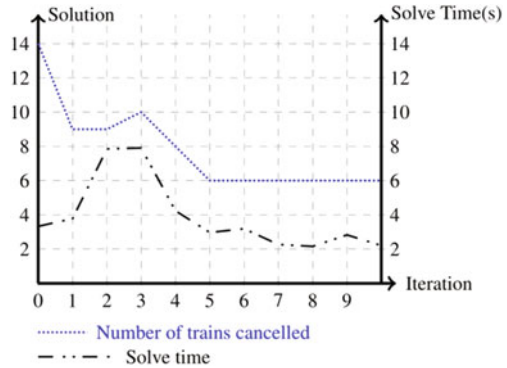


Fig. 4 Recovery of full-day timetable by three-level method



The results are displayed in Fig. 4. The double-ordinate chart represents the solution and the solve time of each iteration. The abscissa is the iteration count. The solve time is read on the right ordinate. The calculation of one loop can be completed in 2–8 s. The calculations in three levels are stopped by the optimal solution while GAP is 0%. Number of trains cancelled in Level III is read on the left ordinate. The best solution found at the 5th iteration is 6 trains cancelled. We can see that three-level method greatly reduces the solve time from 2496 to 30 s. And the number of trains cancelled is 6 trains instead of 9 trains. Compared with results of our previous paper [11], both of quality of solution and computational efficiency are improved greatly.

5 Conclusion

In this paper, we present the train scheduling and routing problem through a railway station. To ensure the feasibility of timetable, cancellation of trains is permitted to eliminate insolvable conflicts. Three-level decomposition algorithm based on sub-problems interaction analysis efficiently improves the performance of adapting timetable in railway stations. The complete algorithm can be easily applied to different railway stations.

In further work, more real cases in different railway stations will be collected and studied to evaluate the efficiency of the complete algorithm.

Acknowledgements We would like to thank the anonymous reviewers for their helpful and constructive remarks. This work is supported in part by Science Foundation for Young Scientists of CARS (Grant no. 2016YJ065).

References

1. D'Ariano A (2008) Improving real-time train despatching: model, algorithms and applications [ph. D], TU Delft; D'Ariano A, Pacciarelli D, Pranzo M (2007a) A branch and bound algorithm for scheduling trains in a railway network. *Europ J Oper Res* 183(2):643–657
2. D'Ariano A, Corman F, Pacciarelli D, Pranzo M (2006b) Real-time train conflict detection and resolution: global sequencing and local rerouting. In: van Zuylen HJ (ed) *Proceedings 9th TRAIL congress, selected papers*. Delft University Press, pp 77–72
3. D'Ariano A, Albrecht T (2006a) Running time re-optimization during real-time timetable perturbations. In: Allan J, Brebbia CA, Rumsey AF, Sciutto G, Sone S, Goodman CJ (eds) *Computers in railways X*. WIT, Southampton, pp 531–540
4. D'Ariano A, Pranzo M, Hansen I (2007) Conflict resolution and train speed co-ordination for solving real-time timetable perturbations. *IEEE Trans Intell Trans Sys* 8(2):208–222
5. D'Ariano A, Pranzo M, Hansen I (2008) Reordering and local rerouting strategies to manage train traffic in real time. *Trans Sci* 42(4):405–419
6. Rodriguez J (2007) A constraint programming model for real-time train scheduling at junctions. *Transp Res Part B* 41(2):231–245
7. Corman F, D'Ariano A, Pacciarelli D, Pranzo M (2010) A tabu search algorithm for rerouting trains during rail operations. *Transp Res Part B* 44:175–192
8. Caprara A (2010) Almost 20 years of combinatorial optimization for railway planning: from lagrangian relaxation to column generation. In: Erlebach T, Lbbecke M (eds) *Proceedings of the 10th workshop on Algorithmic Approaches for Transportation Modelling, Optimization and Systems (ATMOS) 2010*, vol 1. Schloss Dagstuhl, Germany, pp 1–12
9. Caprara A, Galli L (2011) Solution of the train platforming problem. *Trans Sci* 45(2):246–257
10. Cacchiani V, Furini F, Kidd M (2016) Approaches to a real-world train timetabling problem in a railway node. *Omega* 58:97–110
11. Bai L, Bourdeaud'huy T, Castelain E (2015) Automatic one-day timetable generation in busy and complex railway station with cancellation processing: platforming, routing and scheduling. In: *15th IFAC Symposium on Information Control in Manufacturing (INCOM 2015)* Ottawa, Canada, 11–13 May 2015 (EI: 20160201797448)

Train Operation Intervals Optimization Based on a Group Dynamics Model

Xuelei Meng, Bin Zhang and Limin Jia

Abstract Train operation interval is a most important parameter in the railway management work and train dispatching work. It also determines the capacity of the railway sections. In this paper, a train interval optimization model based on group dynamics theory is proposed. It takes the potential energy of the artificial potential field as the optimization object, which is defined with the train operation interval. And the train operation rules, especially the intervals between arrivals and departures, are designed as the constraints of the model. An immune clone algorithm is designed to solve the model. A computing case is presented, and we got the optimal intervals of trains on the timetable. The computing case proves the effectiveness of the approach. This paper presents a new method to design train intervals for timetable designers and dispatchers.

Keywords Train operation · Intervals · Group dynamics model

1 Introduction

Group dynamics is a new science and is closely related with the current research of complex systems. We get the general rule groups through the analysis on individual behavior, the influence between the individuals and the overall characteristics of the groups.

X. Meng

School of Traffic and Transportation, Lanzhou Jiaotong University,
No. 88 Anning West Road, Lanzhou, Gansu Province, China

B. Zhang

Department of Technology and Information, Shuo Huang Railway
Development Co., Ltd., Sunning, Hebei Province, China

L. Jia (✉)

State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University,
No. 3, Shangyuan Cun, Haidian District, Beijing, China
e-mail: jialm@vip.sina.com

At present, the research of group dynamics theory and group dynamics is getting more and more attention from the researchers. And its application in engineering fields is being studied extensively.

The scholars have built several important models to study the group dynamics, such as the Boid Model (BM). The main idea of BM is the velocity matching and collision rule to reach the coordinate status in the condition of lacking global information and central control. Thus, the train dispatching and control problem can be described with the BM. The trains are taken as the group with dynamics, and we should analyze each train movement and the influence between them. Then, we master the characteristics of the train group and give the optimal dispatching strategy and the scheme.

The dynamics theory is more and more utilized in engineering problems. Barreiro-Gomez et al. proposed a novel methodology for solving constrained optimization problems in a distributed way inspired by population dynamics [1]. Schwarz et al. did simulation experiments to show that maintenance of the upgraded infrastructure, the scope of upgrading efforts, and timing (early vs. late investments) affected infrastructure quality [2]. João et al. presented a robust approach to the planning of regional wastewater systems under population dynamics uncertainty [3].

And there are much fuzziness and uncertainty in the engineering computation, which leads to various computing results and different decision consequence. The decision-makers also have subjective preference in their work, which is usually uncertain. It is quite difficult to describe the uncertainty factors in engineering computing fields. It is typical in the train dispatching work. Different dispatchers have different decision preference when dispatching the trains. Some dispatchers prefer to ensure the high on schedule rate of all the trains, other prefer to reduce the total delay time of all the trains. We present a method to describe the fuzzy preference in dispatching work in this paper.

A boundary station refers to a station which is a demarcation point of two dispatching railway sections. In China, the dispatchers want to hand over as much as possible trains to the neighbor dispatching section before 18 o'clock every day. It is because of the economic interest of the related railway bureaus. We must find an equilibrium point for the railway bureaus to receive and send trains on the boundary stations.

When solving the engineering computing problem with the group dynamics theory, the reference vector is the base of decision. The reference vector is attained according to the opinion on the problem to be solved and various rules of the decision-makers. $\sum w_i = 1$, where w_i is the weight of the i th optimization objective in the opinion of a decision-maker.

According to the group dynamics, all of the individuals in a system are in a potential energy field.

Using the fluid mechanics and thermodynamics laws to describe the traffic flow reality, Nakrachi presented a new second order model, which represented different traffic flow phases. The conditions for transition between phases became clear [4]. Nakrachi et al. proposed a new expression of pressure and a second-order model,

which suggested solutions to a number of problems yet to be resolved [5]. Tao et al. took the actions in driving process as the interaction of the efficiency and safety [6]. Shahabpoora addressed group walking management problem by studying the dynamic properties of the combined human structure system [7]. We have studied the dynamic characteristics in recent years [8, 9]. We used an improved cellular automata model to analyze the train flow chaos by statistics on the distances between the neighbor trains and a Poincare section [9].

The status of an individual is influenced by the potential energy field and the potential energy field is also influenced by the individual. The rule can be mapped to the train dispatching work. The stations can be seen as intelligent individuals which can analyze the status of themselves and exchange information with each other. Then, they can select the optimal control strategy.

The potential energy can be defined with the number of the trains which will arrive at the station and the train arriving intervals. The high over high potential energy means overmuch trains and too small intervals between two trains' arrival and departure. Thus, the station cannot receive and send trains on time. On the contrary, small potential energy intends the small number of trains in a period of time, which is a waste of capability of the station. This paper tries to design train intervals based on the group dynamics model.

2 Model Construction

The model is to determine the number of the trains and the intervals between trains. The objective is to reduce the potential energy of the boundary station.

$$E = \min \sum_{i=1}^N \left\{ \rho \times n_i^2 + \mu \times C \times \sum_{\alpha=1}^{n_i^2} \left(\frac{1}{T_{\alpha,1,i}^{\text{arrival}}} + \frac{1}{T_{\alpha,1,i}^{\text{departure}}} \right) \right\} \quad (1)$$

The constraints are generated according to the capacity to receive and send trains of the station, the capacity of the railway section.

(1) Intervals between arrivals

$$\begin{cases} T_{\alpha,1,i}^{\text{arrival}} \geq I_i \\ T_{\alpha,1,i}^{\text{arrival}} \geq T_i^{\text{arrival}} \\ T_{\alpha,1,i}^{\text{arrival}} \geq T_i^{\text{wait}} \\ T_{\alpha,1,i}^{\text{arrival}} \leq T \\ T_{\alpha,1,i}^{\text{arrival}} \times (n_i - 1) \leq T \end{cases} \quad (2)$$

(2) Intervals between departures

$$\begin{cases} T_{\alpha,1,i}^{\text{departure}} \geq I_i \\ T_{\alpha,1,i}^{\text{departure}} \geq T_i^{\text{planned,departure}} \\ T_{\alpha,1,i}^{\text{departure}} \leq T \\ T_{\alpha,1,i}^{\text{departure}} \times (n_i - 1) \leq T \end{cases} \quad (3)$$

E	Potential energy of a boundary station
N	Number of the connecting directions of a boundary station
ρ	Fuzzy preference on number of trains
μ	Fuzzy preference on intervals between trains
α_i	Number of trains of the i th connecting direction, $i = 1, 2, \dots, N$
C	Tune parameter for interval between trains
T	A period of time in the study
I_i	Minimal interval between trains on the i th direction, $i = 1, 2, \dots, N$
t_i^{arrival}	The planned arrival interval between trains on the i th direction, $i = 1, 2, \dots, N$
$t_i^{\text{departure}}$	The planned departure interval between trains on the i th direction, $i = 1, 2, \dots, N$
t_i^{wait}	Waiting time of the trains of i th direction
n_i	The number of planned arriving trains of i th direction
T_i^{arrival}	The adjusted arrival interval between trains on the i th direction, $i = 1, 2, \dots, N$ (decision variable)
$T_i^{\text{departure}}$	The adjusted departure interval between trains on the i th direction, $i = 1, 2, \dots, N$ (decision variable)

In the upper level model, the intervals between the trains are the decision variables. Then, they are transmitted into the lower level model as the constraints. The objective of the lower level model is defined as follows.

$$T(g_{\alpha,\beta,i}^{\text{arrival}}, g_{\alpha,\beta,i}^{\text{departure}}) = \min \sum_{i=1}^N \sum_{\alpha=1}^a \sum_{\beta=1}^b \left\{ W_{\alpha,\beta} \left(g_{\alpha,\beta,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{planned,departure}} \right) + W_{\alpha,\beta} \left(g_{\alpha,\beta,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{planned,departure}} \right) \right\} \quad (4)$$

The constraints of the lower level model are listed as follows:

(1) Intervals between trains

$$\begin{cases} g_{\alpha+1,\beta,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{arrival}} \geq T_{\alpha,\beta,i}^{\text{arrival}} \\ g_{\alpha+1,\beta,i}^{\text{departure}} - g_{\alpha,\beta,i}^{\text{departure}} \geq T_{\alpha,\beta,i}^{\text{departure}} \end{cases} \quad (5)$$

(2) Operation time in sections

$$\begin{cases} g_{\alpha,\beta+1,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{departure}} \geq T_{\alpha,\beta \rightarrow \beta+1,i}^{\text{departure}} \\ g_{\alpha,\beta+1,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{departure}} \leq T_{\text{run}} \end{cases} \quad (6)$$

(3) Dwelling time at stations

$$g_{\alpha,\beta,i}^{\text{departure}} - g_{\alpha,\beta,i}^{\text{arrival}} \geq T_{\alpha,\beta,i}^{\text{wait}} \quad (7)$$

(4) Intervals between departures

$$g_{\alpha+1,\beta,i}^{\text{departure}} - g_{\alpha,\beta,i}^{\text{departure}} \geq t_{\alpha,\beta,i}^{\text{departure}} \quad (8)$$

(5) Intervals between arrivals

$$g_{\alpha+1,\beta,i}^{\text{arrival}} - g_{\alpha,\beta,i}^{\text{arrival}} \geq t_{\alpha,\beta,i}^{\text{arrival}} \quad (9)$$

(6) The earliest departure of the passenger trains

$$g_{\alpha,\beta,i}^{\text{departure}} \geq \text{Flag1} \times t_{\alpha,\beta,i}^* \quad (10)$$

(7) The number of the arrival and departure tracks

$$sl_{\beta} - N \left(g_{\alpha,\beta,i}^{\text{departure}} - g_{\alpha,\beta,i}^{\text{arrival}} \geq T_{\alpha,\beta,i}^{\text{wait}} \right) \geq 1 \quad (11)$$

3 ICA for Optimizing the Arrival and Departure Intervals

(1) Immune clone algorithm

Immune Clone Algorithm (ICA) is a new kind of IA, which hires the proportional replication operator and mutation operator. It is different from IA in that it generates some variant individuals to improve the local search ability and maintain the diversity of the individuals.

We can see that the ICA includes the following steps: particle swarm initialization, group antibody clone, antibody variation, and selection. The key step is group antibody clone, which decides the number of the antibody clone individuals. The steps will be continued until we get the sufficiently accurate results. Then, we will transfer the computing results to the engineering plan.

(2) ICA for optimizing the intervals between trains model

It is obvious that the essence of optimizing the potential energy of the boundary station is to optimize the interval between trains.

Step 1: Initialization

To decide the initial antigens (They determine the optimizing objective and the constraints). To generate the immune body with random number generation. Then, $T_{\alpha,1,i}^{\text{arrival},0}$ and $T_{\alpha,1,i}^{\text{departure},0}$ are generated. Then, a set $P_0(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$ is generated with M members.

$$g(T_{\alpha,1,i}^{\text{arrival}}, T_{\alpha,1,i}^{\text{departure}}) = \frac{1}{(1 + 1/\sqrt{n - 1})H} \tag{12}$$

Step 2: To calculate the affinity of the antibody individuals

For each $V_0(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0}) \in P_0(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$, to calculate the affinity $g(T_{\alpha,1,i}^{\text{arrival}}, T_{\alpha,1,i}^{\text{departure}})$. To generate the clone group $C(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$. The function of affinity is

$$g(T_{\alpha,1,i}^{\text{arrival}}, T_{\alpha,1,i}^{\text{departure}}) = \frac{1}{(1 + 1/\sqrt{n - 1})H}, \tag{13}$$

where H is a constant between 0 and 1

Step 3: To select n antibody individuals from the group $C(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$. The rule is to select the antibody individuals with highest affinity. Then, the antibody set $V(T_{\alpha,1,i}^{\text{arrival}'}, T_{\alpha,1,i}^{\text{departure}'})$ is generated.

Step 4: To clone the antibody with highest affinity

To clone the n antibody individuals which are selected at the previous step. Then to decide the optimal intervals through evaluating the number of the antibody individuals— $f(T_{\alpha,1,i}^{\text{arrival}'}, T_{\alpha,1,i}^{\text{departure}'})$. Thus, the scale of the antibody is enlarged. If $f(T_{\alpha,1,i}^{\text{arrival}'}, T_{\alpha,1,i}^{\text{departure}'}) > f(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$, then to replace $V_0(T_{\alpha,1,i}^{\text{arrival},0}, T_{\alpha,1,i}^{\text{departure},0})$ with $V(T_{\alpha,1,i}^{\text{arrival}'}, T_{\alpha,1,i}^{\text{departure}'})$.

Step 5: To mutate the antibody individuals with the single-point method to update the antibody group.

Step 6: To calculate the affinity of the antibody individuals. Set $n = n + 1$ and go to Step 4.

Step 7: To decide whether to stop the calculation loop. To judge whether the number of loop times reaches to M (M is settled to be the maximal loop times). If yes, stop the calculation, go to Step 8. If no, go to Step 2.

Step 8: To output the optimal solution $T_{\alpha,1,i}^{\text{arrival},*}$ and $T_{\alpha,1,i}^{\text{departure},*}$.

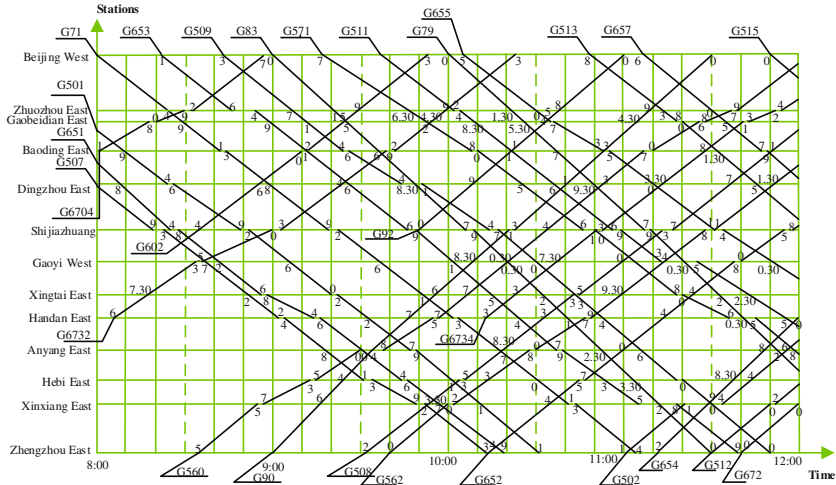


Fig. 1 Timetable from Beijing to Zhengzhou during 8 o'clock and 12 o'clock

4 Computing Case and Results Analysis

We took a section from a main railway line of China. There are ten stations and nine railway subsections in the dispatching section. We cut a part, from 8 to 12 o'clock from the daily train operation timetable, which is shown in Fig. 1. Then, we materialized the model proposed in Sect. 2 by setting the related parameters in the model, according to the planned timetable, assuming that some delays had occurred. Then, we solved the model with the ICA designed in Sect. 3.

Figure 1 is the planned timetable from Station 1 to Station 10 during the period from 8 to 12 o'clock. According to the method presented in this paper, we materialized the optimizing model and design the ICA to solve the model. The destination is to minimize the potential field energy. Set $\rho = 0.5$, $\mu = 0.5$, $C = 1000$, $n_i = 29 \times 2 = 58$, the computing result is that the optimal arrival interval is 12 min and the departure interval is 12 min.

5 Conclusion

This paper proposes a feasible, effective approach to solve the train operation interval optimization problems in railway transportation. We first present a model for optimizing the train operation interval based on the group dynamics theory. The object is to reduce the potential energy exerted on station by trains. An immune clone algorithm is designed to solve the model and the detailed steps to solve the model are given. The computing results are turned into optimal intervals between

trains when they arrive at a station from the same direction and departure from a station toward a same direction. Future study should be focused on the train rescheduling according to the optimized train operation intervals.

Acknowledgements This work is supported by the Natural Science Foundation of Gansu Province (Grant 1610RJZA047) and National Natural Science Foundation of China (Grant: 61563028).

References

1. Barreiro-Gomez J, Quijano N, Ocampo-Matinez C (2016) Constrained distributed optimization: a population dynamics approach. *Automatica* 69:101–116
2. Schwarz N, Flacke J, Sliuzas R (2016) Modelling the impacts of urban upgrading on population dynamics. *Environ Model Softw* 78:150–162
3. Zeferino JA, Antunes AP, Cunha MC (2014) Regional wastewater system planning under population dynamics uncertainty. *J Water Res Plan Manag* 140(3):322–331
4. Nakrachi A (2009) On traffic flow modelling with phase transition: an energy concept. In: *Proceedings of the 12th IFAC Symposium on Transportation Systems Redondo Beach, CA, USA, September 2–4*
5. Nakrachi A, Hayat S, Popescu D (2012) An energy concept for macroscopic traffic flow modelling. *Europ Trans Res Rev* 4(2):57–66
6. Tao P, Jin S, Wang D (2011) Car flowing model based on artificial potential field. *J Southeast Univ Nat Sci Edition* 41(4):854–858 (in Chinese)
7. Shahabpoor E, Pavic A, Racic V, Zivanovic S (2017) Effect of group walking traffic on dynamic properties of pedestrian structures. *J Sound Vib* 387:207–225
8. Meng X, Xiang W, Wang L (2014) Controllability of train service network. *Mathematical Problems in Engineering*, Article ID: 631492
9. Meng X, Xiang W, Jia L, Xu J (2015) Train flow chaos analysis based on an improved cellular automata model. *Chaos, Solitons Fractals* 81:43–51

Study on Energy Saving of Multi-vehicle Operation Based on Genetic Optimization Algorithm

Xuejin Wang, Xiangxiang Zhou, Yong Zhang and Zongyi Xing

Abstract Aiming at the multi-vehicle energy-saving problem of a metro train, this paper presents a research method of multi-vehicle operation energy saving based on genetic algorithm. First, the process of braking energy transfer in multi-train operation is analyzed. Second, taking the least energy consumption, and travel time as the targets, all-day trains, and the high/low peak traffic as the constraints, a multi-vehicle energy-saving model based on a multi-vehicle operation energy saving is established. Finally, the genetic algorithm is used to obtain the optimal stopping time and starting interval, and the total energy consumption, train energy consumption, and line loss are calculated. At the same time, the multi-vehicle energy-saving simulation is carried out by using the short-term of four sections of Rong Jingdong Street Station to Yizhuang Bridge Station of Beijing Yizhuang Line, and it also optimized the stopping time and the starting interval.

Keywords Urban rail transit · Regenerative braking energy · Genetic algorithm
Multi-vehicle energy saving

1 Introduction

Urban rail transit system is a complex multi-target and multi-vehicle dynamic system. There are frequent energy exchanges during the train operation process [1]. When the bicycle is running, the train only considers the situation of the interval operation, regardless of the influence of other vehicles and the impact of the parking time at each station [2], while the braking energy recycling is the most important part of train energy saving. Therefore, the energy-saving research on multi-train operation is of great practical significance to improve energy efficiency.

X. Wang · X. Zhou · Y. Zhang (✉) · Z. Xing
School of Automation, Nanjing University of Science and Technology,
Nanjing 210094, China
e-mail: 843922880@qq.com

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_54

Domestic and foreign experts have made a certain study on the train energy saving optimization. Liu [3] took Guangzhou Metro Line 1 as an example and analyzed the energy consumption of the train during the running of the positive line. The process of regenerative braking energy utilization was studied. It indicates that the traffic density is large and the efficiency of braking energy recycling is very impressive. Zhao [4] introduced the concept of passenger satisfaction. He studied the global energy consumption by changing the train departure interval. The optimal scheme was plotted as a run grab by him. Wang [5] established a single-line and multi-train scheduling model with energy consumption cost, carbon emissions, and train running time as the optimization target, who used the multi-objective fuzzy optimization algorithm to solve the optimization target. Zhao [6] studied the optimization principle of interstation run time based on the rich time allocation scheme, and took the total energy consumption as the objective function, the reasonable time to optimize the rich time, which optimized the wealthy time. The above research has made great contributions to the multi-vehicle energy-saving operation, but there are still some defects in versatility and convergent speed.

Aiming at the shortcomings of the above research, this paper presents a multi-vehicle energy-saving research method based on genetic optimization algorithm. This paper studies the principle of multi-train braking energy utilization, establishes the multi-vehicle multi-objective optimization model, and makes the multi-vehicle energy-saving simulation based on the four sections from Rongzhuang East Railway Station to Yizhuang Bridge Station in Yizhuang line. Verify the rationality and correctness of the model, and provide effective guidance for multi-train energy-saving operation.

2 Multi-vehicle Energy Utilization Principle

Multi-vehicle energy saving can be divided into the train braking energy recycling in the same direction and the train braking energy recycling [7] in the opposite direction. Traveling in the same direction, the two trains are far apart when they are driving in the same direction, which is generally in different power supply range. There will be a lot of line loss when braking energy is conveyed between long distance, resulting in braking energy cannot be effectively used. Therefore, this article only discusses the situation that the trains are in the opposite direction. There are two types of the situation. One situation is that the upstream train is braking, while the downstream train is pulled, which is shown in Fig. 1. Another situation is that the upstream train is pulled, while the downstream train is braking, which is shown in Fig. 2.

That one is the upstream train brake downstream traction, as shown in Fig. 1; one is the train traction down the train, as shown in Fig. 2. The distance between train braking energy utilization is closer, in which they are generally in the same power supply station, the rate of the braking energy utilization is higher.

Fig. 1 Up brake and down traction in the opposite direction

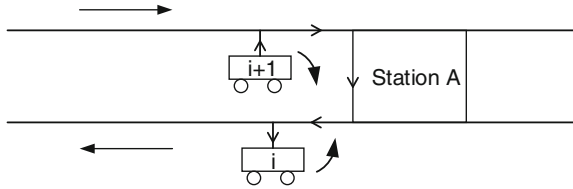
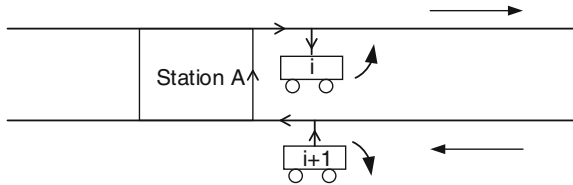


Fig. 2 Up traction and down brake in the opposite direction



3 Multi-vehicle Operation Constraints and Optimization Objective Function

3.1 Multi-vehicle Operation Constraints

Consider the various factors that may affect the operation of multiple-vehicle, we can select several important constraints such as departure interval, dwell time, travel time, and safety distance.

3.1.1 Departure Interval Constraint

Since the “speed–distance” curve is fixed for each train, the greatest impact on train operation is the departure interval, too large or too small departure interval can cause serious consequences. According to the characteristics of passenger flow, the day is mainly divided into morning and evening lower peak, the second peak, morning and evening peak, and normal peak. The specific situation is shown in Fig. 3. The departure time interval is defined as H_{trough} in the morning and evening lower peak during 6:00–7:00 and 22:00–24:00; The departure interval is defined as H_{normal} in the normal peak of daytime; The departure time interval is defined as H_{peak1} in the morning and evening peak during 7:00–9:00 and 17:00–19:00; The departure time interval is defined as H_{peak2} in the morning second peak during 9:00–10:00; Units are seconds.

According to the line of the operation of the plan, the smallest and the maximum departure time interval are defined as H_{min} and H_{max} ; Size relationships are:

$$t_{trace} < H_{min} < H_{peak1} < H_{peak2} < H_{normal} < H_{trough} < H_{max} \tag{1}$$

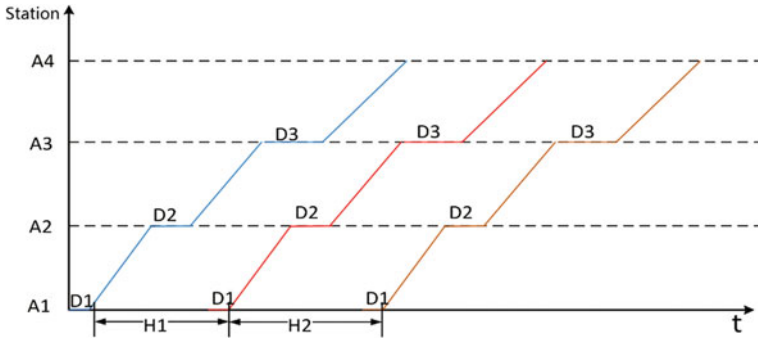


Fig. 3 Stop time and start interval diagram

In addition, the logarithm of train travel all day is NUM_{pair} , which is a fixed value. NUM_{pair} is:

$$\left\lceil \frac{T_{peak1}}{H_{peak1}} \right\rceil + \left\lceil \frac{T_{peak2}}{H_{peak2}} \right\rceil + \left\lceil \frac{T_{normal}}{H_{normal}} \right\rceil + \left\lceil \frac{T_{trough}}{H_{trough}} \right\rceil = NUM_{pair}, \tag{2}$$

where $\left\lceil \frac{T_{peak}}{H_{peak}} \right\rceil$ indicates that $\frac{T_{peak}}{H_{peak}}$ is rounded; T_{peak1} is the total running time in the morning and evening peak; T_{peak2} is the total running time in the second peak; T_{normal} is the total running time in the normal peak during daytime; T_{trough} is the total running time in the morning and evening lower peak.

3.1.2 Each Station Residence Time Constraint

Train stop time consists of three parts: the opening time, passengers on and off time, and closing time [8]. The longer the stop time, the stronger the traffic capacity of the station. Stop time is short, which will cause the travel time for passengers to their destinations become longer, and may cause the accident of pushing. Therefore, the stop time of each station is initially limited according to the average annual traffic, and the stop time range for each station is made, and the stop time is required to be an integral multiple of 5 s. Station number is 1, 2, 3, ..., i , ..., N . The corresponding stop time for each station pair is $D_1, D_2, D_3, \dots, D_i, \dots, D_N$; The lower limit of the stop time for the i station is D_{min}^i . The upper limit of the stop time for the i station is D_{max}^i . The stop time constraint for each station is:

$$D_{min}^i \leq D_i \leq D_{max}^i$$

$$D_i \% 5 = 0, \tag{3}$$

where % means to take the spare operation.

3.2 Multi-vehicle Optimization Objective Function

3.2.1 Multi-vehicle Running Energy Consumption Target

In the urban rail transit system, the train not only consumes electricity but also consumes energy [9] because of the presence of resistance in the catenary and rail. Especially in the line running peak, the line loss is also relatively larger due to the larger load and the larger supply current. Therefore, the goal of multi-vehicle running energy consumption should take into account the loss of the line and the energy consumption of the train. Take the total amount of energy issued by all tractions substation as the goal, and all the traction substations are numbered $P = [P_1, P_2, P_3, \dots, P_i, \dots, P_N]$. The voltage and current of the station P_i are U_{P_i} and I_{P_i} , and the corresponding energy consumption is

$$E_{\text{multi}} = \sum_1^N E_{P_i} \quad (4)$$

3.2.2 Multi-vehicle Travel Time Target

Multi-vehicle travel time refers to the sum of all the time from the originating station to the departure from the terminal. Excessive travel time indicates a possible delay in the line [10]. In day-to-day management, there is a little error between the allowing actual travel time and the specified operation time. The stop time of the train in the line is $D = [D_1, D_2, \dots, D_i, \dots, D_N]$, and the running time of the train in each interval is $t = [t_{1-2}, t_{2-3}, \dots, t_{i-(i+1)}, \dots, t_{(N-1)-N}]$. The travel time target of the train is:

$$T_{\text{travel}} = D + t = \sum_{i=1}^N (D_i + t_i) \quad (5)$$

3.3 Multi-objective Optimization Model for Multi-vehicle Operation

The goal of this paper is to establish an optimal timetable that needs to meet:

- (1) The peak and non-peak departure intervals are separately optimized, and the peak operating time is changed within ± 10 s, and the non-peak departure interval is changed within ± 30 s.
- (2) Change the stop time within ± 5 s.

- (3) One-way running time, all-day running time, and other configurations remain unchanged.
- (4) The optimization process also takes into account the optimization of the train running energy consumption and the optimization of regenerative braking energy, and calculate the equilibrium point of the two, in order to achieve the overall (substation) energy optimal.

According to the constraints and optimization objectives of multi-train operation, the following multi-train optimization model is established:

$$\begin{aligned} & \min \{E_{\text{multi}}, \Delta T\} \\ \text{s.t. : } & \begin{cases} D_{\min}^i \leq D_i \leq D_{\max}^i \\ \left[\frac{T_{\text{peak1}}}{H_{\text{peak1}}} \right] + \left[\frac{T_{\text{peak2}}}{H_{\text{peak2}}} \right] + \left[\frac{T_{\text{normal}}}{H_{\text{normal}}} \right] + \left[\frac{T_{\text{trough}}}{H_{\text{trough}}} \right] = NUM_{\text{pair}}, \\ H_{\min} < H_{\text{peak1}} < H_{\text{peak2}} < H_{\text{normal}} < H_{\text{trough}} < H_{\max} \\ D_i \in N^*, H_{\text{peak1}}, H_{\text{peak2}}, H_{\text{normal}}, H_{\text{trough}} \in N^* \end{cases} \end{aligned} \tag{6}$$

where $D = [D_1, D_2, D_3, \dots, D_i, \dots, D_N]$, N is the total number of stations. $\Delta T = |T_{\text{table}} - T_{\text{travel}}|$.

4 Model Solving Based on Genetic Algorithm

In the multi-train optimization model, it is necessary to establish a full-line operating schedule.

Timetable consists of two aspects, one is the single train in the station of the stop time, the other is between the train interval [11]. In this paper, the genetic algorithm is used to solve the multi-vehicle operation optimization. In this paper, genetic algorithm is used to solve the multi-vehicle operation optimization. The stop time and the departure interval are shown in Fig. 3.

Use the genetic algorithm to optimize the timetable as follows:

Encoding: Based on the number of stations in the actual line and the stop time series, use the genetic algorithm to optimize the timetable as follows:

- (1) Encoding: Based on the number of stations in the actual line, the stop time and the departure interval are integrated coding as $D = [D_1, D_2, D_3, \dots, D_i, \dots, D_N]$ and $H = [H_{\text{peak1}}, H_{\text{peak2}}, H_{\text{normal}}, H_{\text{trough}}]$. That is to define the genetic algorithm for the chromosome as $C = [D_1, D_2, D_3, \dots, D_i, \dots, D_N, H_{\text{peak1}}, H_{\text{peak2}}, H_{\text{normal}}, H_{\text{trough}}]$. According to the actual needs, stop time is mostly in 30–60 s. The departure interval H is in the peak period of about 120–300 s and in a low peak period of about 180–600 s. The length of the gene was determined, and 50 individuals were randomly selected to form a population.

- (2) The 50 groups of solutions in population are respectively introduced into the objective function to obtain the timetable of the whole day. Calculate the day-to-day energy consumption of the train, substation load data, and train travel time. The result of the operation is substituted into the fitness function to calculate the fitness value of each group.
- (3) According to the fitness value, the 50 groups of the populations were sorted and the genetic operation was carried out to obtain a new generation of 50 groups of new feasible solutions, where the ratio of crossover and mutation is 0.8 and 0.005, respectively.
- (4) Repeat Steps (2) and (3) until the feasible solutions that are calculated satisfy one of two following conditions:
 - ① The best fitness values of the two generations are within 0.001.
 - ② The total number of population does not exceed 50 generations.

In order to meet the need that the smaller the energy consumption, the smaller the travel time difference, the following fitness function is used to study the multi-vehicle genetic optimization algorithm:

- (5) and $H = [H_{\text{peak1}}, H_{\text{peak2}}, H_{\text{normal}}, H_{\text{trough}}]$ the departure interval of each station are integrated coding. That is to define the genetic algorithm for the chromosome as $C = [D_1, D_2, D_3, \dots, D_i, \dots, D_N, H_{\text{peak1}}, H_{\text{peak2}}, H_{\text{normal}}, H_{\text{trough}}]$. According to the actual needs, stop time is mostly in 30–60 s. The departure interval H is in the peak period of about 120–300 s and in a low peak period of about 180–600 s. The length of the gene was determined, and 50 individuals were randomly selected to form a population.
- (6) The 50 groups of solutions in population are respectively introduced into the objective function to obtain the timetable of the whole day. Calculate the day-to-day energy consumption of the train, substation load data, and train travel time. The result of the operation is substituted into the fitness function to calculate the fitness value of each group.
- (7) According to the fitness value, the 50 groups of the populations were sorted and the genetic operation was carried out to obtain a new generation of 50 groups of new feasible solutions, where the ratio of crossover and mutation is 0.8 and 0.005, respectively.
- (8) Repeat Steps (2) and (3) until the feasible solutions that are calculated satisfy one of two following conditions:
 - ① The best fitness values of the two generations are within 0.001.
 - ② The total number of population does not exceed 50 generations.

In order to meet the need that the smaller the energy consumption, the smaller the travel time difference, the following fitness function is used to study the multi-vehicle genetic optimization algorithm:

$$\text{fit} = \frac{A - E_{\text{multi}}}{B} + \frac{D}{\Delta t^2 + C}, \tag{7}$$

where, A, B, C, and D are constant coefficients.

5 Simulation Analysis

5.1 Simulation Background

In order to verify the accuracy of the multi-vehicle energy-saving calculation model [12], this paper takes “Rongjing Dongjie–Wanyuan Street–Yizhuang Culture Park–Yizhuang Bridge” to simulate the three intervals. Traction substation and stop time of each station from Yizhuang line Rongjing East Street to Yizhuang Bridge are shown in Table 1.

Yizhuang line is powered by the third rail power supply. Rated voltage level of the third power is 750 V. The third track unit resistivity is 0.125 Ω/km, while rail resistance takes 0.02 Ω/km. According to the bicycle running energy-saving simulation model, the distance curve and power curve of the train in three intervals are obtained as the input of multi-vehicle operation. Each station has a stop time adjustment of ±5 s or remains unchanged. The simulation run time is set to 2 h. The first hour of departure is between 420 and 450 s, and the second hour starts between 540 and 570.

5.2 Simulation Results

First, when the first-hour departure interval [13] is 450 and the second-hour departure interval is 550, the stop time is [40,30,35,30]. Only 1/4 of the overlap can achieve the effective use of brake energy pairs. In the use of multi-vehicle running energy-saving optimization, the front end of the optimal solution is that the train departure interval is adjusted to 463 s in the first hour, which is adjusted to 556 s in the second hour, and stop time is adjusted to [35,35,40,25]. In 2 h of simulation

Table 1 The substation location and stop time from Rongjing East Street to Wanyuan Street

Station	Rongjing East Street	Wanyuan Street	Yizhuang Cultural Park	Yizhuang Bridge
The substation location/km	0	1280	2818	3811
The original stop time/s	40	30	35	30

Table 2 The energy consumption of the interval before and after the optimization

Timetable sequence	[40,30,35,30,450,550]	[35,35,40,25,463,556]
Energy consumption/kWh	2441.35	2249.3
Line consumption/kWh	42.14	29.83
Total energy consumption/kWh	2483.49	2279.13

time, 3/4 can produce the effective use of braking energy. The energy consumption of the interval before and after the optimization is shown in Table 2.

As the simulation interval is shorter, less restricted, multicar energy-saving optimization simulation effect is obvious. Before and after the timetable optimization, the running time of 442 s remains unchanged, the energy consumption of the train decreases by about 7.87%, the line loss decreases by about 29.21%, and the energy consumption of the whole line decreases by 8.23%. Optimized, due to the current in the line does not fluctuate within a wide range, line loss decreased more, and the train can effectively use the braking energy, reducing the overall loss.

6 Conclusion

This paper presents a multi-vehicle energy-saving study based on genetic optimization algorithm, and solves the multi-objective optimization problem of multi-vehicle operation according to the principle of train braking energy utilization. Choose four sections from the Rongdong East Street Station to Yizhuang Bridge Station in Yizhuang line as a short line. Optimize the stop time and start interval through the multi-vehicle energy-saving simulation in this short line. The simulation results show that the optimization model and genetic algorithm can improve the energy consumption of the whole line by adjusting the stop time and the starting interval. The optimization model and the genetic algorithm in the multi-vehicle energy-saving calculation are reasonable and correct.

Acknowledgements This work is supported by Guangzhou Science and Technology Project (201604030061) and National Key R&D Program of China under Grant 2016YFB1200402.

References

1. Wang Y (2013) Using genetic algorithm to optimize the subway multi-interval velocity curve and stop time to realize traction energy saving simulation. Nanjing University of Science and Technology (in Chinese)
2. Li K (2014) Study on multi-train cooperative control energy-saving optimization method. Beijing Jiaotong University, Beijing (in Chinese)
3. Liu B (2007) Analysis of energy consumption of subway train. *Electr Locomot Urban Rail Veh* 4:65–68, 70 (in Chinese)

4. Zhao L (2014) Research on train timetable optimization model and algorithm based on regenerative braking. Beijing Jiaotong University, Beijing (in Chinese)
5. Wang D, Li K, Li X (2012) Economic dispatching model and fuzzy optimization algorithm for multi-target train. *Sci Technol Eng* 12:2869–2873 (in Chinese)
6. Zhao Y (2015) Study on energy-saving operation of multi-train of urban rail transit considering regenerative braking energy. Beijing Jiaotong University, Beijing (in Chinese)
7. Wu Y, Liu S (2004) Study on Train Marshalling Scheme Based on Energy saving and Passenger Service. *Urban Rail Transit Res* 06:27–31 (in Chinese)
8. Xue Y, Ma D, Wang L (2007) Calculation method for traction energy consumption of trains. *China Railway Sci* 03:84–87 (in Chinese)
9. Sidelnikov V (1965) Computation of optimal controls of a railroad locomotive. *Proc State Railw Res Inst* 2:52–58
10. Milroy IP (1980) Aspects of automatic train control[D]. © Ian Peter Milroy
11. Feng J (2014) Study on optimization of urban rail transit train operation behavior considering energy saving target. Beijing Jiaotong University, Beijing (in Chinese)
12. Moritani T, Kondo K (2010) Basic study on a designing method of the traction equipments to save the running energy with an optimization method. In: *Electrical Systems for Aircraft, Railway and Ship Propulsion*, IEEE, 2010, pp 1–6
13. Tian Z, Hillmansen S, Roberts C et al (2015) Energy evaluation of the power network of a DC railway system with regenerating trains. *IET Electr Syst Transp* 6(2):41–49

A Simulation Research for Fare Collection Equipments on Tiyuchang Station in Xi'an Subway

Guang Li and Chaosheng Wu

Abstract The Automatic Fare Collection (AFC) system is a kind of infrastructures for urban rail transit and is widely used in the subway system around the world. It is an important problem to determine the setting number of AFC equipments. While setting too little equipments, there will be congestion and may cause security problem. And setting too many will bring waste. This paper uses simulation method to study this problem for Tiyuchang station of Xi'an subway. It uses its own queuing simulation software for AFC equipments to analyze the passenger data of Tiyuchang station and gives suggestion based on simulation results.

Keywords Fare collection equipment · Urban mass transit · Subway Simulation

1 Introduction

Urban rail transit is a popular urban traffic form in big cities around the world. It has many advantages like large traffic volume, energy saving, and so on. The past decade is a high-speed development period for urban rail transit in China Mainland [1]. The subway has become popular in China.

The Automatic Fare Collection (AFC) system [2, 3] is a very important component for modern urban rail transit. An AFC system is an automated version of manual fare collection. It is usually the basis for integrated ticketing. There are two main parts of the AFC system, which are the AFC station equipment and the computer system. In Chinese subway system, there are mainly two kinds of AFC station equipments. They are the automatic ticket gate machines and automatic ticket vending machines. They support automation of the ticket work and can replace human operator. The ticket data will be collected by the AFC station

G. Li (✉) · C. Wu

School of Electronic and Control Engineering, Chang'an University,
Xi'an 710064, China
e-mail: hit6006@126.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_55

543

equipments in the normal working process. And these data will be transferred, saved, and analyzed by the computer system.

It is an important problem to decide the setting number of the AFC station equipments. Obviously, setting too many AFC equipment will cause equipment idling and waster. But while setting too few, the phenomenon of passengers waiting in the line will be serious. It will produce passenger congestion, increasing the time spending for travel, giving passengers a bad experience and bring great security risk.

The AFC station equipments' reasonable number is closely related with the station passenger traffic. The AFC station equipments form a queuing system. And computer simulation [4, 5] is always used to study it. Because every station has its own unique passenger traffic, it should study and decide the reasonable number of the AFC station equipments for every station, respectively.

This paper studies this problem for the Tiyuchang station of Xi'an subway. It collected the passenger data by field observation. And simulation is done by using its own software. According to the simulation result, it gives suggestion for the setting of the AFC station equipment in Tiyuchang station.

The computer simulation means a simulation, which runs on computers. The purpose of a computer simulation is to reproduce a complex system's behavior. It is a useful tool and can be used in many fields, like electronic engineering [6], astronomy [7], manufacturing system [8], and so on.

Queuing system is a typical complex system that always needs simulation. It can appear in many fields, like logistics [9], traffic [10], aviation [11], and so on. This paper uses queuing simulation to study AFC station equipment setting.

This paper organized its rest part as follows. The queuing simulation method and software this paper used are described in Sect. 2. Section 3 describes the data this paper used and the simulation result. Section 4 presents conclusions.

2 Queuing Simulation Method and Software

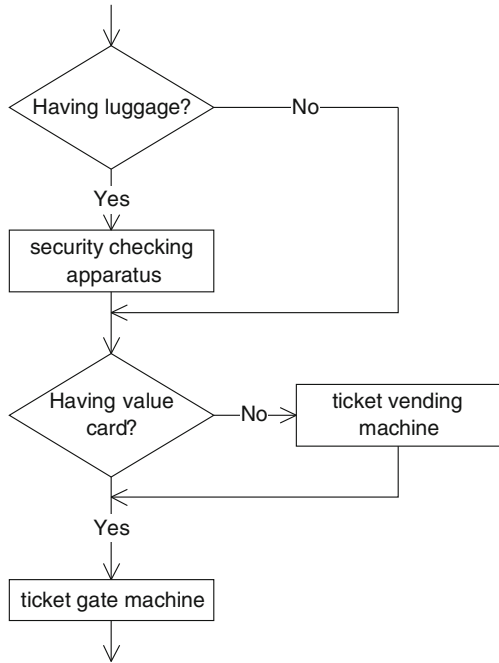
This paper designed and developed its own queuing simulation software and used it to study the AFC equipment setting problem of Tiyuchang station of Xi'an subway.

In most urban rail transit systems in China, including Xi'an subway, queuing can appear in both the outbound process and the inbound process.

In the outbound process, the passengers get off the train and go to ticket gate machine directly. The queuing in the outbound process may appear at the ticket gate machine only.

In outbound process, when the train stops at the station, a group of outbound passengers will appear. There will be no outbound passengers at other times. That means, in subway, the outbound passengers are sudden passenger flow. The busyness and idleness will appear alternately at the ticket gate machine.

Fig. 1 The Xi'an subway's inbound process



As shown in Fig. 1, the typical inbound process in China is much more complex. The passengers need to pass through the security checking apparatus, the ticket vending machines, and the ticket gate machines. At all these three kinds equipment, there may be queues. If the passengers do not carry luggage, they can skip over security checking apparatus. If the passengers have stored value card, they can skip over ticket vending machines.

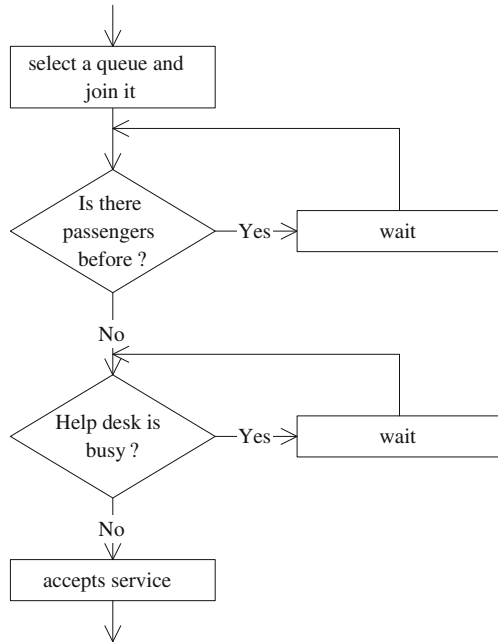
Unlike outbound process, passenger traffic is not sudden passenger flow in inbound process. Its change is more gentle and regular. In working days, the passengers are very few in the early morning and increase gradually. It will reach a peak when it is time to go school or work. The passengers will decrease after the early peak and will increase again in the afternoon. It will reach another peak when it is time to go home from school or work. After late peak, the passengers decrease again. It will have very few passengers in the late at night.

Our queuing simulation software concentrates on analyzing the inbound process. As shown before, the outbound process is simple and the outbound passenger traffic is typical burst passenger traffic. So, there is little valuable to analyze the outbound process.

There are mainly three parts in a queuing process. They are service law of the help desk, queuing discipline, and input process.

In this paper, the service law of the help desk means how much time the desk needs to serve a passenger. This paper observes the actual AFC operation to get it.

Fig. 2 Passenger behavior in queuing system



The queuing discipline means the criterion which the passengers must adhere to. In this paper, it is the first come first service (FCFS) and the waiting criterion. The waiting criterion means if a passenger comes, he will wait in his queue. He will not change to another queue to wait. He also will not go away no matter how much time he needs to wait. The FCFS criterion means the earlier a passenger coming, the earlier he will be served. The input process means the law the passengers coming. This paper gets it through field observations.

The behavior of passengers is shown in Fig. 2. Each service desk corresponds to a queue. When a new passenger is coming, he selects a shortest queue to join. Then, he checks if there are other passengers before him in this queue. If there are, he waits and checks again. If there are no such passengers, he checks if the help desk corresponding to him is busy. If the help desk is busy, he waits and checks again. If not, he goes to this desk and accepts its service.

This paper's software uses event-driven policies. The software defined some events. And it just waits until an event happened. Then, it does some processing according to the happening events. There are two kinds of events in this software. They are leaving event and arriving event. A leaving event is happen when a passenger finished service and leave the help desk. An arriving event is happen when a new passenger is coming and adding into the queuing system. Only when event happened, the queuing system will change accordingly.

This paper's software is programmed by R language [12], which is free and has a good performance.

3 Data and Simulation Results

This paper studies the AFC equipment setting problem for the Tiyuchang station of Xi'an subway. It collected data from field observation, and used its own queuing simulation software. According to simulation result, it gives suggestion to improve the current setting. As shown above, this paper concentrates on the inbound process.

Figure 3 shows the current setting of AFC equipment in Tiyuchang station. This station has four gateways: A, B, C, and D. On A–D side, there are seven ticket vending machines, one security checking apparatus, four inbound ticket gate machines, and six outbound ticket gate machines. On B–C side, there are six ticket vending machines, one security checking apparatus, three inbound ticket gate machines, and five outbound ticket gate machines.

Obviously, the number of the security checking apparatus is small. In fact, the security checking apparatus is insufficient. It is often serious queuing before the security checking apparatus when the traffic is slightly larger. But because the space is not enough, it is difficult to add new security checking apparatus. So, the ticket vending machines and the ticket gate machines are the research focus of this paper.

Around Tiyuchang station, there are mainly residential area and the Shaanxi provincial stadium. This stadium often holds large-scale activities, like concert or sports competition. Traffic generated from residential area has a relatively strong regularity. The stadium usually almost does not generate passenger flow, but will generate big huge passenger flow when there is large-scale activity.

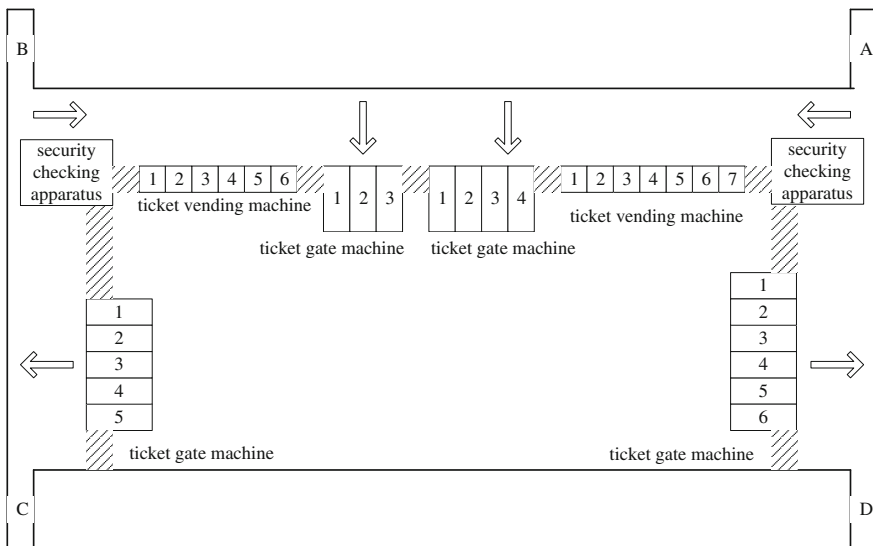


Fig. 3 The current setting of AFC equipments in the Tiyuchang station

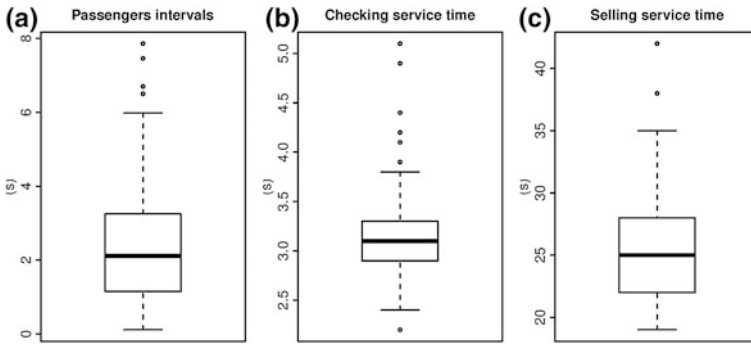


Fig. 4 The observation data. **a** Shows the intervals of passengers arriving. **b** Shows time spent by ticket gate machines for one passenger. **c** Shows time spent by ticket vending machines for one passenger. The data unit is seconds

After observation of several large-scale activities, this paper finds that existing equipment cannot respond the large sudden passenger flow well. The bottleneck is still the security checking apparatus. The queuing waiting before the security checking apparatus is very seriously when there is big huge passenger flow and often needs many policemen to maintain order. It is not suitable to add the number of security checking apparatus for solving this problem. There is not enough space for more security checking apparatus as shown above. And, in order to prevent waste, the equipments cannot setting to adapt the large sudden passenger flow, because it does not often appear. So, the simulation in this paper does not consider the sudden passenger flow.

Figure 4 shows the observation data. Three kinds of data have been collected. They are the intervals of passengers arriving and time spent for one passenger of ticket gate machines and ticket vending machines. The data is collected in morning peak of a working day in 2016. And there is no large event in the Shaanxi provincial stadium that morning. According to field observation, data of A–D side and B–C side has no obvious difference if there is no large event. So, this paper used data of B–C side only. About one-third of passengers do not carry luggage and do not need security checking. About two-third of the passengers have stored value cards and do not need ticket vending machines.

Tables 1 and 2 showed the simulation results for different number of ticket vending machines and ticket gate machines respectively. In this experiment, the

Table 1 The simulation result for ticket vending machine

Equipments' number	Equipment utilization rate	Average wait time (s)	Rate of passengers need waiting
3	0.96	111.04	0.95
4	0.92	48.61	0.90
5	0.85	21.05	0.79
6	0.58	3.01	0.26

Table 2 The simulation result for ticket gate machine

Equipments' number	Equipment utilization rate	Average wait time (s)	Rate of passengers need waiting
2	0.98	51.12	0.98
3	0.85	1.55	0.59
4	0.65	0.13	0.10

number of equipments not stated in tables is consistent with the current actual situation.

According to the simulation result, it can be seen that the appropriate number of the ticket vending machine and the ticket gate machine is about 5 and 3, respectively. The existing setting number is enough. And in fact, it slightly sets too many AFC equipments.

The reason of setting so many equipments should be considering the sudden passenger flow from the Shaanxi provincial stadium. It can be seen that the A–D side has more AFC equipments than the B–C side as shown in Fig. 3. This paper thinks it has the same reason because the A–D side is closer to the stadium. But as shown before, the bottleneck is the security checking apparatus and it is difficult to add because the space is not enough. So, this paper thinks there is no need to set a little many AFC equipments if the security checking apparatus cannot be added. If there is no plan to add the security checking apparatus, this paper suggests removing one and two ticket vending machines for B–C side and A–D side, respectively, and removing one inbound ticket gate machines for A–D side.

4 Conclusion

In the modern subway system, the AFC station equipments are popular. It is an important problem to decide the setting number of them, because setting too much will cause wasting and too few will cause crowding. It is obvious that the passenger traffic decided the AFC equipments' reasonable number mainly. The passenger traffic is different from one station to another. So for each station, it is necessary to study this problem, respectively.

This paper uses queuing simulation method to study this problem for Tiyuchang station in the Xi'an subway. It used its own simulation software and collected data by field observation. According to the simulation result, it gives the suggestion of the AFC equipment setting for Tiyuchang station of Xi'an subway.

Acknowledgements This paper is funded by Natural Science Basic Research Plan in Shaanxi Province of China (2016JQ6078), the Basic Research Funds of Chang'an University (0009-2014G6114024) and the National Nature Science Foundation of China (61601059, 51407012).

References

1. Gu M (2011) Status of urban rail transit development and prospects. *Chin Railways* 2011 (10):53–56 (in Chinese)
2. Zhu W, Hao H, Huang Z (2014) Calibrating rail transit assignment models with genetic algorithm and automated fare collection data. *Comput Aided Civ Infrastruct Engineering* 29(7):518–530
3. Jiang Y, Luo N (2013) Urban rail transit station automatic fare collection configuration research and simulation. In: *Proceedings of the ICTE 2013*. pp, 1142–1149
4. Law A (2014) *Simulation modeling and analysis*, 5th edn. McGraw-Hill Education, New York City, USA
5. Ross SM (2012) *Simulation*, 5th edn. Elsevier Academic Press, Waltham, Massachusetts, USA
6. Marlier C, Videt A, Idir N (2015) NIF-Based Frequency-domain modeling method of three-wire shielded energy cables for EMC simulation. *IEEE Trans Electromagn Compat* 57 (1):145–155
7. Vogelsberger M, Genel S, Springel V, Torrey P, Sijacki D, Xu D, Snyder G, Bird S, Nelson D, Hernquist L (2014) Properties of galaxies reproduced by a hydrodynamic simulation. *Nature* 509:177–182
8. Negahban A, Smith JS (2014) Simulation for manufacturing system design and operation: literature review and analysis. *J Manufact Syst* 33(2):241–261
9. Ma X, Qin X, Cheng C (2012) Simulation optimization based on production of queuing theory of logistics system. *Comput Eng Appl* 48(20):244–248 (in Chinese)
10. Zhang W, Li D, Luo W (2013) Simulation for speed spatial distribution and queuing characteristics on traffic accident section of freeway. *J Syst Simul* 1:158–163 (in Chinese)
11. Zhang J, Minghua H, Zhenya W (2013) Airspace capacity evaluation of terminal area based on flight queuing simulation. *Aeronaut Comput Tech* 43(4):9–13 (in Chinese)
12. Kabacoff R (2015) *R in action: data analysis and graphics with R*, 2nd edn. Manning Publications, Greenwich, Connecticut, USA

An Integer Programming Model of Reinvestment Strategy Based Project Portfolio Selection and Scheduling with Constrained Resource

Baolong Wang

Abstract Selecting an appropriate series of projects among numerous candidates and assigning start and finish time for every selected project under the constraints of limited resource does matter in the project management. With the introduction of annual budget and time value of capital, the proposed problem becomes more complex. In this paper, an integer programming model is proposed and constructed to model the situation of selecting and scheduling suitable project portfolio at the same time with constrained resource, in which reinvestment strategy is also applied in the process. In reality, annual budget is time-dependent, so it is necessary to take time value of capital into account during the modeling process. The use of reinvestment strategy means that the profit earned by accomplished projects is put into the resource pool to finance more projects to be selected and executed in order to make the total profit of project portfolio maximal. During the modeling progress, others constraints are also considered, like project interrelationship, to represent closely practical situations.

Keywords Integer programming · Project interrelationship · Project portfolio Selection and scheduling · Reinvestment strategy

1 Introduction

Since the introduction of project management from the foreign, it has become an important tool in the daily operation of organization or enterprise. Its wide application also reveals its huge value in the practice. Within the enterprises, whether the daily work or the production operation are both regarded as a whole project to deal with, like business operation. At first, the project management for single project does make a big difference in the enterprises and organizations over decades.

B. Wang (✉)

China Three Gorges Corporation (CTG), No. 335 Shuangyong North Road,
Nyingchi, Tibet, People's Republic of China
e-mail: wangbaolong13@mails.ucas.ac.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_56

551

However, as the global market becomes more and more complex and the competition more and more drastic, the tool of single project management shows its weakness in coping with multiple projects. So the multi-project management begins to attract the interest of practitioners and scholars. Usually, during a given time horizon does exist many alternative projects, called available projects or investment opportunities, to be considered and selected for an organization or a company. But the resource of an organization usually is not infinite, so it is not possible to choose all ones to invest simultaneously, which means that it is necessary to select a right series of projects, also called project portfolio, to execute to realize the objects of organization strategy. Considering the fact that 26% of information technology projects are finished punctually meeting the given budget [1], this makes big effect on the realization of organization strategy in the long term. All in all, the research and application of portfolio management theory will lead the field of project management to a more wide application, and has a potential space in the operation of the organizations and other institutions.

2 Literature Review

Project portfolio is a series of projects which are collected as a fixed group to assure the realization of organizations' strategy objectives, and these projects consume the same resource from a certain organization during a given time period, among which many kinds of constraints may exist in the proposed situation. Selecting a suitable project portfolio and separately assigning a starting and finishing time for every chosen project is another form of resource-constrained project scheduling problems (RCPS), which has wide application in many kinds of situations, like project management and financial management. Project portfolio selection and scheduling problem (PPSSP) is the kind of organization optimization problem in which several projects are composed into an optimal portfolio to meet the organization strategy by determining appropriate project elements and reasonably assigning every project element with a starting and finishing time within the planned time horizon. The element of chosen project portfolio may differ in many aspects, like the total consumed resource, the time limit, and even the relationship between them, and its like. Mostly, it is the best to choose all or most of the alternative projects and execute them simultaneously to assure the maximal produced profits, but the amount of resource within an organization is usually limited or constrained, it is necessary to choose the optimal portfolio and determine the start–finish time for every selected one. So the return or profit produced by completed projects can be made good use of and can be reinvested to finance other ones and alleviate the load of the organization to a certain extent. Consequently, the execution rank and the use of reinvestment strategy matter in the organizing process of project management.

The concept and theory of portfolio occurred in the field of financial investment and were used to evaluate the earn and the corresponding risk for every potential asset, being the base of modern portfolio [2]. Further researches were conducted by

McFarlan, which combined the concepts of portfolio and project management by considering individual and overall risk to make a better choice, laying the foundation of modern portfolio selection system [3]. As another form of RCPSP, the research of PPSSP has attracted many interests from scholars and practitioners for decades. Subsequently, a great progress is made about the development of project portfolio theory; and many extensions have enriched the original project portfolio theory and practice, like deterministic and static circumstances [4]. In addition to the traditional methods of evaluating and selecting projects, like AHP and PERT, more mathematical programming ways were developed to cope with more complex optimization models because of other kinds of constraint. Based on the constructed PPSSP and its variants, some resolution methods have been developed to solve the optimization models. Lean and Six Sigma method is constructed to select suitable projects [5]. As the problems become more complex, the flexible heuristic methods, like priority rule, have gradually taken place of mathematical programming, like zero-one programming [6, 7]. A branch and bound algorithm is also proposed to solve RCPSP [8].

The proposed PPSSP is analyzed under the circumstance of limited resource, like annual budget, by making full use of constrained conditions in order to realize the proposed object. During the model formulation, the strategy of reinvestment is also taken into account. So which projects are carried out first will make a big difference on optimal project portfolio selection and scheduling, as well as the final proposed objectives. To represent the practical situations, the capital time value, relationships among projects are also incorporated into the optimization model. A deadline is given to assure that all selected projects should be selected completed within the certain periods.

3 Problem Description and Modeling

During a certain time horizon, there are numerous and available projects with comparable property for an organization to make decision of investment, that is, selecting a suitable project portfolio and determining the detailed starting–finishing time for each element of the portfolio to ensure the realization of organization strategy. Many kinds of resource can be transformed to capital. Necessary capital may reach by one time or multiple periods during a planned time intervals, like the 3-year plan of an organization or company. Here, it is assumed that the budget reaches by many times to closely reflect the practice. And, usually the budget cannot finance all selected project to execute at the same time. In other words, the project element of proposed portfolio has to be launched or started in certain order to meet the given conditions. So the proposed PPSSP under certain above considerations comes to be a problem which has to be tackled with by practitioners and scholars. In addition, when a project is completed, it would generate fixed earning within a certain period of time, all of which may be constant or various, but it is usually a fixed value. The start-up fund is generally supported by first-year budget. The budget needed by following projects is composed of the balance of previous

input, newly reaching budget and the earn from completed ones. Except for the kind of resource which can be transformed to be money, there is the category resource which can just be provided with a certain every periods, called renewable resource. Within a certain period of time, it is supplied by a fixed amount every period interval. With the purpose of simplification and brief, the fixed value of resource amount is assumed to be constant; the budget provision necessary for a whole project is put into the project by once; once a project is started, it cannot be ceased until the whole project is completed; if a project is selected, it will be commenced at the beginning time, not the middle or the end; we suppose that the profit from previous completed project is constant and produced at the end time of every period within the period of profit, and is produced a year later when the project is done.

On the basis of above description, this article proposes a mathematical optimization model to realize the optimal total profit of the given problems under consideration.

Parameter introduction:

- N the number of available projects, $i = 1, 2 \dots N$;
- T_1 the given time period of accomplishing project, meaning that all selected projects must be accomplished within T_1 ;
- T_2 the time periods of budget reaching;
- M_j the budget during period j , occurring at the moment $j - 1$ ($j = 1, 2 \dots T_2$);
- H_i the return period of project i ;
- a the rate of discount;
- P_{ik} the annual earn of completed project i within period k under immediate launch of project i ;
- Q_i investment cost of project i if it starts immediately at the initial moment;
- D_i the implement duration of project i ;
- S_t the series of started projects before or on period t , $S_t = \{i \mid 0 < Y_i \leq t\}$;
- F_t the series of accomplished projects before the moment t , $F_t = \{i \mid Y_i - 1 + D_i \leq t\}$.
- X_i a binary decision variable,

$$\text{where } X_i = \begin{cases} 1 & \text{if project } i \text{ is included in the portfolio} \\ 0 & \text{if project } i \text{ is excluded in the portfolio} \end{cases};$$
- Y_i an integer decision variable, representing the starting time of project i ,

$$\text{where } Y_i = \begin{cases} \text{PI} & \text{if project } i \text{ is selected in the portfolio} \\ 0 & \text{if project } i \text{ is excluded in the portfolio} \end{cases}$$
- PI is an integer more than 0;

Before the modeling, it is necessary to define some relationships among alternative projects to represent closely the realistic conditions: partial exclusivity (E) and precedence constraint (P). Accordingly, E denotes that at most one of each pair with partial exclusivity relationship can be selected; P means that one project is completed, the other, belonging to the pair with precedence constraint, can start.

An integer programming model for PPSSP under considerations is formulated as follows:

$$\text{Maximize } Z = \sum_{i=1}^N \left(X_i \left(\sum_{k=D_i+1}^{D_i+H_i} P_{ik} e^{-ak} - Q_i \right) \right) \tag{1}$$

Subject to

$$\sum_{i \in S_t} \left(X_i Q_i e^{a(Y_i-1)} \right) \leq \sum_{j=1}^{\min\{T_2, t\}} M_j + \sum_{i \in F_{t-1}} \sum_{k=Y_i+D_i}^t X_i P_{ik} e^{a(Y_i-1)} \quad t = 1, 2, \dots, T_1 \tag{2}$$

$$X_i + X_j \leq 1 \quad (i, j) \in E \tag{3}$$

$$Y_i + D_i \leq Y_j \quad (i, j) \in P \tag{4}$$

$$Y_i - 1 + D_i \leq T_1 \quad i = 1, 2, \dots, N \tag{5}$$

$$Y_i = 1, 2, \dots, T_1 \quad i = 1, 2, \dots, N \tag{6}$$

$$X_i = 0, 1 \quad i = 1, 2, \dots, N \tag{7}$$

Equation (1) is defined as the objective function to maximize total profit of optimal project portfolio. Equation (2) denotes that the required investment must meet the provision of available resource, consisting of reaching budget and return generated by completed projects. Equation (3) presents the set of partial exclusivity relationship between each pair of projects. Equation (4) presents the set of precedence constraint between each pair of projects. Equation (5) means that all selected projects need to be finished within given time period. Equations (6) and (7) show decision variable Y_i is an integer and project i , if selected, must start before T_1 , and X_i is a binary variable for project selection.

Above, an integer programming model is constructed according the proposed situation. Considering all the decision variables are Boolean and Integer, the problem is not very complex, so it can be optimally solved by standard optimization technology. Many commercial softwares for integer programming technique can solve the proposed problems to obtain global optimal solution and ensure its precision.

4 Case Analysis

Within the chapter, a simple case of project selection and scheduling is taken out to illustrate the detailed application of above model. Besides, new product R&D can also be applied as another case of this paper.

To simply operation progress, we assume that annual profit of a project is the same after its completion, i.e., $P_{ik} = P_i$. The project attributes consist of project cost, duration, annual profit, and profit duration, which are shown in Table 1.

Table 1 Sample information of project attributes

Project name i	Project cost(\$) Q_i	Project duration (y) D_i	Annual profit (\$) P_i	Profit duration (y) H_i
P1	9 M ^a	3	5 M	4
P2	8 M	4	4 M	5
P3	6 M	2	3 M	6
P4	3 M	1	2 M	4
P5	3 M	1	2 M	3
P6	4 M	2	2 M	6
P7	4 M	2	1 M	8
P8	8 M	2	3 M	7
P9	9 M	3	5 M	4
P10	2 M	1	1 M	6

^aThe symbol of M means a million dollars

Table 2 Annual budget Limit (*unit* Million Dollars)

Year	First	Second	Third	Fourth
Annual budget	8 M ^a	9 M	8 M	9 M

^aThe symbol of M means a million dollars

Let the discount rate, $a = 10\%$. And the given deadline, $T_1 = 5$, is set to make sure selected projects completed within the planned time horizon.

Annual budget reaches by multiple periods are shown in Table 2, accordingly, $T_2 = 4$. Besides, interrelationship among projects is also considered, including partial exclusivity and precedence constraint, where P5 and P6 belong to the set of partial exclusivity relationship, and P5 can start if P4 has been completed.

For simplicity, the result of optimal selection and schedule under reinvestment strategy is called as Scenario I. Scenario II is defined as the same to Scenario I without reinvestment strategy. The difference of optimal selection and schedule under two scenarios can be analyzed in detail, such as optimal project portfolio, total profit, the element of optimal portfolio and its relative starting and finishing time, and the change trend of cumulative profit

Here, standard optimization technology is applied to solve the mathematical models optimally. The optimal selection and scheduling under Scenario I are shown in Fig. 1, where P1, P3, P4, P5, P7, P8, and P10 are included in the portfolio. Correspondingly, P4 commences at period $t = 1$; P1 and P10 at $t = 2$; P3 and P7 at $t = 3$; P8 at $t = 4$; P5 at $t = 5$ separately. For Scenario II, the optimal resolution is shown in Fig. 2, where P3, P4, P5, P7, P8, and P10 are included in the final project portfolio, and P4 and P10 commence at $t = 1$; P8 at $t = 2$; and P3 and P5 at $t = 3$; P7 at $t = 4$. More other information, like the starting–finishing time, can also be obtained. By contrasting Figs. 1 and 2, the conclusion is that reinvestment strategy not only makes a big difference on optimal portfolio selection but it also affects portfolio schedule. For example, P1 is included into the optimal portfolio under

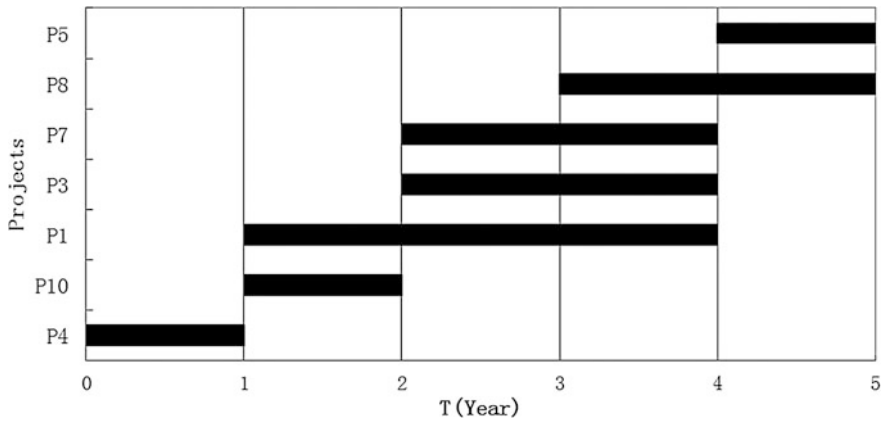


Fig. 1 Optimal selection and schedule for Scenario I

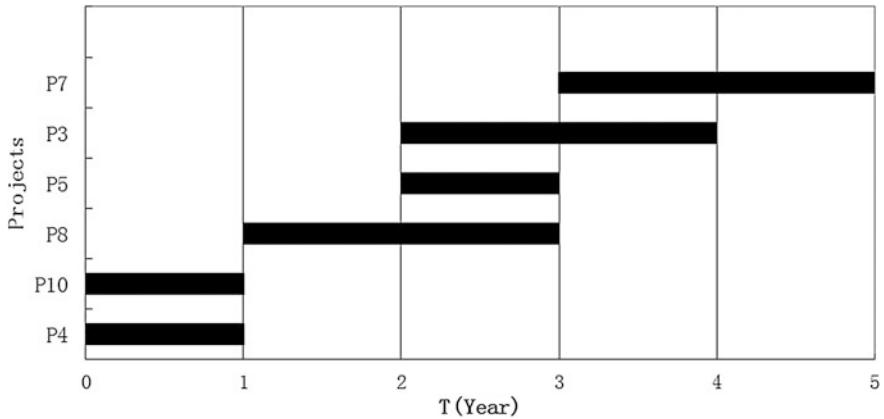


Fig. 2 Optimal selection and schedule for Scenario II

Scenario I and excluded in Scenario II. As shown in Fig. 3, the total return is \$18,429,000 for Scenario I and \$15,521,000 for Scenario II. The use of reinvestment strategy increases the total profit by 18.75%. The increment of revenue can be considered as the profit of additional project P1 by the use of reinvestment strategy. As shown in Fig. 3, the implement of reinvestment strategy provides more chance for projects to be selected and executed in order to maximize the final profit under the circumstance of same budget. In summary, the introduction of reinvestment strategy is reasonable and profitable for the selection and scheduling of optimal project portfolio under the constraint of limited resource. It not only finances more projects to be executed but improves the efficiency of capital remarkably.

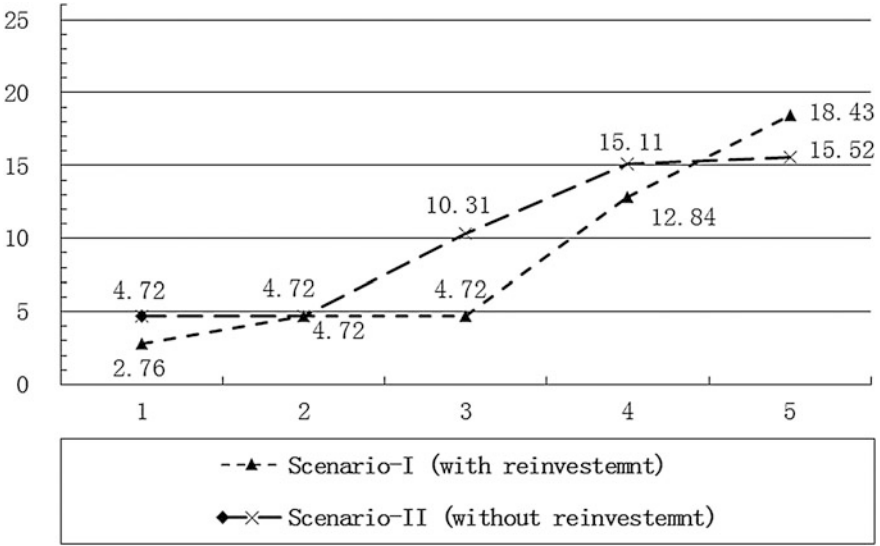


Fig. 3 Cumulative profit under both scenarios

5 Conclusion

An integer programming optimization model of reinvestment strategy-based project portfolio selection and scheduling with constrained resource is constructed to resolve optimally combinatorial optimization problems in the field of project management in this paper. Under the considered circumstance, annual budget reaches by multiple periods, that is, time-dependent; capital time value and relationships among project elements are also considered in order to closely reflect practical situations. The fund, a special category of resource, is consumed at the starting of a project and generated at the finishing time of the project. The special property of the fund provides more possibility to be reused and reinvested into other projects, so it finance more projects to be selected and incorporated into the final project portfolio. A simple case with two different scenarios confirms the application of proposed optimization model. In the paper, the optimal results of with and without reinvestment strategy are made an analytic comparison to demonstrate the effect of reinvestment strategy, which relieve the pressure of constrained resource, fund other projects, and then make overall profit of optimal project portfolio increase, enriching the research of theory and practice in project management field.

References

1. Chen J, Askin RG (2009) Project selection, scheduling and resource allocation with time dependent returns. *Eur J Oper Res* 193(1):23–34
2. Jafarzadeh M, Tareghian HR, Rahbarnia F et al (2015) Optimal selection of project portfolios using reinvestment strategy within a flexible time horizon. *Eur J Oper Res* 243(2):658–664
3. Markowitz HM (1952) Portfolio selection. *J. Finance* 7(1):77–91
4. Mcfarlan FW (1981) Portfolio approach to information-systems. *Harvard Bus Rev* 59(5): 142–150
5. Hu G, Wang L, Fetch S et al (2008) A multi-objective model for project portfolio selection to implement lean and Six Sigma concepts. *Int J Prod Res* 46(23):6611–6625
6. Kolisch R (1996) Efficient priority rules for the resource-constrained project scheduling problem. *J Oper Manag* 14(3):179–192
7. Tseng LY, Chen SC (2006) A hybrid metaheuristic for the resource-constrained project scheduling problem. *Eur J Oper Res* 175(2):707–721
8. Brucker P, Knust S, Schoo A et al (2000) A branch and bound algorithm for the resource-constrained project scheduling problem. *Math Method Oper Res* 52(3):413–439

Track Assignment Adjustment Problem in Complex Railway Passenger Stations

Yinggui Zhang, Min An and Li Wang

Abstract Track assignment plans are often adjusted for the purpose of safety and the minimum delay in real-time conditions, especially in busy complex railway passenger stations. First, taken trains as nodes, tracks as colors, a graph coloring model is designed with the objective function of the minimum total weighting delay. Then, a heuristics algorithm based on the two-stage solution method is put forward. Finally, a simple case is used to demonstrate the effectiveness and high efficiency of the proposed methodology based on graph coloring theory about how to solve the track assignment adjustment problem.

Keyword Track assignment · Adjustment problem · Railway passenger stations
Graph coloring model · Heuristic algorithm

1 Introduction

In recent years, the rapid development of high-speed railway network and its integration with existing passenger railway networks enhance the intricacy of track assignment problem in busy complex passenger stations, which has sustained high density in arrivals and departures, complexity in station layout and variety in train types. In real-time conditions, railway dispatchers sometimes have to adjust the given track assignment plans made in advance immediately due to various reasons, i.e., train early or delay, train accidents, which may further impact the given timetable of the whole railway network. However, track assignment adjustment

Y. Zhang (✉)

School of Traffic and Transportation Engineering, Central South University,
No. 22 Shaoshan South Road, Changsha, Hunan, China
e-mail: ygzhang@csu.edu.cn

M. An · L. Wang

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

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_57

problem is to make the reasonable track arrangements for all train passing through these complex railway passenger stations safely in real-time conditions.

An integrated approach is generally hopeless for real-world instances due to the complexity of train timetabling problem and train platforming problem [1], also namely track assignment problem, and track assignment adjustment problem studied in this paper can be categorized into the real-time level [2]. Many scholars have concentrated on studying track assignment problem with various objectives (i.e., delay) and constraints in different situations [3–5]. Also, Lusby and Samà et al. studied the real-time junction train routing and scheduling problems [6, 7]. It should be noted that less computation time for making reasonable track assignment adjustment plans is crucial for railway passenger transport safety organization. Since the reassignment of tracks to trains is time-critical in real-time level, heuristic approaches can be permitted to run in limited iterations in order to meet the time requirement [8]. In this paper, we will design a graph coloring model, in which trains are regarded as nodes and tracks are viewed as colors for track assignment adjustment problem in complex railway passenger stations. Also, we will design a heuristic algorithm to make reasonable track assignment adjustment plans within limited computation time to guarantee trains running on schedule.

The remainder of the paper is organized as follows. Taken the minimum total weighting delay as the objective function, a graph coloring model is built in Sect. 2. Section 3 designs a heuristics algorithm based on the two-stage solution method. A simple case is discussed in Sect. 4, and the paper concludes with a summary and discussion of some extensions for further research in Sect. 5.

2 Graph Coloring Model

The core task of track assignment adjustment problem is to find a feasible and safe track for each train with the constraint of mapping relationship between tracks and trains in the real-time situation, which can be gained from rail operations rules and regulations in railway passenger stations. Let $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{C})$ express an undirected graph for the track assignment adjustment problem. Nodes are also trains passing through a passenger station in certain period, $\mathbf{V} = \{v_1, v_2, \dots, v_i, \dots, v_j, \dots, v_n\}$ (n is the total number of trains). These nodes can be covered by various different colors (tracks in the real world), and thus $\mathbf{C} = \{c_1, c_2, \dots, c_k, \dots, c_m\}$ (m is the total number of tracks in the station). Also, there is an optional coloring set \mathbf{C}_i to describe the mapping relationship between tracks and trains for each train. For any pair of trains v_i and v_j , if there is a time conflict phenomenon, calculated by arrival or departure time in given timetable and station operations requirements, an edge e_{ij} exists between them, which means that these trains cannot be reassigned the same potential possible tracks at the same time and two conflicting routes during the same period for safety reasons. Routes releasing and occupying safely are crucial for

transport safety. There may be several routes connecting with each track in the station interlocking system. Thus, for adjacent nodes, there is a color taboo pairs set, each pair of which cannot be colored on these adjacent nodes at the same time due to potential trains and routes conflicts. In this way, the track assignment adjustment problem can be turned into a graph coloring problem.

Next, we will discuss about how to design edges and color taboo pairs sets. Let x_i and y_i denote the scheduled arrival and departure time respectively from the original timetable, x'_i and y'_i denote the actual arrival and departure time, respectively. Considering station operations safety requirements, there are standard service times in the rail arrival–departure layout for originating passenger trains and terminating passenger trains, which can be expressed as τ_{sf} and τ_{zd} . Besides, let τ_{jc} , τ_{fc} , τ_{tg} , τ_{crk} and τ_{xh} denote the general arrival routes occupied time, the general departure routes occupied time, the nonstop passing through time, the general routes occupied time between the arrival–departure layout, and passenger technical station or other layouts, the general signaling time opening in advance (min).

According to different train services and technical operational processes, passenger trains can be classified into originating passenger trains (ORPTs), terminating passenger trains (TEPTs), turnaround passenger trains (TUPTs), and through passenger trains (THPTs). Specifically, THPTs can be classified into nonstop through passenger trains (NSTHPTs) and stop through passenger trains (STHPTs) according to whether or not they will stop at the station. The initial actual arrival and departure time of the train v_i can be calculated by Eqs. (1) and (2).

$$x'_i = \begin{cases} y_i - \tau_{sf} & \text{for ORPT} \\ x_i & \text{for others} \end{cases} \tag{1}$$

$$y'_i = \begin{cases} x_i + \tau_{zd} & \text{for TEPT} \\ x_i & \text{for NSTHPT} \\ y_i & \text{for others} \end{cases} \tag{2}$$

For any two trains or nodes v_i and v_j , if $[x'_i - \sigma, y'_i + \sigma] \cap [x'_j - \sigma, y'_j + \sigma] \neq \emptyset$ (namely as tracks conflicts pre-judgement formula), the edge between nodes v_i and v_j exists (σ is a safe time allowance, minutes).

In addition, we can calculate much more precisely about the starting occupied time s_i and the releasing time r_i of tracks which are reassigned by trains, which can be rewritten as Eqs. (3) and (4), respectively.

$$s_i = \begin{cases} y_i - \tau_{sf} - \tau_{crk} - \tau_{xh} & \text{for ORPT} \\ x'_i - \tau_{jc} - \tau_{xh} & \text{for TEPT} \\ x'_i - \tau_{jc} - \tau_{xh} & \text{for TUPT and STHPT} \\ x'_i - \tau_{xh} & \text{for NSTHPT} \end{cases} \tag{3}$$

$$r_i = \begin{cases} y'_i + \tau_{jc} & \text{for ORPT} \\ x_i + \tau_{zd} + \tau_{crk} & \text{for TEPT} \\ y'_i + \tau_{jc} & \text{for TUPT and STHPT} \\ x'_i + \tau_{tg} & \text{for NSTHPT} \end{cases} \quad (4)$$

When the train v_i and v_j arrive at or depart from the track c_k and $c_{k'}$ through the selected routes R_{ik} and $R_{jk'}$, and if these two different operation routes conflict with each other during the time period $[s_i, r_i]$ or $[s_j, r_j]$, they should be prohibited coloring on these nodes for trains safety running, namely color taboo pairs sets ${}^{\circ}\mathbf{T}$, whose element can be expressed as $((v_i, c_k, R_{ik}), (v_j, c_{k'}, R_{jk'}))$.

Let x_{ik} denote whether the node v_i is covered by the color c_k (which also means whether the train is reassigned on the track or not), and it is a decision parameter. If the node is colored by c_k , then $x_{ik} = 1$; otherwise, $x_{ik} = 0$. Also, let e_i denote the earliness of the actual arrival time (its tardiness of the actual arrival time can be reflected by the tardiness of its actual departure time), and let t_i denote the tardiness of the actual departure time (In general, passenger train cannot be allowed to leave the station before its scheduled departure time due to service requirements), which can be calculated by Eqs. (5) and (6), respectively.

$$e_i = x_i - x'_i, \quad e_i > 0 \quad (5)$$

$$t_i = y'_i - y_i, \quad t_i > 0 \quad (6)$$

In order to avoid or alleviate the influences caused by delayed trains in certain stations on the whole network, it is better to use the minimum total weighting delay as the objective function to guarantee train running on schedule as much as possible, which can be expressed by (train weighing ω_i) Eq. (7).

$$f(\mathbf{G}) = \sum_{i=1}^n (\omega_i(e_i + \alpha t_i)), \quad (7)$$

where $\alpha(\alpha > 1)$ is the weight for the departure time of trains' tardiness (Compared with the earliness, tardiness is much easier to cause other delays in other stations in the whole network), and if $x_{ik} = 0(\forall k, c_k \in \mathbf{C}_i)$, $e_i = +\infty$, $t_i = +\infty$.

Thus, a mathematics optimal model with the objective function of minimum total weighting delay for track assignment adjustment problem can be written as

$$\text{Min } f(\mathbf{G}) = \sum_{i=1}^n (\omega_i(e_i + \alpha t_i)) \quad (8)$$

s.t.

$$\sum_{\forall k, c_k \in C_i} x_{ik} \leq 1 \quad \forall i \tag{9}$$

$$\sum_{\forall i} x_{ik}(t) \leq 1 \quad \forall k, \forall t \tag{10}$$

$$x_{ik} + x_{jk} \leq 1 \quad \forall k, \forall e_{ij} \tag{11}$$

$$x_{ik} + x_{jk'} \leq 1 \quad \forall k \neq k', \forall e_{ij}, ((v_i, c_k, R_{ik}), (v_j, c_{k'}, R_{jk'})) \in \circ\mathbf{T} \tag{12}$$

$$x_{ik} = 1 \text{ or } 0 \quad \forall k, \forall i \tag{13}$$

Formula (9) expresses that each node can be only covered by one color at most within the optional coloring set (one train can occupy one track at most according to their mapping relationship). Formal (10) means that each color can be used by one node at most at any time (one track can be occupied by one train at most at any time). Formal (11) demonstrates two adjacent nodes can be only covered by one color. Formal (12) shows that adjacent nodes can be covered with the constraint of the color taboo pairs set (the routes safety constraints of track assignment adjustment problem). Equation (13) is the value constraint of decision parameters.

3 Heuristic Algorithm Design

Track assignment adjustment problem in railway passenger stations is a real-time level problem, and the reassignment of tracks to trains is time-critical in real-time level. We can adopt a heuristic approach to solve such problem and meet the time requirement in limited iterations at the same time. A two-stage solution method is introduced in this paper. The first stage is to try to find an initial solution and gain much smaller values of the objective function with the help of appropriate rules. Then, the second stage is dedicated to solve some potential conflict trains and find a reasonable solution on basis of the results generated from the first stage. As a result, the proposed heuristic algorithm can be described as follows.

Step 1: Initiation. Input timetable, track and routes data, i.e., x_i, y_i, ω_i of the train v_i , track c_k and mapping relationship between trains and tracks C_i . Input various operation time standard parameters such as $\tau_{sf}, \tau_{zd}, \tau_{jc}, \tau_{fc}, \tau_{lg}, \tau_{crk}, \tau_{xh}$ and a safe time allowance σ , and tardiness weighing α , and so on.

Step 2: Parameter calculations. Calculate the initial actual arrival and departure time (x'_i, y'_i) by Eqs. (1) and (2) respectively, and the starting occupied and release time (s_i, r_i) by Eqs. (3) and (4), respectively, and the earliness and tardiness e_i, t_i by Eqs. (5) and (6). $x_{ik} = 0$, the unassigned trains set $\mathbf{J} = \emptyset$.

Step 3: Undirected graph construction. Taken trains as nodes, tracks as colors, mapping relationship as optional coloring sets, an undirected graph \mathbf{G} can be designed with the tracks conflicts pre-judgement formula proposed in Sect. 2.

Step 4: Color taboo pairs set design. For any two trains and two different tracks, each possible route should be checked and taken into consideration for safety reasons. If they conflicts, thus, ${}^{\circ}\mathbf{T} = ((v_i, c_k, R_{ik}), (v_j, c_{k'}, R_{jk'}))$.

Step 5: Prescheduling. With the rules of biggest weighting train first, smallest limitation (the number of optional colors for certain train) train first and train scheduled arrival time first in and first out at the same time, a new arrangement order of these nodes or trains can be gained as $\mathbf{V}' = \{v'_1, v'_2, \dots, v'_i, \dots, v'_n\}$, in which $\omega_i \geq \omega_{i+1}$, $|\mathbf{C}_i| \leq |\mathbf{C}_{i+1}|$, and $x_i \leq x_{i+1} (i = 1, 2, \dots, n - 1)$.

Step 6: Coloring procedures. Color each node in the undirected graph according to the new order \mathbf{V}' . For any node or train v_i , a color c_k can be selected from its optional coloring set \mathbf{C}_i on basis of the following rules: (i) the color c_k has not been colored on its adjacent nodes during the period $[s_i, r_i]$; (ii) according to the color taboo pairs set ${}^{\circ}\mathbf{T}$, the node v_i can be colored by c_k only when there are no other taboo colors selected in its adjacent nodes before; (iii) $c_k \in \mathbf{C}_i$; (iv) select the original track in its original track assignment plan as much as possible. If there is no feasible and safe color available for such node, then $\mathbf{J} = \mathbf{J} \cup \{v_i\}$.

Step 7: Judgement and recoloring. If $\mathbf{J} = \emptyset$, go to Step 8. Otherwise, adjust the actual arrival or departure time or reassign or exchange different tracks for unassigned trains or conflict trains, and then go to Step 2 (the second stage).

Step 8: Results output. Calculate the total weighting delay by Eq. (7). And then output all results, i.e., occupied tracks, platforms and routes, and their starting occupied and releasing time, actual arrival and departure time, and function values.

Without Step 7, the proposed algorithm can be viewed as the first stage one. Solution improving strategies such as adjust the actual arrival or departure time for these unassigned trains, and exchange or reassign tracks for conflict trains are introduced in Step 7 and it can be named as the second stage one. Its main purpose is to find a reasonable solution as quickly as possible to meet the time requirement in limited iterations. However, if there is enough time for railway dispatchers to adjust track assignment plans, we can also use other methods, i.e., using the methodology put forward in the Ref [4].

4 Case Study

In the real-time condition in someday at a busy and complex railway passenger station, there is a delayed train, whose new actual arrival time has been gained by the railway dispatcher in advance. There may be six trains that are needed to be adjusted in the next 30 min in the station. In order to avoid or alleviate other trains' tardiness caused by the delayed train, how to adjust these trains to make them pass through the station safely on schedule as more as possible?

Figures 1 and 2 show their original track assignment plans and graph coloring results respectively. With the proposed methodology in this paper, the new track assignment adjustment plan and its graph coloring result for these trains can be gained, which are shown in Figs. 3 and 4 respectively.

As shown in Figs. 3 and 4, the train v_2 has delayed (t_2 min). Since there are conflicts between the arrival route of the train v_2 and the arrival route of train v_4 , the train v_2 can arrive at the given track c_3 safely when the train v_4 has stopped by the track c_4 . Also, we should note that the train v_5 must be readjusted to other tracks from its origin scheduled stopping track c_3 , because there is not enough safe time allowance between the train v_2 and v_5 if they stop at the track c_3 at the same time. Since the train v_5 conflicts with the train v_4 on c_4 and the train v_6 on c_1 according to the color taboo pairs set, it should be adjusted to the track c_2 (reassigning tracks for conflict trains). In such way, there is no tardiness of the train v_5 ($e_5 = 0, t_5 = 0$) and it can run on schedule. Most trains can run on schedule except for the delayed train v_2 . The value of objective function can be calculated by w_2t_2 . Also, original delay t_2 is much more serious than the new delay t_2' with the new adjustment plan

Fig. 1 Track assignment original plans

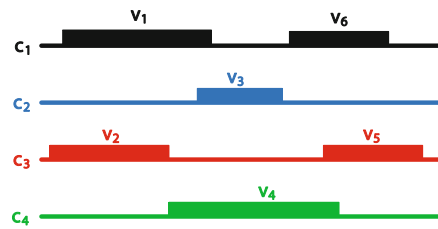


Fig. 2 Graph coloring original results

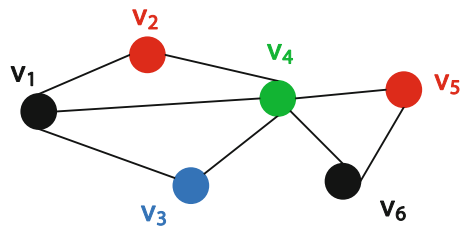


Fig. 3 Track assignment adjustment plan

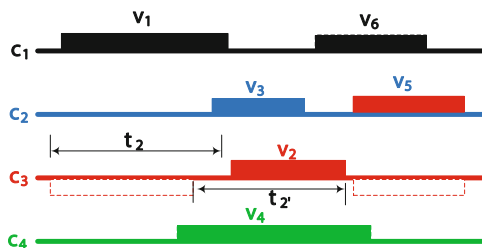
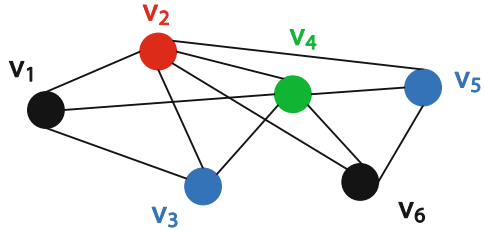


Fig. 4 Graph coloring adjustment results



(adjusting the actual departure time and reducing its operation time to some extent of delayed trains), which means the delay degree of the train v_2 has further alleviated at the railway station. It can be concluded that the track assignment adjustment plan can take the same tracks as many as possible compared to its track assignment original plan, which can help dispatchers to organize railway operations much more easily. From above analysis, the proposed methodology based on graph coloring theory can solve the problem effectively and efficiently.

5 Conclusion

In this paper, a graph coloring model and its mathematic optimal model with the minimum total weighting delay are proposed to describe the track assignment adjustment problem in busy and complex railway passenger stations. Also, a heuristic algorithm for such problem is designed. Case analysis results show that the proposed methodology can make a reasonable track assignment adjustment plan as quickly as possible which can meet the time requirement in limited iterations in the real-time conditions. However, since a train delay can lead to other trains' delay and such delay can also widespread among the whole railway network, it is better to take both track assignment adjustment problem in railway passenger stations in the real-time level and timetable rescheduling in the network level into considerations at the same time, which is our next research work.

Acknowledgements This study is funded by the National Key Research and Development Programme of China (2016YFB1200401), National Natural Science Foundation of China (71501190 & 61374157), and Beijing Jiaotong University State Key Laboratory of Rail Traffic and Control and Safety (RCS2017K001).

References

1. Cacchiani V, Galli L, Toth P (2015) A tutorial on non-periodic train timetabling and platforming problems. *Euro J Trans Logist* 4(3):285–320
2. Sels P, Vansteenwegen P, Dewilde T et al (2014) The train platforming problem: the infrastructure management company perspective. *Trans Res Part B: Methodol* 61(1):55–72

3. Carey M, Carville S (2003) Scheduling and platforming trains at busy complex stations. *Trans Res Part A Policy Pract* 37(3):195–224
4. Zhang Y, Lei D, Tang B et al (2011) Due windows scheduling model and algorithm of track utilization in railway passenger stations. *J China Railway Soc* 33(1):1–7 (in Chinese)
5. Lei D, Guo C, Zhang Y et al (2016) Reasonable scheduling for arrival-departure track operations in railway stations. *Trans Plann Technol* 39(6):624–639
6. Lusby RM, Larsen J, Ehrgott M et al (2013) A set packing inspired method for real-time junction train routing. *Comput Oper Res* 40(3):713–724
7. Samà M, Meloni C, D’Ariano A et al (2015) A multi-criteria decision support methodology for real-time train scheduling. *J Rail Trans Plann Manage* 5(3):146–162
8. Fang W, Yang S, Yao X (2015) A survey on problem models and solution approaches to rescheduling in railway networks. *IEEE Trans Intell Transp Syst* 16(6):2997–3016

Study on Dynamic Characteristics of Metro Stray Current Based on CDEGS

Mingjie Liu, Sheng Lin, Liping Zhao, Xiaohong Lin and Liang Chen

Abstract In this paper, a stray current dynamic distribution model was established by Current Distribution, Electromagnetic Field, Grounding and Soil Structure Analysis (CDEGS) software. The rectangle cut-and-cover model is based on a practical construction in Shijiazhuang Metro Line 3. It showed the dynamic distribution of conductor-to-earth voltage and stray current under constant-current-strategy condition and target-speed-strategy condition. The traction current of the former one is changeless, but the later one should take into account passenger comfort and operational efficiency. The simulation results give several conclusions: Traction operation strategy and locomotive position have great impacts on the distribution of voltage and current. Through the analysis of various parameters, some efficient measures can be used to avoid the overvoltage condition and serious metal corrosion.

Keywords Stray current · CDEGS · Target-speed-strategy · Corrosion

1 Introduction

The rapid development of metro system has contributed significantly to resident travel. However, the mass density operation has also brought numerous problems, e.g., metal corrosion problem. Since the rail itself have finite resistance and poor insulation from the earth, parts of traction current leak from the rail into the soil and finally flow back to the negative side of the traction substation, which is called “stray current” [1]. This part of nontarget current even affects the main transformer of urban power grid, which will result in direct current magnetic bias phenomenon. Thus, the research on stray current has a vital significance of the protection of the metro system and the surrounding metal devices.

M. Liu · S. Lin (✉) · L. Zhao · X. Lin · L. Chen
School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China
e-mail: slin@swjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_58

Constructing a suitable tunnel model is the prerequisite for analyzing the stray current. At present, the research models of stray current in metro system can be divided into three forms, such as “rail-earth” two-layer structure, “rail-current drainage net (CDN)-earth” three-layer structure and “rail-CDN-structural metal-earth” four-layer structure [2, 3]. This paper will choose the four-layer network model to research the stray current.

Instead of analytical method, CDEGS produced by Safe Engineering Services & technologies Ltd. (SES) [4] is a convenient tool for the operational simulation. This paper introduces a new way to research the stray current by modeling on CDEGS. By observing the current and conductor-to-earth voltage, it reveals the distribution of dynamic stray current. Finally, this research proposes a useful method for judging the corrosion degree at different locations, so the corresponding solutions can be used for maintenance and repair.

2 Impact of Stray Current and the Modeling

2.1 Original Stray Current

DC 750 V/1500 V voltage is mostly applied to the metro traction power system. The current flows from overhead line system and pantograph to the locomotive, then into the steel rail and finally back to the negative of the DC rectifier. Because running rail is not completely insulated from the earth, there is a potential difference between them. Some stray current leaks from the rail to soil before returning back to the rail or the anode of traction substation [5, 6].

In order to reduce the negative influence of stray current, CDN is usually laid under the rails to provide circulation path for it. Another way of minimizing the negative impact is increasing the resistivity between rails and soil. It is usually more than $15 \Omega \cdot \text{km}$. Normally this can be achieved by lying insulation pads under the rails [7]. While as time goes by, the resistivity will drop to $3 \Omega \cdot \text{km}$ or even lower due to the attachment of large dirt and dust. It significantly affects the resistivity of insulation pads.

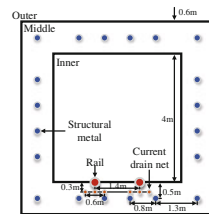
2.2 Metro Tunnel Modeling

The rectangle structure model is built using Frequency Domain Grounding Analysis (MALZ) in CDEGS. This model is more suitable for assessing the dynamic stray current in cut-and-cover project. It contains rails, CDN, and structural metals. Table 1 shows the basic parameters of each section based on actual project data and, Fig. 1 shows the metro construction model.

Table 1 The parameters of components

Parameter	Value	
Track length and power supply	2 km single track with a traction substation at either end	
Rail	Resistance	40 m Ω/km
	Radius	0.035 m
	Amount	2 lines
Drain net conductor	Resistance	660 m Ω/km
	Radius	0.008 m
	Amount	6 lines
Structural metal	Resistance	320 m Ω/km
	Radius	0.01 m
	Amount	20 lines
Insulation pad resistivity	450,000 Ω · km	
Soil model resistivity	Inner (air): $1 \times e^{18}$ Ω · km	
	Middle (concrete): 2000 Ω · km	
	Outer (soil): 200 Ω · km	

Fig. 1 Metro tunnel modeling



In CDEGS, the resistivity of insulating pads must be converted to a coating of a given resistivity that is placed uniformly along the rails. A rail coating resistivity of 450,000 Ω · km is equivalent to an insulating pad resistivity of 15 Ω · km [8].

In this paper, the continuous power supply current curve will be divided into multiple intervals depending on the running conditions and current changes. A continuous simulation will be divided into numerous static processes. Except for the numerical difference, the distribution of stray current on CDN and structural metal is basically the same, so this paper only analyzes the distribution of CDN.

3 Constant Current Strategy

In order to observe the variation in rail parameters, before considering the actual operation strategies, we presume that the traction current on the rail is equal at any position. In this section, the supply current is 2000 A constantly in bilateral power supply system.

Figures 2 and 3 show the current and the voltage of unilateral rail and the outermost conductor of CDN when the locomotive is in different sites. Among them, each dashed line represents the result of entire conductor in each simulation. Assume that the direction of current flows to the left end or out of the conductor is positive.

From Fig. 2a, it is obvious that the longitudinal current of the rail in bilateral power supply system is not a smooth curve but there exists a mutation at the locomotive location. The amount of the traction current from both sides depends on the distance between them. However, the current is always equal to 2000 A. Figure 2b shows that the rail-to-earth voltage reaches its maximum positive value at the locomotive location on each curve. By contrast, the values of both ends are negative and almost equal. Another difference to the longitudinal current in Fig. 2a is that the voltage curve in this picture is continuous without obvious drops.

According to Fig. 3a, b, for the longitudinal current of drain net conductor (DNC), there exists a valley between two distinct peaks. In fact, the path of the valley coincides with the time–distance curve of the locomotive. The maximum values of stray current on both sides of the DNC do not appear at the same time but in the vicinity of 700 and 1300 m, respectively, when locomotive runs to the center of the rail. While the total current of the whole DNC reaches its maximum value when locomotive runs to the middle of the rail.

Considering the change of DNC-to-earth voltage in Fig. 3c, the voltage near the locomotive is smoother than that in Fig. 2b. Different from the rail, the voltage at the end of the DNC is not absolutely negative if locomotive approaches the traction substation on that side.

Fig. 2 Rail simulation results under constant current strategy. **a** Current of rail. **b** Rail-to-earth voltage

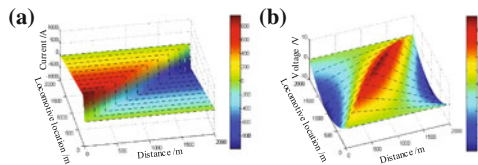


Fig. 3 DNC simulation results under constant current state. **a** Current of DNC (3-D diagram). **b** Current of DNC (projection diagram). **c** DNC-to-earth voltage

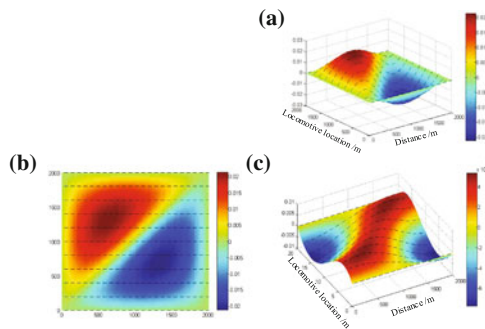
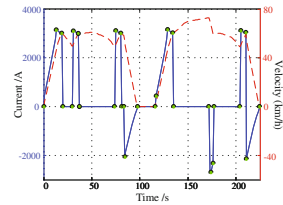


Fig. 4 Current–time curve and velocity–time curve



4 Target-Speed-Strategy

4.1 Time–Current Curve

In this paper, the part power supply sections of Shijiazhuang Metro Line 3 is a good case which combines three independent stations (Xisan Village, Aquatic Park, and Bolin Village). DC traction power system (DCTPS), a traction calculation software produced by Southwest Jiaotong University [9], is a convenient tool to calculate the time–current curve and time–distance curve. Thus, the distance–current curve can be obtained. By filtering the characteristic points in the last curve, the injection currents at each point on the rail can be used as the inputs to CDEGS to complete the dynamic stray current simulation.

According to the simulation of the DCTPS, current–time curve and time–speed curve under target-speed-running strategy are shown in Fig. 4. This strategy is defined by setting a target speed interval for those locomotives, where the range is 50–75 km/h in this case.

The whole movement process consists of three statuses: acceleration, coasting, and braking. At the beginning, the locomotive speed increases slowly. Then by controlling the traction current, the locomotive runs in the speed limit interval until it close to the next station. Concerning the passenger comfort and operational efficiency, the reverse current will slow down the locomotive. The green points in Fig. 4 indicate the current mutation. These points will be brought into CDEGS as input parameters so that a constant movement will be divided into finite independent static simulations.

4.2 Stray Current Distribution Under Target-Speed-Strategy

After integrating simulation results in CDEGS, Figs. 5 and 6 show the current and conductor-to-earth voltage of the rail and the DNC, respectively.

Different from the constant current condition, the practical working condition is more complicated. Due to the discontinuity of the supply current for the locomotive, the currents of the rail and the DNC are also intermittent. Sometimes the traction currents at two moments are not equal, but the total stray current is almost

Fig. 5 Rail simulation results under target-speed-strategy.
a Current of rail.
b Rail-to-earth voltage

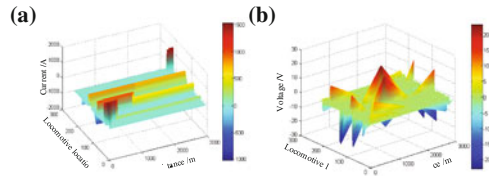
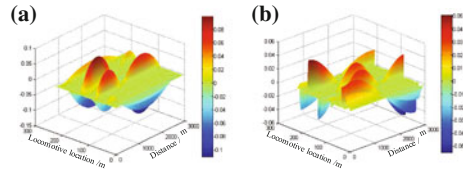


Fig. 6 DNC simulation results under target-speed-strategy.
a Current of DNC.
b DNC-to-earth voltage



the same. This condition also proves that locomotive position has great influence on the output of the transformer and the distribution of stray current.

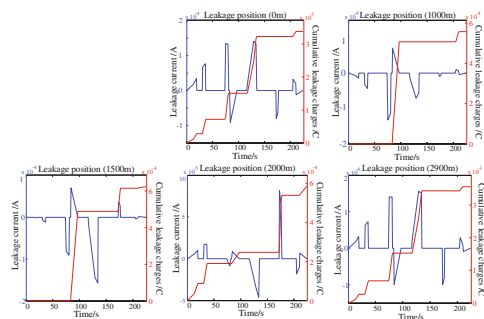
Rail-to-earth voltage is the reflection of leakage current density. The maximum voltage is near 1400 m of the rail at 129ths in Fig. 5b and at the left end of the DNC at 173ths in Fig. 6b. In fact, if the rail-to-earth voltage is higher than 90 V under the normal circumstances, the overvoltage protection device requires work to avoid electric shock or corrosion [10].

When locomotive operates in the state of acceleration or breaking, the large traction current will directly aggravate the leakage of stray current and bring intense fluctuation. Experience proves that no damage occurs on the tracks over a period of 25 years if the average stray current per unit length does not exceed the value of 2.5 mA/m [5].

Figure 7 shows the leakage current per unit length and the cumulative leakage charges versus time curves of seven observation points at the DNC under once complete operation.

According to the corrosion theory, when charge flows out of the DNC, it will cause metal corrosion. At this time, the leakage current is usually positive. The amount of charge flowing out of DNC is the integral of the positive leakage current

Fig. 7 Leakage current and cumulative leakage charges on each observation point



versus time and it is proportional to the amount of metal corrosion. In Fig. 7, the red curve is the integral of the positive value of the blue curve versus time. Compared with other five points, positions at 0 and 2900 m are more vulnerable to corrosion. The same method can be used for other conductors.

In this way, the amount of corrosion at each point of the conductor can be calculated. Then combined with different operating strategies and daily train diagram, we can estimate the life of the rail.

5 Conclusion

In this paper, we proposed a metro stray current distribution model using CDEGS and analyzed the influence of stray current. The research of constant current strategy showed that locomotive position has a great impact on the output of the substations on both sides. Furthermore, the supply current will, in turn, affect the distribution of stray current combing locomotive position. The total amount of stray current leaking into the soil reaches its maximum value at the moment that the locomotive runs to the center of the rail. At this time, the stray current of DNC is the largest and the absorption rate is about 96%. The maximum value of rail-to-earth voltage is always at the locomotive position. DNC is the same. However, the trend of rail current is opposite to that of DNC. These phenomena are more prominent in the actual process. According to the analysis of target-speed-running strategy, some measures should be taken to prevent the conductor from exceeding the voltage threshold, and help them to avoid being corroded through calculating the accumulated leakage charges at each point. These necessary methods will contribute to the metro safe operation.

Acknowledgements The authors would like to express our thanks to the editor and anonymous reviewers for their help in revising the manuscript. This research is sponsored by the National Natural Science Foundation of China (No. 51677153), Science & Technology Project of Guangdong Province (2015B010128002), and Collaborative Innovation Centre of Zhuhai (approval numbers: 55560307).

References

1. Zhao LP, Li JH, Liu MJ (2016) Simulation and analysis of metro stray current based on multi-locomotives condition. In: Control Conference (CCC), 2016 35th Chinese, IEEE, pp 9252–9258
2. Yu K (2015) The simulation of metro stray current based on CDEGS. Southwest Jiaotong University (in Chinese)
3. Charalambous C, Cotton I, Aylott P (2013) Modeling for preliminary stray current design assessments: the effect of crosstrack regeneration supply. *IEEE Trans Power Deliv* 28(3): 1899–1908

4. Charalambous C, Aylott P (2014) Dynamic stray current evaluations on cut-and-cover sections of dc metro systems. *IEEE Trans Veh Technol* 63(8):3530–3538
5. Pires CL (2016) What the IEC tells us about stray currents: guidance for a practical approach. *IEEE Electr Mag* 4(3):23–29
6. Lucca G (2015) Estimating stray current interference from DC traction lines on buried pipelines by means of a Monte Carlo algorithm. *Electr Eng* 97(4):277–286
7. Zhang DL (2012) Research on protection technique of urban rail transit DC traction reflux system. China University of Mining and Technology (in Chinese)
8. Cotton I, Charalambous C, Aylott P et al (2005) Stray current control in DC mass transit systems. *IEEE Trans Veh Technol* 54(2):722–730
9. Xu LL, Liu W, Liao J et al (2016) The analysis on the measured traction energy consumption and braking energy consumption of urban rail transit. *J Railway Sci Eng* 13(9):1818–1824 (in Chinese)
10. Zhang DL (2014) Rational distribution of rail potential and OVPD's switch in multiple metro intervals. *Urban Mass Transit* 17(9):53–56 (in Chinese)

Analysis of Influencing Factors of Energy Saving and Emission Reduction for Railway Transportation of China

Liang Sun, Huiting Guo, Li Wang, Liang Chen, Jianhua Chen, Mei Liu and Wei Bao

Abstract The current status of energy saving and environment protection in railway sector is analyzed in this paper. The passenger transport-related energy consumption has degraded dramatically, and the freight-related energy consumption on the decline in general, but the energy consumption per unit grows slightly in last two years, due to the shrinking of rail freight volume. Then, the key effective factors of energy saving and environment protection are presented from the aspects of key facilities and equipment and operation management. Finally, based on the development status of China railway, some new technologies and measures of energy saving and environment protection in four perspectives are proposed, such as life cycle management of railway system, lightweight cabin design and manufacture, comprehensive energy consumption control and energy recycle, and high-efficiency operation.

Keywords Railway transportation · Energy saving · Emission reduction Evaluation index system

1 Introduction

As the most sustainable way of transportation, railway transit has been seen as the main artery of the national economy and the backbone of modern city operation, the development of Chinese social economy and national security have been fully supported by which. Recently, China's railway network has been increased rapidly,

L. Sun · H. Guo · L. Chen · J. Chen · M. Liu · W. Bao
China National Institute of Standardization, 100191 Beijing, China

L. Wang (✉)
School of Traffic and Transportation, Beijing Jiaotong University, 100044 Beijing, China
e-mail: wangli@bjtu.edu.cn

L. Wang
Beijing Engineering Research Center of Urban Traffic Information Intelligent Sensing and Service Technologies, 100044 Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_59

this also leading the trend of high-speed railway construction in the world. End of 2016, China's rail operational mileage has been reached 124,000 km, among this, the mileage of high-speed train has been reached 22,000 km, exceeds 60% of the worldwide mileage. According to the medium and long-term Plan for rail network (2016–2015), till 2020, Chinese rail network mileage will be reached 150,000 km, and the mileage of high-speed train will be reached 30,000 km, 80% more megacities will be covered. Till 2025, rail network mileage will be reached 175,000 km, and the mileage of High-speed train will be reached 38,000 km, more area will be covered, and the structure of network will be optimized more, and the role of backbone will appear more outstandingly. Optimizing the wide range national rail network, building up the modern high-speed rail network, creating integrated and comprehensive transportation junction will be the key contents for the construction of Chinese rail transportation infrastructure in the next 10–15 years.

Under the background of a wide range of network running, the energy consumption of railway transportation has been paid more and more attention to. Fernández-Rodríguez et al. designed the robust and energy-efficient ATO speed profiles of metropolitan lines considering trainload variations and delays [1]. Huang et al. proposed an improved ant colony optimization methodology for energy optimization of train operation [2]. Hamid et al. presented a single train trajectory optimization method to minimize overall energy consumption. It then explores how errors in train position data affect the total consumed energy of railway trains, with regard to the tractive force due to gradient when following the optimized trajectory. Also, a genetic algorithm was used to optimize the train speed profile in [3]. Hu et al. investigated the voltage profiles and harmonic impacts of high-speed trains on the traction power supply system and primary utility system, and studied power-quality impact assessment models, methods, and applications using train timetable [4, 5].

The main energy consumption factors of railway were also analyzed by some scholars. Shi et al. carried out analysis on the influence of energy consumption from time, station, line, area, such as vehicles, equipment, season 7 elements, and established the evaluation system of energy consumption [6]. Wang et al. analyzed the energy consumption influence and measures of the line design, technology such as driving speed selection, marshaling technology, strategy, and application of new technology [7]. Sun et al. established the gray correlation degree model in order to quantitatively determine the factor of energy consumption intensity, based on the identification of the factors of railway passenger transport energy consumption [8]. De et al. gave an innovative optimization framework for the definition of energy-efficient speed profiles within rail traffic by means of supply design models [9], taking into account the operational requirements of the services, such as trains scheduling, absence of or small allowance for delays, and respect for buffer times for passenger transfer at connecting stations.

The current status of energy saving and environment protection in railway is analyzed. Some key factors of energy saving and environment protection are present from the aspects of key facilities and equipment and operation management.

Finally, some new technologies and measures of energy saving and environment protection are proposed based on the development status of China railway.

2 Energy Consumption Analysis of Rail Transportation

Compared to other types of transportation, rail transportation has the incomparable advantages in the aspect of sustainability development, such as, first, low energy consumption, counts per passenger-kilometer as the unit, the energy consumption of road transit is 2.3 times than railway, and the air transit is 7 times more than the railway. Second, low environmental pollution, counts per passenger-kilometer as the unit, the emission of air pollutants of road transit is 10–20 times than railway, and the air transit is 100 times than railway. Third, low accident rate, the accident rate of bus service is 2 times more than railway, and cars are 20–30 than the railway. Fourth, lowland occupation, under the same transport volume, the land occupation of road transit is 3.5 times than railway. Fifth, low noise, the noise effect to the city of rail is lower than both road transit and aviation. Sixth, high transportation capacity, the single carrying capacity of the railway is bigger, and the departure density is larger, also with a better ability of integrated transit than the other ways of transportation. Seventh, better adaptability to the environment, the railway will be affected less by external factors, such as weather and traffic jam, etc., could be operated around the clock, has a higher on-time rate. Eighth, low cost of construction, the railway is cost-effective on traveling.

Refer to national railway statistical bulletin of the year 2003–2016, we can get the energy consumption data of railway transportation of China, which including the passenger capacity (10^2 million person), the passenger turnover (trillion person-kilometers), the freight capacity (10^2 million tons), the freight turnover (trillion ton-kilometers), the operating mileage (10 thousand kilometers), the electrification mileage (10 thousand kilometers), the comprehensive energy consumption per unit transportation workload (tons of standard coal/million converted ton-kilometers), and the data of the sulfur dioxide emission (10 thousand tons). All these items can be represented by B, C, D, E, F, G, H, and I, respectively, shown in Table 1. The losing data about sulfur dioxide emissions between 2003 and 2006 can be obtained by energy consumption prediction regression model according to [10].

Based on the passenger turnover and freight turnover, we get the energy consumption per unit, shown in Table 2 and Fig. 1.

It can be found that the ratio of comprehensive energy consumption to passenger turnover and the ratio of sulfur dioxide emission to passenger turnover decrease dramatically from 2003 to 2016, while the ratio of comprehensive energy consumption to freight turnover and the ratio of sulfur dioxide emission to freight turnover are generally a downward trend, but they increased slightly in the past two years.

Table 1 Energy consumption and related factors of railway

Year	B	C	D	E	F	G	H	I
2003	9.73	0.48	22.35	1.71	7.30	1.88	7.32	4.38
2004	11.18	0.57	24.81	1.91	7.44	1.93	6.65	4.54
2005	11.56	0.61	26.83	2.05	7.54	2.02	6.48	4.61
2006	12.57	0.66	28.71	2.17	7.70	2.44	6.12	4.85
2007	13.17	0.72	31.45	2.4	7.80	2.55	5.78	4.36
2008	14.61	0.78	32.59	2.48	8.00	2.80	5.6	4.23
2009	15.25	0.79	33.2	2.49	8.60	3.60	5.3	4.02
2010	16.76	0.87	36.42	2.76	9.12	4.25	5.01	4.03
2011	18.62	0.96	39.32	2.95	9.33	4.61	4.76	4.01
2012	18.93	0.98	39.04	2.92	9.76	5.10	4.74	3.78
2013	21.06	1.06	39.66	2.92	10.31	5.58	4.66	3.53
2014	23.05	1.12	38.13	2.75	11.20	6.50	4.55	3.17
2015	25.35	1.20	33.58	2.37	12.10	7.40	4.68	2.88
2016	28.14	1.26	33.32	2.38	12.40	8.00	4.71	2.39

Table 2 Energy consumption ratio

Years	H/C	H/E	I/C	I/E	H/(C+E)	I/(C+E)
2003	15.2500	4.2807	9.1250	2.5614	3.3425	2.0000
2004	11.6667	3.4817	7.9649	2.3770	2.6815	1.8306
2005	10.6230	3.1610	7.5574	2.2488	2.4361	1.7331
2006	9.2727	2.8203	7.3485	2.2350	2.1625	1.7138
2007	8.0278	2.4083	6.0556	1.8167	1.8526	1.3974
2008	7.1795	2.2581	5.4231	1.7056	1.7178	1.2975
2009	6.7089	2.1285	5.0886	1.6145	1.6159	1.2256
2010	5.7586	1.8152	4.6322	1.4601	1.3802	1.1102
2011	4.9583	1.6136	4.1771	1.3593	1.2174	1.0256
2012	4.8367	1.6233	3.8571	1.2945	1.2154	0.9692
2013	4.3962	1.5959	3.3302	1.2089	1.1709	0.8869
2014	4.0625	1.6545	2.8304	1.1527	1.1757	0.8191
2015	3.9000	1.9747	2.4000	1.2152	1.3109	0.8067
2016	3.7381	1.9790	1.8968	1.0042	1.2940	0.6566

Along with the electrification reconstruction of national railway, and the development of high-speed railway, and the growth of passenger turnover, the energy consumption of passenger transit is downgraded rapidly. Although the freight-related energy consumption on the decline in general, the energy consumption per unit has been increased slightly in last two years, due to the shrinking of rail freight volume.

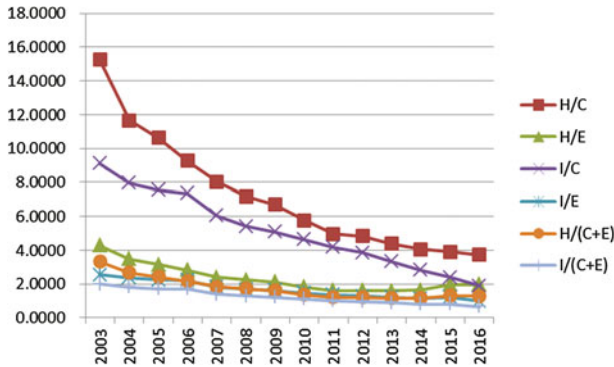


Fig. 1 Energy consumption ratio

3 Factors Analysis of Energy Saving and Emission Reduction

With the rapid development of national high-speed railway in China, the scale of railway station enlarges constantly. Although the railway freight turnover declines continuously during last three years, the demand of the railway passengers increases substantially. The energy consumption of railway is only next to water transportation ranked fourth in the five kinds of transportations. The main factor of the energy consumption of railway is to accomplish the various types of train operations for the demand of passengers and freights. Meanwhile, factors, such as the level of the transportation operation management and the level of the transportation equipment technic, can affect the railway energy consumption too.

- (i) Effect of the train operation. The energy consumption in the process of the train operation is due to various resistances of the train, the circuit conditions, the braking, and the acceleration system of the train and another kind of train operation mainly. There are two kinds of resistances of the train operation: the mechanical resistance and the aerodynamic resistance. The mechanical resistance contains the bearing drag resistance, the wheel rolling resistance, the wheel slip resistance, and the shock and vibration resistance. The aerodynamic resistance mainly caused by the relative movement of the ambient air due to the train operation, which is relative to the speed, the shape, and the size of the train. It is considering that the basic resistance of the train operation is directly proportional to the square of the train speed generally, which means, the higher the speed is, the greatest of the driving resistance and the higher of energy consumption will be. Moreover, the circuit condition is an important factor for energy consumption. The large line slope and huge line transformation can not only cause the energy consumption to overcome the gravity of the trains but also increase the shock and vibration resistance. The emergencies often cause the waste of energy, such as the

temporary parking and limited speed due to the emergency equipment failures or the boundary intrusion. To keep the safety of the train operation, the drivers need to operate the acceleration and the braking system frequently, which can cause more waste of energy.

- (ii) Effect of the management level of the transportation operation. In the condition of the fixed transportation capacity, if the number of transport production was higher, then the energy consumption of per unit production is lower, which means the efficiency of the transportation operation management is higher. However, the transportation organization management contains the transportation operation models, the operation scheme, the operation plan, the maintenance plan, the crew plan, and contents of each level. For the freight transportation, it involves the loading plan, the freight plan, the shunting plan, and other contents. The transportation operation scheme with high efficiency can take full advantage of the network transportation capacity. In the same condition, improving the passenger turnover and the freight turnover can reduce the energy consumption of per unit transportation capacity efficiently.
- (iii) Effect of the technical level of the transport equipment. The technology level of China's railway transportation equipment has improved a lot compared to the past. The electrified railway has taken over 60% of the national railway. The energy consumption can be affected by the efficiency of the key technical equipment, such as all kinds of basic signal devices, the operation blocking system, the station chain system, the traction control system, and the train control system.

4 Strategies for Energy Saving and Emission Reduction

Along with the promoting of system construction and operation for a high-speed train and intercity fast line, the management of Chinese railway operation has transformed from line independent operation to network operation. For promoting the overall efficiency level of Chinese railway, the development of new technology in energy saving and emission reduction should be paid a tension on. Four perspectives are described in this section, life cycle management of railway system, lightweight cabin design and manufacture, comprehensive energy consumption control and energy recycle, and high-efficiency operation management.

- (i) Energy consumption control technology within life cycle. From design, construction, operation, maintenance and disposal 5 stages of the railway system, the systematic energy consumption and carbon emission in life cycle perspective should be calculated, and the specific calculation method for each stage shall be established, the lowest energy consumption and emission throughout the full life cycle could be achieved.

- (ii) Design and manufacture technology for lightweight cabin. The air resistance could be reduced by optimizing cabin shape design, and the energy consumption of operation could be lowered effectively. The main energy consumption being used is to overcome the air resistance while train operation. By optimizing the shape design, using aluminum alloy and composite material and all kinds of lightweight structure, in order to relieve the cabin weight, the mechanical resistance of train operation could be reduced.
- (iii) Hybrid power technology. By studying electricity and diesel dual-power mode, the electricity-drive mode train could be applied on the truck line, and the diesel mode could be used in branch line that without a circuit, the exchange of passenger will be reduced, the travel speed will be increased, and the travel time will be shorter.
- (iv) Energy comprehensive control technology. By applying smart technology, aiming the best energy saving as the target, according to the train operational line condition, climate condition, passenger flow volume, etc., the comprehensive control on the train traction, air conditioning, lighting, door switch, energy recovery and storage, etc., should be applied, the maximum energy saving could be achieved.
- (v) Energy recovery technology. Under considering the braking safety requirement of the railway, the methods on friction braking, regenerative braking, eddy-current braking, resistance braking will be studied continually, and the energy recovery technology on energy dissipation type, energy reclaiming type and energy storage type will be also studied. Under rational matching, the optimal status will be achieved within the whole system.
- (vi) High-performance operation management technology. The low energy consumption operation management mode which is suitable for China should be built up. By arranging the different level of speed high-speed train, the passenger attendance rate could be raised. By operating an express train on normal speed railway, fulfilling the transportation market for scattered and small packages, and developing fast freight services, developing high-speed train express service, the freight turnover will be increased. By studying cargo train transport planning technic, cargo real-time tracking, and dispatch control technics, trunk line railway and cargo fast receive-delivery coordination and other technics, the level of energy consumption and emission could be lower during the train operation.

5 Conclusions

The energy consumption of passenger transit is downgraded rapidly compared with the freight-related in recent years, due to the continuous growth of passenger turnover and decline of the freight volume. The construction of high-speed freight railway will resolve the problem in the future. Furthermore, some new technologies

and strategies for energy saving and emission reduction for railway transportation should be developed along with the high-speed railway network infrastructure.

Acknowledgements This study is funded by the National Key Research and Development Programme of China (2016YFB1200401), the Fundamental Research Funds for the Central Institute (542016Y-4479), National Natural Science Foundation of China (71701010), and the Fundamental Research Funds for the Central Universities (2017JBM033).

References

1. Fernández-Rodríguez A, Fernández-Cardador A, Cucala AP et al (2015) Design of robust and energy-efficient ATO speed profiles of metropolitan lines considering train load variations and delays. *IEEE Trans Intell Transp Syst* 16(4):2061–2071
2. Youpeng H, Yang C, Shaofeng G (2016) Energy optimization for train operation based on an improved ant colony optimization methodology. *Energies* 9(8):626
3. Hamid HA, Nicholson GL, Douglas H, et al (2016) Investigation into train positioning systems for saving energy with optimised train trajectories. In: 2016 IEEE International Conference on Intelligent Rail Transportation, Birmingham. 2016.8.23-25
4. Hu H, He Z, Li X et al (2016) Power-quality impact assessment for high-speed railway associated with high-speed trains using train timetable—part I: methodology and modeling. *IEEE Trans Power Deliv* 31(2):693–703
5. Haitao H, Zhengyou H, Ke W et al (2016) Power-quality impact assessment for high-speed railway associated with high-speed trains using train timetable—part II: verifications, estimations and applications. *IEEE Trans Power Deliv* 31(4):1482–1492
6. Shi JingYa S, Yongqing YJ (2008) Factors analysis of energy consumption of railway transportation and establish of energy consumption estimation system. *Railway Trans Econ* 30(9):46–49 (in Chinese)
7. Wang Y, Liu M, Zhao Y (2011) Study on energy-saving measures of urban rail transit system. *Comprehensive transportation* (4):57–59 (in Chinese)
8. Sun Q, Xu C (2012) Analysis on running organization factors and Gray relation grade measure of railway energy consumption. *Railway Trans Econ* 34(8):56–61 (in Chinese)
9. De Martinis V, Weidmann UA (2015) Definition of energy-efficient speed profiles within rail traffic by means of supply design models. *Res Trans Econ* 54:41–50
10. Sun L, Lin L, Chen L, et al (2016) Railway energy consumption analysis based on regression model. In: The 7th International Conference on green intelligent transportation system and safety, Nanjing. 2016.7.1.-4

Study on the Method of Passenger Transport Organization in Peak-Hour of Railway Station

Yuyang Nie and Xiaoning Zhu

Abstract In recent years, the passenger capacity of the railway continues to soar, increasing the pressure of passenger transport organizations. Although many stations would increase passenger capacity by building a new equipment and facilities, the effect is still not satisfactory, if the new technology and equipment can be applied at the same time, and we give full consideration to the characteristics of passenger flow, concern about improving passenger transport efficiency, then the passenger capacity can be enlarged greatly. This essay aims to find out take effective organizational methods to disperse the streamline of people to avoid the peak-hour crowding and introduces the concept of the peak-hour crowding, and passenger flow characteristics. And then, the passenger streamline organization and passenger capacity of equipment were discussed and studied. It focuses on the composition of passenger and characteristics of peak-hour crowding in domestic railway station. And analyze how to computing station equipment capacity, the way to optimize the management of passenger streamline in peak-hour.

Keywords Railway station · Peak-hour of transportation · Organization of passengers · Streamline of passenger

1 Introduction

Passenger flow is that based on the passenger's own travel needs to choose their own vehicles, within a certain time and space for the purpose of moving the location [1]. Passenger flow peak is in a certain period of time, passenger flow increase rapidly, so ordinary passenger organization method cannot meet the requirements of passenger travel, it should find effective organization method, to

Y. Nie · X. Zhu (✉)

School of Traffic and Transportation, Beijing Jiaotong University,
Haidian District, Beijing, China
e-mail: 16120867@bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_60

587

maximize station equipment capacity [2]. The work of the passenger station is the key to improving the competitiveness of the other modes of transport.

For the large station is relatively crowded in peak-hour, so the passenger capacity of the train station facing a serve test. Therefore, it is the key to improve the efficiency of passenger transportation and alleviate the pressure of capacity of the station.

2 Passenger Flow Characteristics in Peak-Hour

Peak passenger flow is a special form of passenger, is in a special period of time, such as the Chinese New Year holiday, winter and summer vacation period, and is a large-scale passenger flow. We can classify passenger flow into Spring Festival passenger flow, summer transport passenger flow, holiday flow and other different categories. Because the number of passengers required to transport is huge and the period of time is very short so there are some other passenger flow characteristics compare with normal passenger flow.

The traditional peak passenger flow generally occurs during the Spring Festival, but with the masses of people's quality of life greatly improved, the people's consumption concept has changed dramatically, the growing demand for travel is growing rapidly. So the peak-hour is also become frequent, which means more challenges to passenger organization.

2.1 *The Composition of Peak Passenger Flow*

In general, for the peak passenger flow from its composition structure can be classified into five types of passenger flow [3]:

- (1) Student passenger flow, the beginning and end of the summer and summer vacation and the beginning and end of each festival are the students when the passenger flow occurs.
- (2) Migrant workers passenger flow, with the pace of economic construction faster and faster, a large number of rural population come into the city to work, during the Spring Festival holiday, most migrant workers commute to the city and town.
- (3) Tourist flow, travel passenger flow, including group travel and personal travel, as well as the travel agencies organized by travel agency.
- (4) Business passenger flow, thanks to the rapid development of high-speed rail, high-speed rail became an indispensable choice for the businessmen.

2.2 Passenger Flow Characteristics in Peak-Hour

(1) The concentration of passenger travel time

Peak passenger flow is generally concentrated in the Spring Festival and Labor Day and the National Day and other long holidays, and there is a strong regularity, such as the peak transportation before the Spring Festival and the peak transportation after the Spring Festival, National Day, before and after the Labor Day are when the peak passenger flow occurs.

(2) The travel pressure of passenger

The station is very crowded at the peak-time, so generally make passenger anxious, they eager to quickly leave the station, especially for those who carry valuables asset, they are more worried about the safety of their property. So they are likely to have serious psychological problems and other excessive behavior.

(3) All-weather

When in the peak-hour passenger flow, the transport capacity is generally unable to meet the requirements of passenger, most passengers must spend a lot of money to buy tickets, and often stay in the station for a long time.

(4) Guidance

Because the transport capacity cannot be stored, it is easy to have a shortage of transport capacity. At this time the government can call the passengers through the news media choose a reasonable time to travel.

(5) Travelers carry more luggage

Ordinary passengers are usually only with ordinary daily necessities, but as to a particular holiday people tend to carry gift, specialty or travel supplies, it is increased the risk of the luggage Loss, criminal cases may increase.

3 Passenger Station Transport Organization in Peak-Hour

3.1 The Facility of Station

The passenger train station is the station responsible for the management of the passenger transport service and is the closest link between the passenger and the railway, and the railway Legal rights and obligations arise when passengers arrive at the station and disappear after the passenger leaves the station. Passenger station consists of station room, station and station square.

- (1) Station room. There are specifically for the passenger services of the room, but also specifically for the management of office space, as well as for the package, mail room.

Passenger station should be in the city's traffic network, to facilitate the transfer of passengers. All types of streamlines must ensure that each other does not interfere with others. The station should be practical and beautiful and show the local style characteristics, but also have good ventilation and lighting, and have a good heating equipment and air conditioning equipment.

- (2) Station. Station is a place for passenger technology operations, it consists of lines, platforms, canopy, crossline equipment and other components.
- (3) Station Square. Consist of the roadway, parking and passenger activities and other areas. Station square can also be used as a temporary rally place. In order to ensure the safety and convenience of passengers and vehicles, the construction of the station should be closely coordinated with the city planning.

3.2 Analysis of the Capacity of Passenger Station

Analysis of Capacity of waiting room

The waiting room is the place that passenger staying for longtime in the terminal. And there should have a convenient and fast access between station entrance and ticket gate.

The layout of the waiting room is usually divided into two forms:

- (1) Centralized waiting way: that is a large number of passengers waiting at the same time, this way is convenient and simple. The utilization rate is relatively high. It is also the most basic and most common way. However, if the passenger flow is large, you want to maintain the normal waiting order is not very realistic, it is very likely to have a negative effect on the passenger's waiting time.
- (2) Subline waiting way: This arrangement is the opposite with the first, the purpose is to divert passengers to ease the pressure of a single waiting port. So the terminal with relatively large passenger flow will use this arrangement.

The main factors that affect the capacity of waiting rooms are as follows:

- (1) The highest number of passengers gathered, which is the key to the waiting room waiting capacity.
- (2) The area of the waiting room.
- (3) The average occupancy area of each passenger.
- (4) The average time of staying in the waiting room of each passenger.

(3) and (4) are related not only related to the nature of the passenger flow, the peak period, and the improvement of the passenger's facilities, but also related to

whether the layout of the premises is reasonable and whether the ticket gate is efficient or not. So according to the specific circumstances of the station to get waiting room capacity is the method we should take.

The capacity of waiting room:

$$N = H \tag{1}$$

$$N = \sum n = \sum_{i=1}^n H\alpha_i \tag{2}$$

- N, n The capacity of waiting room that required in unit time;
- H The maximum number of passengers gathered at the station;
- α_i The number of passengers in each waiting room accounts for the maximum number of people gathered at the station.

The equality of capacity of waiting room now available:

$$N = \frac{M}{W} \tag{3}$$

$$N = \sum n = \sum \frac{m}{w} \tag{4}$$

- N, n The ability of waiting room, waiting room that required in unit time now available;
- M, m The area of waiting room;
- W, w The index of area of waiting room. The following table values are recommended.

Usually, the index of area for individuals in each waiting room is as follows (Table 1).

3.3 The Analysis of Crossline Equipment Capacity

Crossline equipment is the connection between the platform and the path. For the accurate calculation of the capacity, the device must be a one-way operation, and not allowed to have the opposite cross flow.

The main factors affecting the ability to crossline equipment are as follows:

- (1) The width of a walkway. The ability for crossline equipment is usually seen as the sum of the ability of several walking channels, and the width of each

Table 1 The index of area for individuals in each waiting room

Normal	Group	Soft seat	Mother and infant	Solider	Disabled	VIP
1.1	1.1	2.0	2.0	1.1	1.3	3.0

walkway and the passenger status (such as whether the empty hand, with no objects, push things, etc.) are closely linked;

- (2) Passenger travel speed. Passenger arrival and departure speed are not exactly the same, the passengers who arrive at the station are often eager to get on the train, and because passengers from all directions rushed to the station so its density of crowd is small, and the stride of passengers who departure from station is more casual, slower.
- (3) The area be taken by the passenger in the walkway. At the time of the station, the crowd gathered from the various directions to the station, and the distance between the passengers to passengers was different.

It is noteworthy that according to the terrain the travel speed of passengers on the crossline equipment to be divided into several cases to discuss. For example, in the flat or on the slope, the passenger speed has a significant difference.

Crossline equipment capacity calculation method:

$$N = \frac{A}{t} \quad (5)$$

- N Crossline equipment required capacity per minute;
 T The time of arrival or departure of a train of passenger crossing the line equipment, min;
 A Within t minutes, the number of passengers of one train enters into the line equipment.

Crossline equipment existing capacity:

$$N = \sum \frac{B}{b} \cdot n \quad (6)$$

- N Crossline equipment available capacity per minute;
 B The width of crossline equipment;
 b Width of a walkway;
 n Passenger passing capacity of a pedestrian passage per minute.

3.4 The Analysis of the Capacity of Entrance and Exit

Table 2 take manual selling as a prerequisite; this table represents the status of most of the passenger station ticket-check ability [4].

Calculation the ability of exits and entrances

Table 2 The index of the ticket checking ability of entrance and exit

Category	Unit	Capacity range	Median value of capacity
Ticket-check ability of entrance	Person/minute	30–40	35
Ticket-check ability of exit	Person/minute	32–42	37
Free-pass ability of exit	Person/minute	45–50	47

$$N = \frac{A}{t} \quad (7)$$

N The ability of exits and entrances required throughput per minute;

t The time standard of release a train of passenger;

A The number of passengers in a train or trains during the t period.

The ability of exits and entrances at now:

$$N = m \cdot n \quad (8)$$

N The ability of exits and entrances at now per minute;

m The number of exits and entrances;

n The ability of each exits and entrances at now.

3.5 The Ability of Selling in Wicket

The ticket organization is related to selling method, and the degree of proficiency of the conductor, the level of the main equipment of the ticket window of the station, and the number of tickets for the purchase of passengers and the specification of the ticket work. In the way of manual ticketing, Table 3 [5] is basically in line with the reality.

The selling ability of booking office:

$$N = A \cdot \alpha \quad (9)$$

N The selling ability of booking office required per day;

A The highest number of passengers on the month;

α The percentage of ticket number accounted for the number of transportation.

The selling ability of booking office at now:

Table 3 The table of ticket selling ability in various types of ticket office

The category of wicket	Long-distance	Short-distance	Suburb	Mixture	Transfer-signature	Transfer-acceleration
The average ability of wicket	80-100	80-200	280-320	120-180	180-200	120-140

$$N = m \cdot n \cdot T \quad (10)$$

- N The selling ability of booking office at now;
 m The number of wicket;
 n The average selling ability of each wicket;
 T Daily selling hour.

4 Study on the Streamline Organization of Peak-Hour

4.1 Streamline Profile

In the station, passenger luggage movement and many distribution activities were produced by certain flow process and flow lines. The layout design of passenger transport equipment should be based on the relief of the streamline circulation, as far as possible to reduce the possibility of crossing each other. Reasonable organization can effectively alleviate the pressure of streamline organization, accelerate the circulation efficiency of passengers. Passenger organizations need to study the characteristics of various types of streamlines, to strengthen streamline management [6].

During the peak-hour, passengers were crowded together in station, which makes capacity of station up to a maximum level. In order to successfully complete the task of passenger transport, we need to summarize the rules of the various types of streamline. The flow line of the passenger station can be divided into the entrance streamline and the exit streamline, and the streamline can be divided into three types: the passenger streamline, the package flow line and the vehicle streamline.

4.2 The Categories of Streamline

The flow line of the passenger station can be divided into the entrance streamline and the exit streamline, and the streamline can be divided into three types: the passenger streamline, the package flow line and the vehicle streamline.

Entrance streamline:

Since the flow of people from all directions gathered, so the streamline of passenger is relatively random. The incoming passenger streamlines are divided into the following types by passenger type [7]:

- (1) Ordinary passenger streamline: This is a key component of the entrance streamline, which is characterized by a large number of passengers.
- (2) Special passenger streamline: because passengers often have a lot of inconveniences, in the large passenger station have maternal waiting room and special services for special passengers.

- (3) Transit passenger streamline: different passenger transit procedures in accordance with the length of transfer.
- (4) Suburban passenger streamline: the average passenger commute time is short, so it would not cause much pressure for the passenger organization.
- (5) VIP streamline: large train station in order to allow guests to have a good experience will be set up VIP room.

In the station streamline, such as passengers have bought tickets or have checked items, you can directly get on the train. This can simplify the passenger's entry procedures, save passengers time, but also reduce passenger interference. The key is to reduce the number of passengers at the highest gathering, greatly easing the load of the passenger station. So the way to strengthen the presale tickets and baggage and other business work is very helpful for passenger organizations.

Exit streamline:

Exit passenger streamline has a relatively density. Exit streamlines should be smooth and fast, passengers must immediately out of the station, and scattered, which requires the station to keep the channel smoothly. If not it may cause congestion, it is likely to have mutual interference between, greatly interfere with the normal conduct of passenger organizations.

Vehicle streamline:

The vehicle streamline refers to the streamlines of the bus vehicles, the streamlines of the postal vehicles, the streamline of the special vehicles, the streamline of the taxi vehicles and other social vehicles, as well as non-motor vehicles.

4.3 Streamline Organization Principles and the Basic Way of Passenger Flow Optimization

(1) Organizational principles

There should avoid interfere between the streamlines. We should separate the entrance streamline and exit streamline, it is worth noting that the package line should be separated from the postal streamline. In addition, in order to ensure the normal transport order, it must separate the arrival package line and the departure package line.

(2) Passenger flow line optimization

Passenger flow line optimization of passengers and passengers in the station, transit passengers, railway and other modes of transport in the flow line optimization, as well as the optimization of passenger guide signs, etc. Streamline organization optimization aims to shorten the passenger walking distance from the station to the train, the possibility of reducing the mutual interference between the various types of streamline, and to prevent clogging, streamline detour. And then to

provide convenience to other types of passengers. In order to facilitate the transfer of passengers and to save valuable time, it is necessary to achieve a reasonable passenger streamline design, the focus is to reduce passenger transit time. If not, the streamline design will appear a lot of interference. Streamline cross interference will lead to massive congestion, even easily lead to conflict between passengers, increase the passenger organization work difficulty. When make the passenger streamline design should pay attention to shorten the passenger walking time, and take into account all kinds of interference caused by streamline.

(3) The basic method to disperse passenger streamline

The easing of the traffic line of the passenger station is mainly divided into three basic ways: the flow line is uncrossed on the plane, the flow line is uncrossed in space, and the streamline is uncrossed in plane and space simultaneously [8]. The flow line is uncrossed on the plane refers to the station and the layout of a variety of passenger equipment make the streamline uncrossed.

The flow line is uncrossed on the plane refers to the entrance streamline to go the second floor, exit passenger streamline to go ground floor.

And the streamline is uncrossed in plane and space simultaneously is the mixture of the plane and the space. This combines the advantages of them. It is not only separate streams but also can shorten the streamline distance. So most of the passenger station take this method.

In addition to the use of the abovementioned three-dimensional methods, we also use the following ways to uncross the streamline:

Cross-out method. It means to remove the intersection of the streamline generated in peak-hour from station to prevent the streamline too close.

Physical cutting method. It means we separated variety of streamlines by physical means such as viaducts, tunnels and other methods to reduce the crosscut, you can also use the notice to guide the various types of flow direction to deal with crosscutting problems, such as the import and export we should set different guide board. There is a simple way is to directly use the fence to cut the flow.

Source control method. It means by controlling the flow of a variety of streamlines to ease the flow line cross. Specific measures can be taken: increase the size of presale tickets to reduce the proportion of passengers directly to the station ticket to ease the load on the station; in addition to the government can also forbid a variety of public transport to reach the square, reduce the idler flow.

The above methods have a good effect on the streamline optimization during the peak-hour, so most of the passenger stations have adopted these control methods.

5 Conclusion

This paper introduces the concept of peak passenger flow, and then discusses the streamline organization and passenger capacity of passenger terminal.

- (1) Analyze the composition and characteristics of the passenger streamline in peak-hour, and introduced the passenger organization in the peak-hour. Beside we conclude the various types of formulation for calculation of station equipment capacity.
- (2) Introduced the classification of various types of streamlines, and make a summary of different types of streamline. We make some methods for optimization of passenger streamline. And we analyze the main features of various streamlines

Acknowledgements This work was supported by the Major Program of the National Natural Science Foundation of China (grant number 71390332), and the Specialized Research Fund for the Doctoral Program of Higher Education (grant number 20130009110001).

References

1. Hao Y (2010) Railway transport organization. China Railway Press, Beijing (in Chinese)
2. Hua W (2004) the characteristics and measures of passenger flow in the Peak-hour. Railway Transport and Economy 10 (in Chinese)
3. Hui J (2007) Study on passenger flow organization in the peak-hour. South Jiaotong University graduate degree thesis (in Chinese)
4. Peng MZ (2014) Railway passenger transport organization. Southwest Jiaotong University Press (in Chinese)
5. Ye H (2005) Passenger transport organization. Southwest Jiaotong University Press (in Chinese)
6. Liu Z, Zhao Y (2010) Transportation marketing. China Railway Publishing House (in Chinese)
7. Wang R (2004) Optimization of grassroots passenger on the. Railway Trans Econ 26(2) (in Chinese)
8. Wu X (2014) Organization of peak passenger flow in railway intermediate station. Shanghai Railway Science and technology, Shanghai (in Chinese)

Bayesian Network Inference on Departure Time Choice Behavior

Xian Li, Haiying Li, Linqiao Qin and Xinyue Xu

Abstract Departure time choice behavior plays an important role in travel decision for metro passengers during morning peak hours. Different from statistical models, this paper proposed Bayesian networks (BNs) to model the departure time choices of metro passengers. The structure of BNs is learned through K2 algorithm and its parameters are estimated by maximum likelihood estimation (MLE) method using the stated preference (SP) survey data. Main results are obtained as follows: (1) passengers can accept departure earlier than usual in the range of 0–20 min; (2) passengers will prefer to choose departure earlier if they enjoy a 20% or more discount on metro fare; and (3) passengers are willing to depart at usual time with slight crowding while they prefer to depart earlier under serious crowding. These findings contribute to making strategies for passenger flow control and safety operation for metro stations.

Keywords Departure time choice · Bayesian network · Peak-hour congestion
Metro passengers

1 Introduction

Passenger congestion has caused a safety operation problem during morning peak hours in metro stations. Capacity improvement is indeed an effective way to solve the problem. However, it is an impossible task at present because of its long

X. Li (✉) · H. Li · X. Xu
State Key Laboratory of Rail Traffic Control & Safety,
Beijing Jiaotong University, Beijing, China
e-mail: 15120846@bjtu.edu.cn

X. Li
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

L. Qin
Traffic Operations and Safety (TOPS) Laboratory, Department of Civil and Environmental
Engineering, University of Wisconsin-Madison, 53706 Madison, WI, USA

construction period, high cost, and physical restrictions of stations [1]. An alternative way is to motivate passengers to change their departure time and ultimately decrease the travel demand during peak hours [2]. A lot of studies have analyzed departure time choice behavior based on discrete choice model such as multinomial logit (MNL) [3], nested logit (NL) [4], mixed logit (ML) [5], and probit model [6]. Bajwa et al. [5] analyzed the departure time choice of car and rail commuters using NL, cross-nested logit (CNL), and ML model, which concluded that accounting for the random taste heterogeneity in the model can improve its performance. Habib et al. [3] analyzed the joint choice of trip timing and mode choice based on MNL, using the 2001 travel survey data in Greater Toronto Area. The results indicated that the estimated model can capture the peak spreading phenomenon and be applied in the future. Thorhauge et al. [2] accounted for the Theory of Planned Behaviour and analyzed the departure time choice of drivers and public transport commuters based on the structural equation model and MNL, which drew a conclusion that fixed starting times at work has a strong effect on departure time choice.

Another approach to analyze travel choice behavior is machine learning such as decision tree models, neural network models, and BN. Machine learning focuses on connection of data, instead of the constraints with linear structure by discrete choice model [7]. One of the major advantages for BN is the intuitive and graphical representation of the causal relationships between data, which allows better understanding [8]. Thus, the BN is employed to model passengers' departure time choices based on SP survey data in this study. BN is a factorization of a probability distribution along a directed acyclic graph, which has been successfully applied in the field of artificial intelligence, reliability theory, system security, and bioinformatics.

This article is organized as follows: Sect. 2 presents a brief introduction on BN; Sect. 3 provides a new modeling approach that utilizes BN for departure time choice inference of metro passengers, where the structure and parameter learning of BN are dynamically determined. Detailed discussion on metro passengers' preferences under different attribute levels is presented in Sect. 4. Then conclusions are drawn in Sect. 5, as well as future research.

2 Theoretical Background of Bayesian Network

This section presents a brief introduction about BN. BN is one of the members for probabilistic graphical models and belongs to the directed graphical model. A usual definition of BN is $BN = (X, G, P)$, where

- X is a set of discrete variables with finite different values.
- $G = (V, E)$ is an acyclic and directed graph with a set of nodes V and a set of edges E .
- P is the set of conditional probability distributions corresponding to each variable in set X .

Let x_1, x_2, \dots, x_n be an enumeration of all the nodes in X and let π_i be the set of parents of a node x_i , the joint probability distribution of BN can be expressed as follows:

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \pi_i) \tag{1}$$

where $P(x_i | \pi_i)$ represents the conditional probability distribution of node i .

BN learning includes two basic aspects: structure learning and parameter learning. K2 algorithm is a greedy search algorithm which can rapidly reconstruct a moderately complex belief network using a small number of cases. This algorithm is used for structure learning, whose steps are given as follows:

- Step 1 Generate an ordering set with n nodes by assumptions;
- Step 2 Input a database D containing m cases;
- Step 3 Traverse from node 1 to node n and determine the parent node set of each node according to $P(i, \pi_i)$:
 - Step 3.1 Let maximum number of the parent node be u ;
 - Step 3.2 Set $\pi_i = \emptyset$ and search for the probable parent node for node i ($1 < i \leq n$);
 - Step 3.3 If $|\pi_i| < u$, select one node before node i into π_i and calculate the $P_{\text{new}}(i, \pi_i)$ by formula (2);
 - Step 3.4 If $P_{\text{new}}(i, \pi_i) > P_{\text{old}}(i, \pi_i)$, set $P_{\text{old}}(i, \pi_i) = P_{\text{new}}(i, \pi_i)$ and repeat Step 3.2–Step 3.3; otherwise, search for next node ($i + 1$).
- Step 4 Output the parent node set of each node.

$$P(i, \pi_i) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \tag{2}$$

where N_{ijk} is the number of cases in D where node i has r_i values, $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$; q_i is the number of parent for node i .

On the other hand, MLE is used for parameter learning, which is calculated as follows:

$$L(x_i | \pi_i) = \sum_{i=1}^n \log[P(x_i | \pi_i)] \tag{3}$$

3 Bayesian Network Departure Time Choice Inference

3.1 Data Collection and Variable Definition

The sample data used in this paper was collected through SP survey in Beijing metro stations from November to December in 2016. In the stated preference experiment, four factors including metro fare, departure time change, travel time saving, and crowding of metro stations were selected as scenario attributes analyzed by the Technique for Order Performance by Similarity to Ideal Solution method. One of the scenario choice sets in the survey questionnaire is illustrated in Fig. 1. The images in Table 1 were obtained from passengers flow survey from April to June in 2016 approved by Beijing Subway.

Informed consent was obtained from 522 participants approved by ethics committee of Beijing Jiaotong University. Finally, a total number of 1044 observations were collected, as well as the socioeconomic attributes of passengers such as gender and age. In order to verify the effectiveness and representativeness of the data collected, the gender and age distributions are compared with the 2010 census data as shown in Fig. 2. The variables used in this study are defined in Table 1.

3.2 Structure Learning and Parameter Learning of BN

There are five nodes in the structure learning result, including departure time change, travel time saving, metro fare, crowding, and departure time choice. By the BNT toolbox, the maximum number of parent node is set as two for each node, and adopt K2 algorithm to structure learning. The structure learning result of BN is given in Fig. 3, which shows that departure time choice is dependent on departure time change, metro fare, and crowding. Nonetheless, travel time saving has an indirect connection with departure time choice through crowding.

As for parameter learning, MLE is employed to calculate the posterior probability distribution. Further, receiver operating characteristic (ROC) curve is used to compare BN with MNL based on a small number of observations. Because the false positive and true positive rates of ROC curves are insensitive to the class

Fig. 1 A scenario choice set in survey questionnaire






Choice(ν)	Metro fare	Travel time saving	Crowding
Departing at usual time <input type="checkbox"/>	Full fare		
Departing 20min earlier than usual <input type="checkbox"/>	30% discount	20% less	
Departing 20min later than usual <input type="checkbox"/>	20% discount	10% less	

Table 1 Variables description and definition

Description	Design code for levels of attributes			
Travel time saving	1 for full fare; 2 for 10%; 3 for 15%; 4 for 20%			
Metro fare	1 for full fare; 2 for 10% discount 3 for 20% discount; 4 for 30% discount			
Departure time change	1 for departure at usual time 2/-2 for departure 15 min earlier/later than usual 3/-3 for departure 20 min earlier/later than usual 4/-4 for departure 25 min earlier/later than usual			
Crowding	1 for		2 for	
	3 for		4 for	
Departure time choice	1 for departure at usual time 2 for departure earlier than usual 3 for departure later than usual			

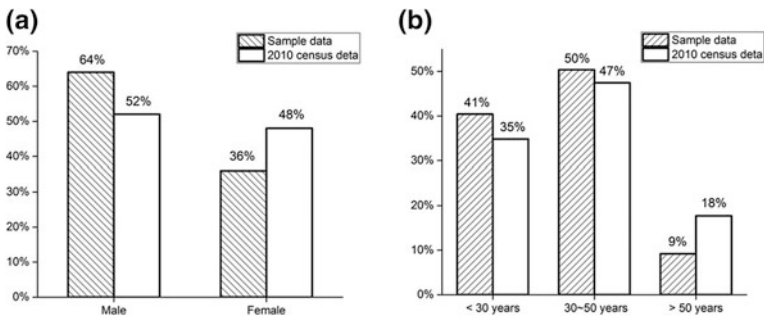


Fig. 2 Socioeconomic variable distribution: **a** gender and **b** age

Fig. 3 Structure learning result of BN

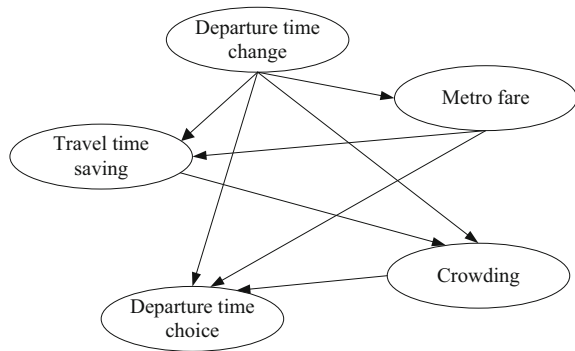
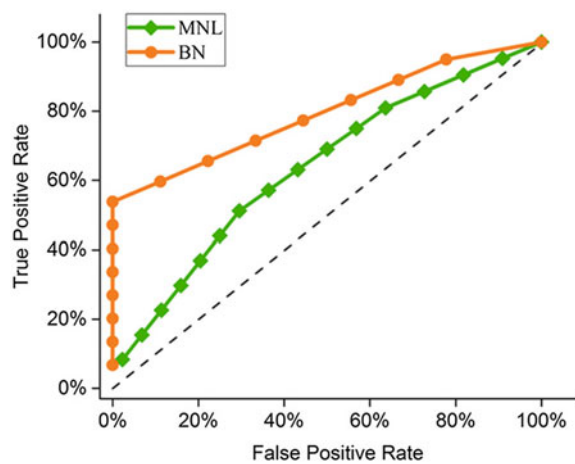


Fig. 4 Prediction comparison between BN and MNL



distribution, it is a proper measure for evaluating predictive models. The area under the ROC curve ranges from 0.5 to 1, with larger values representing higher system performance. The result shows predicting precision of BN is better than MNL as shown in Fig. 4.

4 Result Analysis

This section analyzes metro passengers’ preferences for choosing departure time under different attribute levels. The conditional probability distribution is calculated by formula (1) on the basis of BN structure.

The conditional probability distribution of departure time choice is illustrated in Fig. 5. An obvious characteristic is that passengers prefer to choose departure earlier or at usual time instead of departure later than usual, which may be the

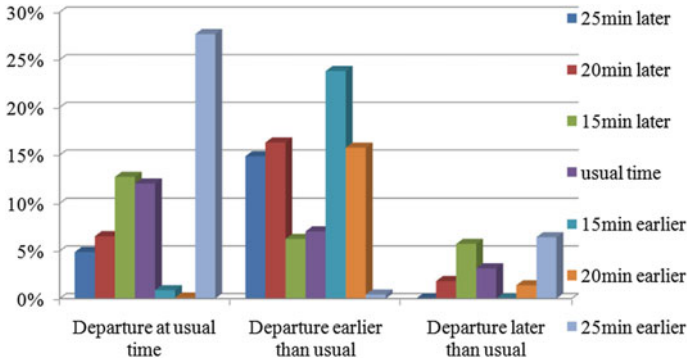


Fig. 5 Conditional probability distribution of departure time choice for departure time change

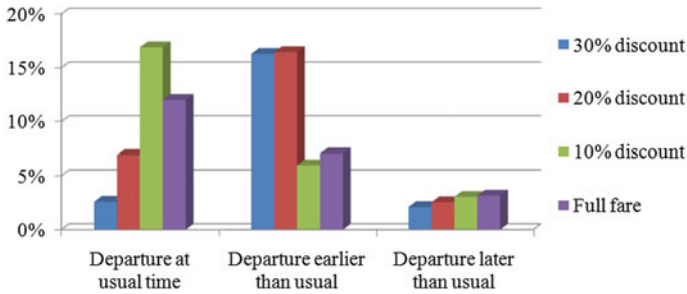


Fig. 6 Conditional probability distribution of departure time choice for metro fare

reason that most passengers are restricted by the working or school time. When passengers face a scenario with departure 25 min earlier, they tend to choose departure at usual time compared with departure 15 or 20 min earlier. A possible explanation is that passengers do not own much time, and departure earlier means that they have to pay more time cost.

In conclusion, most passengers accept departure earlier no more than 20 min and rarely choose departure later than usual.

In terms of metro fare, its conditional probability distribution of departure time choice is illustrated in Fig. 6. When passengers are confronted with 20 or 30% discount on metro fare, they tend to choose departure earlier than usual. But if the discount is less than 10%, passengers are likely to departure at usual time. Therefore, if there is a discount more than 20% on metro fare, some passengers will change their departure time, decrease travel demand during morning peak hours.

As for crowding, its conditional probability distribution of departure time choice is shown in Fig. 7. The result indicates that passengers prefer to choose depart at usual time when the crowding is not very serious. This is reasonable because they

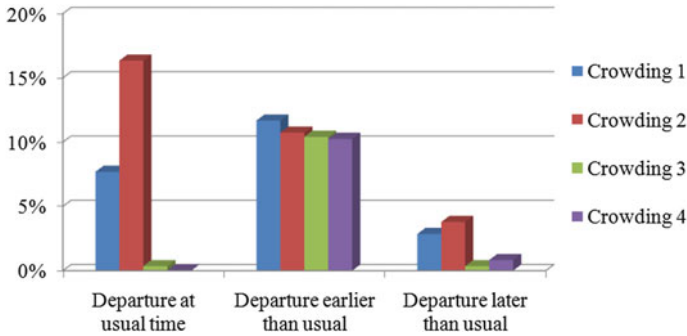


Fig. 7 Conditional probability distribution of departure time choice for crowding

may have a chance of getting a seat. Instead, when the crowding becomes serious, passengers incline to departure earlier than usual.

In short, if there is a discount above 20% on metro fare before peak hours, it will be useful to decrease passengers' travel demand and release congestion in metro stations.

5 Conclusion

In this study, BN is proposed to analyze the departure time choice of metro passengers. K2 algorithm and MLE are, respectively, used for structure and parameter learning of BN, with the help of the SP data collected in metro stations. Finally, passengers' preference toward departure time choice under different attribute levels is discussed.

Main findings of this study are summarized as follows:

1. Passengers can suffer from departure earlier than usual in the range of 0 to 20 min; on the contrary, passengers are unwilling to departure later than usual.
2. Passengers will prefer to choose departure earlier if they enjoy a 20% or more discount on metro fare.
3. Passengers are willing to departure at usual time with slight crowding while they prefer to departure earlier under serious crowding.

The results of this paper can be applied to make strategies for passenger flow control and safety operation in metro stations. In the future studies, other factors including weather and incidents will be considered further.

Acknowledgements The authors gratefully acknowledge the support provided by the national key research project "Key techniques of medium-speed maglev transportation system" (Grant No.2016YFB1200601) in China.

References

1. Xu XY, Liu J, Li HY, Jiang M (2016) Capacity-oriented passenger flow control under uncertain demand: algorithm development and real-world case study. *Trans Res Part E Logist Trans Rev* 87:130–148
2. Thorhauge M, Haustein S, Cherchi E (2016) Accounting for the theory of planned behaviour in departure time choice. *Trans Res Part F Traffic Psychol Behav* 38:94–105
3. Habib KMN, Day N, Miller EJ (2009) An investigation of commuting trip timing and mode choice in the greater toronto area: application of a joint discrete-continuous model. *Trans Res Part A Policy Pract* 43(7):639–653
4. Hess S, Daly A, Rohr C, Hyman G (2007) On the development of time period and mode choice models for use in large scale modelling forecasting systems. *Trans Res Part A Policy Pract* 41(9):802–826
5. Bajwa SU, Bekhor S, Kuwahara M, Chung E (2008) Discrete choice modeling of combined mode and departure time. *Transportmetrica* 4(2):155–177
6. Jou R (2001) Modeling the impact of pre-trip information on commuter departure time and route choice. *Trans Res Part B Methodol* 35(10):887–902
7. Zhu Z, Chen X, Xiong C, Zhang L (2017) A mixed bayesian network for two-dimensional decision modeling of departure time and mode choice. *Transportation* 1–24
8. Águila IMD, Sagrado JD (2016) Bayesian networks for enhancement of requirements engineering: a literature review. *Requirements Eng* 21(4):461–480
9. Cooper GF, Herskovits E (1992) A bayesian method for the induction of probabilistic networks from data. *Mach Learn* 9(4):309–347

Metro Station Facility Layout Optimization Model Based on the Level of Service Reliability

Qingwen Jia, Jianrui Miao and Xu Chen

Abstract Facility layout in metro stations is closely related to the facility utilization, level of service (LOS), and even the pedestrian safety. Aiming at this problem, the concept of reliability is incorporated, and followed by the definition of LOS reliability of single facility, single streamline, and whole station. Among these, pedestrian flow arrival characteristic and facility heterogeneity are considered. Then a facility layout optimization model based on the LOS reliability is proposed, and the decision tree cycle algorithm combining simulation is used to solve the problem. Finally, a case study of Zhichunlu Station in Beijing metro is presented to validate the effectiveness of the proposed model. Compared with reality, the optimization scheme can improve the performance of the station significantly including its operation efficiency and security.

Keywords Reliability · Level of service · Facility layout optimization
Metro station

1 Introduction

Existing metro stations are suffering from low service reliabilities due to the uneven passenger demand during the daily operation [1, 2]. It is urgent to solve this problem by practical strategies such as the facility layout optimization. The core of the facility layout problem (FLP) in metro stations is finding a configuration of pedestrian facilities to optimize specific goals within given constraints [3], including the best number of each facility and the corresponding physical parameters.

Q. Jia · J. Miao · X. Chen (✉)
State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University,
Beijing 100044, China
e-mail: xuchen@bjtu.edu.cn

Q. Jia · X. Chen
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_62

Existing related studies are divided into mathematical and simulation method. On the one hand, mathematical methods will be complicated and inefficient under the complex situations with more constraints [4], such as the LOS restriction of each streamline during rush hours. On the other hand, simulation methodology has become a popular method for providing visual pedestrian movement process and addressing different scenarios efficiently [5]. However, seldom simulation-based model focuses on the reliability of the pedestrian facility, which is directly related to the security of metro stations.

Fortunately, many researchers incorporate the concept of reliability into the road traffic system of transportation field [6]. Existing researches mainly focus on the reliability of the road traffic system from the macroscopic point of view [7]. However, seldom attention has been paid to FLP considering reliability in metro stations from a microscopic view.

This paper first introduces the concept of reliability into the LOS evaluation of pedestrian facility in metro stations, and defines the LOS reliability of single facility, single streamline and whole station on the consideration of pedestrian flow arrival characteristic and facility heterogeneity. Then a facility layout optimization model with the objective of maximal LOS reliability is proposed, and the decision tree cycle algorithm is used to solve the problem and provide possible optimization schemes. Finally, a case study of Zhichunlu Station is presented to illustrate the effectiveness of the proposed model. The main contributions of this paper are as follows: (1) the incorporation and definition of LOS reliability in metro stations; (2) specific facility layout optimization model and its solving algorithm.

2 Definition of the LOS Reliability

In this section, LOS reliability of the single facility, single streamline, and whole station is defined in sequence.

2.1 LOS Reliability of the Single Facility

According to the definition of reliability [8], LOS reliability of the single facility is defined as the probability of its LOS achieving the expected level in a given time period, which is computed as

$$R(x) = P(M < Q) = P(M - Q < 0) = \frac{N_{M-Q < 0}^x}{N_{\text{total}}^x}, \quad (1)$$

where $R(x)$ is LOS reliability of the facility x , M is the actual LOS evaluation index of the facility x , Q indicates the upper bound of the expected LOS of the facility, $N_{M-Q < 0}^x$ denotes the occurring frequency of the event $M - Q < 0$ in a given time

Table 1 Detailed index and level classification of facilities in metro stations

Facility type	Facility name	Index M	A	B	C	D	E
Service facility	SCM	q_x	[0, 5]	(5, 10]	(10, 20]	(20, 30]	(30, ∞)
	TVM		[0, 3]	(3, 5]	(5, 10]	(10, 20]	(20, ∞)
	AFG		[0, 2]	(2, 4]	(4, 7]	(7, 10]	(10, ∞)
Walkway	Corridor	d_x	[0, 0.43]	(0.43, 0.71]	(0.71, 1.43]	(1.43, 2]	(2, ∞)
	Stair		[0, 0.71]	(0.71, 1.11]	(1.11, 1.83]	(1.83, 2.5]	(2.5, ∞)
	Escalator		[0, 0.62]	(0.62, 1.24]	(1.24, 1.86]	(1.86, 2.46]	(2.46, ∞)

Note SCM means security checking machine; TVM denotes ticket vending machine; AFG represents automatic fare gate; q_x denotes average queue length (person); d_x represents pedestrian flow density (person/m²)

period, and N_{total}^x is the total number of the arrival pedestrian at facility x during the given time period.

According to [9], pedestrian facilities can be classified into service facilities X_s and walkways X_w . A detailed index and its level classification of each kind of facility are shown in Table 1.

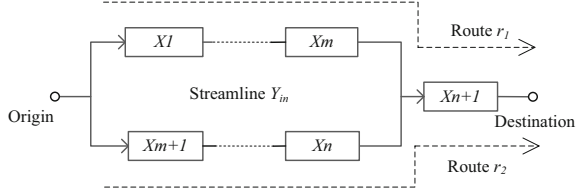
2.2 LOS Reliability of the Single Streamline

LOS reliability of the single pedestrian streamline is the probability that LOS of pedestrian streamline reaches the expected LOS during the given time period. Although streamlines can be divided into inbound Y_{in} , outbound Y_{out} , and transfer Y_{tr} streamlines according to the origin and destination of pedestrians, they have the same component (i.e., origin point, destination point, series-connected routes and parallel-connected routes). In Fig. 1, an example of the inbound streamline is presented. Here, $\phi(l)$ represents the route set of each streamline l , and $F(r)$ represents the facility set of each route $r \in \phi(l)$. For the series- and parallel-connected facilities, the LOS reliability is calculated according to formula (2) and (3), respectively [8]. Then, the final LOS reliability can be calculated using formula (4).

$$RSS(l) = \prod_{X \in F(r)} R(X) \tag{2}$$

$$RSS(l) = 1 - \prod_{X \in F(r)} (1 - R(X)) \tag{3}$$

Fig. 1 Inbound streamline in metro stations



$$RSS(Y_{in}) = \left\{ 1 - \left[1 - \prod_{X \in F(r_1) \setminus X_{n+1}} R(X) \right] \cdot \left[1 - \prod_{X \in F(r_2) \setminus X_{n+1}} R(X) \right] \right\} \cdot R(X_{n+1}) \tag{4}$$

2.3 LOS Reliability of the Whole Station

LOS reliability of the whole station is the probability that LOS of the station reaches the expected LOS during the given time period, which can be calculated according to the LOS reliability of each streamline. Considering the difference between the different kinds of streamlines, three weight parameters (i.e., γ_{in} , γ_{out} , and γ_{tr} , are related to the percentage of each kind of pedestrian flow) are proposed. Then, the LOS reliability of the whole station $RS(S)$ can be calculated as follows:

$$RS(S) = \gamma_{in} \cdot RSS(Y_{in}) + \gamma_{out} \cdot RSS(Y_{out}) + \gamma_{tr} \cdot RSS(Y_{tr}). \tag{5}$$

3 Modeling and Solving Algorithm

3.1 Facility Layout Optimization Model

On the basis of the definition of LOS reliability, a facility layout optimization model is proposed, and the main idea is maximizing the LOS reliability of the whole station under the constraint of that each streamline l reaches the expected LOS (i.e., C-level) [9]. According to the optimal reliability theory [8], both the number of service facilities ($\mathbf{c} = \{c_x | x \in X_s\}$) and width of walkways ($\mathbf{w} = \{w_y | y \in X_w\}$) are regarded as decision variables. The detailed facility layout optimization model is formulated as follows:

$$\text{Max } RS(S) = \gamma_{in} \cdot RSS(Y_{in}) + \gamma_{out} \cdot RSS(Y_{out}) + \gamma_{tr} \cdot RSS(Y_{tr}) \tag{6}$$

$$\text{s.t. } RSS(l) \geq \delta_l, l \in Y_{in} \cup Y_{out} \cup Y_{tr} \tag{7}$$

$$c_x^{low} \leq c_x \leq c_x^{upper}, \forall x \in X_s \tag{8}$$

$$w_y^{low} \leq w_y \leq w_y^{upper}, \forall y \in X_w, \tag{9}$$

where δ_l is the threshold of LOS reliability of each streamline l ; $RSS(l)$ is calculated by formula (2)–(4) for the streamline l ; c_x^{low} and c_x^{upper} is the lower and upper bound of the number of the service facility x , respectively; w_y^{low} and w_y^{upper} represents the lower and upper bound of the width of walkway y , respectively.

3.2 Solving Algorithm

Due to the nonlinear constraints of formula (7), it is difficult to solve the problem with a numerical algorithm. Therefore, the decision tree cycle algorithm combining simulation is proposed. The detailed process is as follows (see Fig. 2):

Step 1: Facility assignment

The existing facility layout (c^0, w^0) is regarded as the initial assignment scheme, which should satisfy the total number constraint and walkway width constraint [i.e., formula (8)–(9)]. Here, Z^0 is the value of LOS reliability for the initial scheme (c^0, w^0) ; $\theta_i = (Z^i - Z^{i-1})/Z^{i-1}, i > 0$ is the rate of change of LOS reliability, i is the cycle time, and its initial value is 0.

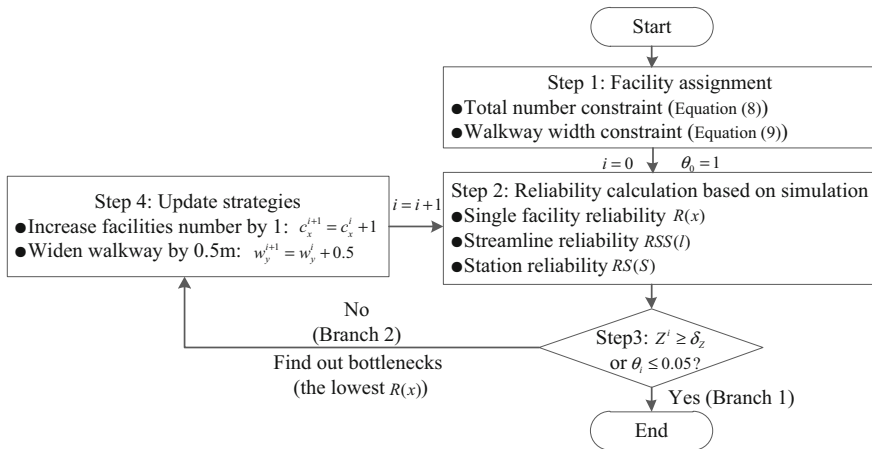


Fig. 2 Solving process of the problem

Step 2: Reliability calculation

Run the station simulation with AnyLogic and calculate the LOS reliability Z^i according to formula (2)–(5).

Step 3: Optimization process

Two termination conditions are considered in this process, including the LOS reliability Z^i which reaches the anticipated level ($\delta_z = 0.9$) and the rate of change θ_i is less than 0.05. Then, a decision tree cycle algorithm is applied to the optimization process. Each decision point leads to two branches:

Branch 1: if any one condition is reached, the algorithm terminates.

Branch 2: if not, there must be a bottleneck facility set B_i , $B_i \subset X_s \cup X_w$, that is, the aggregation of the facilities with lower LOS reliability.

Step 4: Update strategy

Step 4.1: If there is only one kind of facility in the bottleneck set B_i , two facilities x_1 and x_2 are randomly selected from B_i , and set $c_{x_1}^{i+1} = c_{x_1}^i + 1$, $c_{x_2}^{i+1} = c_{x_2}^i + 1$ (or $w_{x_1}^{i+1} = w_{x_1}^i + 0.5$, $w_{x_2}^{i+1} = w_{x_2}^i + 0.5$). Then a new strategy \mathbf{c}^{i+1} (or \mathbf{w}^{i+1}) is generated, and go to Step 2 after setting $i = i + 1$.

Step 4.2: If not, one service facility x and one walkway y are randomly selected from the facility set B_i in turn; then, set $c_x^{i+1} = c_x^i + 1$ and $w_y^{i+1} = w_y^i + 0.5$. Finally, a new strategy $(\mathbf{c}^{i+1}, \mathbf{w}^{i+1})$ is generated, and go to Step 2 after setting $i = i + 1$.

4 Case Study

Zhichunlu Station hall in Beijing metro is regarded as the research object (see Fig. 3). There are seven origin/destination (OD) points that are represented as P1 to P7 and ten pedestrian facilities. Seven video recorders were fixed at these points between 8:30 and 9:00 am on March 18, 2016, and then 10,402 passengers' data was collected.

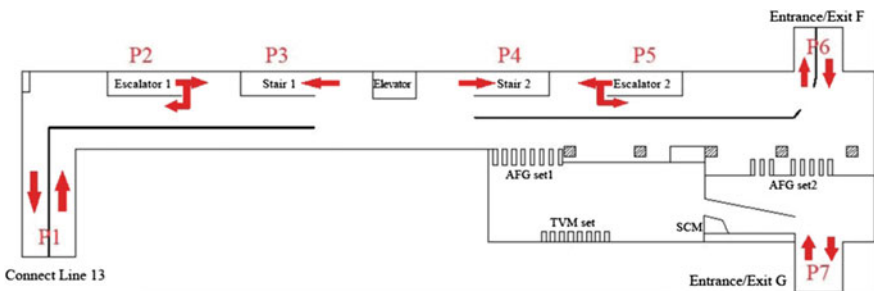


Fig. 3 Layout of the station hall of Line 10 at Zhichunlu Station in Beijing metro

4.1 Model Application

To calculate the LOS reliability of the station, a Monte Carlo simulation using the collected data was conducted by AnyLogic. The parameters $\gamma_{in} = 0.58$ and $\gamma_{out} = 0.42$ were selected, which are determined by the volume ratio of each kind of streamline. According to Sect. 3, LOS reliability of the station hall is lower than 0.9 when the expected LOS is at C-level, then Branch 2 of the decision node is selected. Moreover, the bottleneck set B_0 includes SCM, Stair-1, Stair-2, and Corridor-1 (marked by ellipses in Fig. 4). Then, according to Step 4.2, following measures are proposed (optimization scheme 1):

- (1) $c_x^1 = c_x^0 + 1$ ($x = \text{SCM}$), adding one SCM;
- (2) $w_y^1 = w_y^0 + 0.5$ ($y = \text{Corridor-1}$), increasing the width of Corridor-1 by 0.5 m.

Then the algorithm is jumped to Step 2. LOS reliability of the station hall is still lower than 0.9 due to the insufficient capacities of Stair-1 and Stair-2, which are included in bottleneck set B_1 . As there is no service facility in B_1 , Step 4.1 is conducted, and further optimization is needed (optimization scheme 2):

- (1) $w_{y_1}^2 = w_{y_1}^1 + 0.5$ ($y_1 = \text{Stair-1}$), increasing the width of Stair-1 by 0.5 m;
- (2) $w_{y_2}^2 = w_{y_2}^1 + 0.5$ ($y_2 = \text{Stair-2}$), increasing the width of Stair-2 by 0.5 m.

The scenario of optimization scheme 2 is shown in Fig. 5. The station hall LOS reliability reached to 0.9597 on the constraint of expected LOS, which is better than other scenarios (see Fig. 6). Therefore, the optimization goal is achieved, and the algorithm ends.

4.2 Result Analysis

To validate the effectiveness of the proposed model, station LOS reliabilities of the optimization scheme 1 and optimization scheme 2 are compared with that of original scheme in sequence.

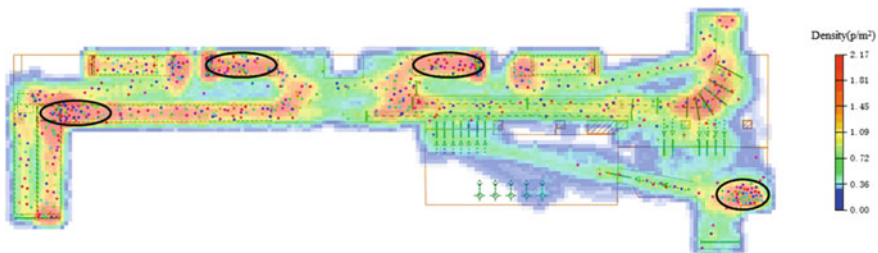


Fig. 4 Simulation scenario of original scheme

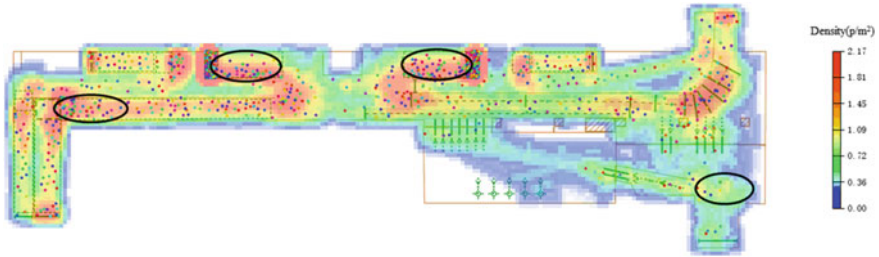
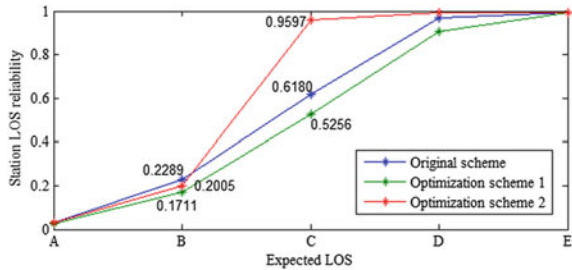


Fig. 5 Simulation scenario of optimization scheme 2

Fig. 6 LOS reliability of the station hall in three schemes



From Fig. 6, we can conclude that the station LOS reliability of the optimization scheme 1 is always smaller than that of the original scheme, which means the original facility layout is better. The main reason is that the optimization scheme neglects the bottlenecks of Stair-1 and Stair-2, resulting in these facilities becoming less reliable under the condition of increasing pedestrian flow volume. Therefore, the optimization scheme 2 is proposed later.

Similarly, we can get the conclusion that the station LOS reliability of the optimization scheme 2 is bigger than that of the original scheme in most cases except for the B-level, which means the facility optimization model and its solving algorithm are useful in most cases. With respect to the abnormal case (i.e. the B-level), detailed reasons are analyzed as follows: the width of Corridor-1 is increased, and meanwhile that of Corridor-2 is reduced, so the bottleneck is transferred from Corridor-1 to Corridor-2, resulting in the LOS reliability is also at a lower level. Aiming at this problem, the efficient solution is reducing the magnitude of the change of the width in each optimization step. Currently, the magnitude of the change is 0.5 m, and it should be adjusted in the future research.

5 Conclusions and Future Research

Aiming at FLP, the concept of reliability is incorporated, followed by the definition of LOS reliability of the single facility, single streamline, and the whole station. On the basis of LOS reliability definition, a facility layout optimization model is proposed with the objective of maximal station LOS reliability, followed by the decision tree cycle algorithm combining simulation. In the case study, two optimization schemes are proposed and compared with the original scheme. Overall, LOS reliability can accurately reflect the usage of the facilities and can be regarded as an important evaluation index, and the proposed algorithm can provide efficient optimization schemes.

Current research has laid a good foundation for the application of reliability in the field of pedestrian traffic. Future work is necessary to compare the performance of LOS reliability with that of other indexes such as the average time of mustering and evacuation.

Acknowledgements This work was financially supported by the national key research project in China under Grant No. 2016YFB1200601. The video data in the metro station is provided by the College of Basic Education, National University of Defense Technology, Master, Dong Liang.

References

1. Cats O, West J, Eliasson J (2016) A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. *Transp Res Part B: Methodological* 89:43–57
2. Xu XY, Liu J, Li HY et al (2014) Analysis of subway station capacity with the use of queuing theory. *Transp Res Part C: Emerg Technol* 38(1):28–43
3. Anjos MF, Vieira MVC (2017) Mathematical optimization approaches for facility layout problems: the state-of-the-art and future research directions. *Eur J Oper Res* 26(1):1–16
4. Hillier FS, Connors MM (1966) Quadratic assignment problem algorithms and the location of indivisible facilities. *Manage Sci* 13(1):42–57
5. Farhan J (2015) An agent-based multimodal simulation model for capacity planning of a cross-border transit facility. *Transp Res Part C: Emerg Technol* 60(4):189–210
6. Asakura Y (1999) Evaluation of network reliability using stochastic user equilibrium. *J Adv Transp* 33(2):147–158
7. Chen A, Yang H, Lo HK et al (2002) Capacity reliability of a road network: an assessment methodology and numerical results. *Transp Res Part B: Methodological* 36(3):225–252
8. Kuo W (2001) *Optimal reliability design: fundamentals and applications*. Cambridge University Press, Cambridge
9. Hunter-Zaworski K (2013) *Transit capacity and quality of service manual*. Washington, D.C.: Transportation Research Board (ed), Transit cooperative research program

Monte Carlo Simulation Method Used in Reliability Evaluation of Railway Overhead Contact Line

Jianfeng Xu, Yuan Zhong and Shibin Gao

Abstract This paper presents an application of Monte Carlo Simulation method for reliability evaluation of the railway overhead contact line system by establishing a repairable system model with four components placed in series. Weibull distribution is adopted to characterize the failure rate of system component. Multinomial distribution is adopted to characterize the occurrence probability of different failure mode of system component. The parameters of probability distribution are obtained from failure data collected in the maintenance of actual overhead contact line system. The system reliability indices are calculated by Monte Carlo simulation method. The simulation result is acceptable compared to the result obtained from Markov analysis method.

Keywords Overhead contact line · Reliability · Monte Carlo simulation
Repairable system

1 Introduction

Electric traction has the function of safely transporting people and/or goods with the aid of the traction power supply system. The objective of the traction power supply system is to ensure uninterrupted, reliable, and safe operation of the electric traction vehicle.

Overhead contact line installations form the traction power feeding system and must provide uninterrupted traction at the pantographs of the traction vehicles. The overhead contact line installation is an outdoor large-scale distributed system, which interacts with traction vehicles and numerous environmental factors and is

J. Xu (✉) · Y. Zhong · S. Gao
Southwest Jiaotong University, 111, North 1st Section,
2nd Ring Road, Chengdu, Sichuan, China
e-mail: xu.jianfeng@my.swjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_63

not installed redundantly for economic and technical reasons; therefore, the required high reliability, high availability of the contact line system is a big challenge for maintenance operation [1, 2].

The following maintenance strategies are adopted by overhead contact line system maintainer of China railway to optimally assure its function:

- Corrective maintenance, i.e. replacement or repair of the failed component;
- Preventive(periodic) maintenance, i.e. regular inspections and possibly repair based on a structured maintenance plan;
- Condition-based maintenance, i.e. performance of the repair action upon detection of the degraded conditions of the system.

Therefore, overhead contact line system is a repairable system. The ability of overhead contact line system to fulfill its operational function can be characterized by the index of system availability. The instantaneous availability is defined as the probability that the system is operating at time t . For systems under periodic maintenance such as overhead contact line system, steady-state availability is not defined, instead, the average availability over a given period of time is introduced as an indicator of performance. It represents the expected proportion of time that the system is operating in the considered period of time.

There has been considerable interest in the method of evaluating the reliability and availability of overhead contact line system. Wan et al. [3] modeled the reliability of the overhead contact line system by fault tree analysis. Based on credibility theory, Yang et al. [4] evaluated the reliability of electrified railway catenary system by probabilistic fuzzy functions. Wan [5] and Wang [6] built a mathematical analysis model for the reliability of the catenary system by Markov analysis method and derived the resolution result of some important reliability indices. However, the Markov analysis method suffers from one major disadvantage. It is necessary to assume constant probabilities or rates for all occurrences (failures and repairs). It is also necessary to assume that events are statistically independent. These assumptions are hardly ever valid in real life.

Monte Carlo Simulation (MCS) method is a powerful modelling tool for the analysis of complex systems, due to its capability of achieving a closer adherence to reality. The advantage of the MCS approach comes from the fact that it allows taking into account, in a realistic manner, the many phenomena that can occur, without additional complications in the modelling and in the solution procedure [7].

In this paper, a repairable system with 4 components placed in series is adopted to model the reliability of overhead contact line system. Monte Carlo method is applied to simulate the state transition of the system from the starting time to the final time of interest. The simulation results show not only numerical system indices such as reliability, availability and average failure time but also the evolution history of the system life.

2 Reliability Model of the Overhead Contact Line System

2.1 Fault Tree Analysis of Overhead Contact Line System

Liu et al. [8] analyses the failure mode of the overhead contact line system via fault tree analysis (FTA) technique. FTA result of the overhead contact line system is shown in Fig. 1. The overhead contact line system is composed of four components: pole and foundation, support device, positioning device and suspension. Any component failure causes the top event (failure of overhead contact line system).

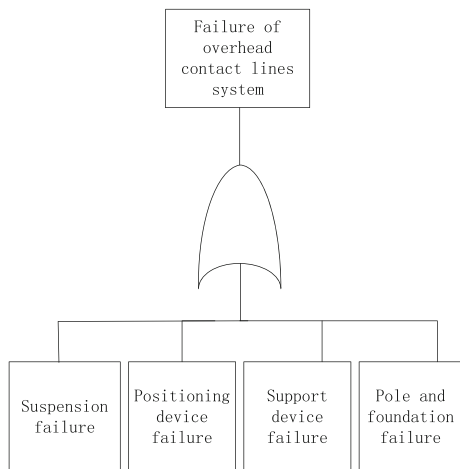
2.2 Failure Rate of System Components

Consider the overhead contact line system is composed of four independent components: pole and foundation, support device, positioning device and suspension. The system is represented by a four components series reliability model, i.e. the failure of any of the components will result in failure of the system. The failure rate of the overhead contact line system is the sum of the failure rate of four components, i.e.

$$\lambda_{\Sigma} = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4, \tag{1}$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ represent the failure rates of pole and foundation, support device, positioning device, and suspension component, respectively, and λ_{Σ} is the failure rate of the overhead contact line system.

Fig. 1 FTA result of overhead contact line system



Consider the two-parameter Weibull distribution function is most frequently used in metal material fatigue design and testing and the overhead contact line system is mostly composed by metal parts, the two-parameter Weibull distribution function is adopted to model the failure characters of the components of the overhead contact line system. The Weibull distribution with shape parameter and scale parameter has density function given by

$$f(t) = \frac{m}{\eta} \left(\frac{t}{\eta}\right)^{m-1} \exp\left[-\left(\frac{t}{\eta}\right)^m\right] \quad m \geq 0, \eta \geq 0. \quad (2)$$

The reliability of a Weibull distributed component is defined by

$$R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^m\right] \quad m \geq 0, \eta \geq 0, \quad (3)$$

where m is the shape parameter and η is the scale parameter of Weibull distribution. Each component of the overhead contact line system has its specific parameters of the Weibull distribution. The parameters are calculated according to the actual failure data of the overhead contact line system.

The failure rate of the Weibull distributed component is defined by

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{m}{\eta} \left(\frac{t}{\eta}\right)^{m-1} \quad m \geq 0, \eta \geq 0. \quad (4)$$

Then the failure rate of the overhead contact line system can be calculated by the sum of four components as follows:

$$\begin{aligned} \lambda_{\Sigma}(t) &= \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \\ &= \frac{m_1}{\eta_1} \left(\frac{t}{\eta_1}\right)^{m_1-1} + \frac{m_2}{\eta_2} \left(\frac{t}{\eta_2}\right)^{m_2-1} + \frac{m_3}{\eta_3} \left(\frac{t}{\eta_3}\right)^{m_3-1} + \frac{m_4}{\eta_4} \left(\frac{t}{\eta_4}\right)^{m_4-1}, \end{aligned} \quad (5)$$

where m_i is the shape parameter of i th component, η_i is scale parameter of i th component ($i = 1, 2, 3, 4$).

2.3 Repair Rate of System Component

The ability of a repairable system to fulfill its operational function depends not only on reliability but also on the maintainability of the system. The ease with which repairs and other maintenance work can be carried out determines a system's maintainability.

Consider the repair of the overhead contact line system is carried out mostly by replacement of the failed parts, the repair time of a failed component is defined as working hours of the maintenance staff used to replace the failed part. As shown in

Table 1 Failure mode summary of 2009–2011 in Changsha

Failure mode	Number of failure	Failure category	Avg. repair time (h)
Positioning tube shedding	8	Positioning device failure	1
Positioning tube rupture	6		1
Positioning wire clip failure	3		0.5
Insulator breakdown	22	Support device failure	3
Transition tube rupture	3		2
Insulator flashover	3		3
Positioning ring failure	3		3
Pole fracture	7	Pole and foundation failure	8
Structure tilt	4		8
Catenary breakage	8	Suspension failure	6
Catenary burned down	5		6
Contact line breakage	4		4
Dropper breakage	3		0.5
Wire clip fracture	4		0.5

Table 1, each component of the overhead contact line system has its own failure modes. Different failure modes need different repair time.

Assuming there are k failure modes for one component of the overhead contact line system and the probability of the component failed due to certain failure mode is multinomial distributed, i.e. if the probability of certain failure mode i occurs is p_i ($i = 1 \dots k$), then $p_1 + p_2 + \dots + p_k = 1$. The multinomial distribution of the failure modes of one component can be calculated according to the actual failure data of the overhead contact line system.

3 Monte Carlo Simulation (MCS) Method

The common task of the application of Monte Carlo Simulation (MCS) method in system reliability analysis is to calculate the probability of the events of interest, e.g. the system availability to function state and the mean of the random variable of interest such as mean time of repair.

Take the calculation of the probability of the events of interest, for example assuming the probability of the event A occurs $P(A)$ equals p (p is to be determined), p can be calculated by steps (1)–(3):

- (1) Assume in N times independent random sampling, the count of the event A of interest is n_A .

Let random variable x_i be the count of event A in i th sampling, i.e.

$$x_i = \begin{cases} 0 & \text{event } A \text{ occurs in } i\text{th sampling} \\ 1 & \text{event } A \text{ doesn't occur in } i\text{th sampling.} \end{cases} \quad (6)$$

- (2) Calculate the occurrence frequency of event A in N times independent random sampling f_N by

$$f_N = \frac{n_A}{N} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (7)$$

- (3) When N is sufficiently large, the estimation of the probability of event A occurs is

$$p = \frac{n_A}{N} = \frac{1}{N} \sum_{i=1}^1 x_i. \quad (8)$$

Random sampling on the computer is to generate random number obeying a cumulative distribution function of a specific probability distribution. Among all the distributions, the uniform distribution in the interval $[0,1)$, denoted as $U(0,1)$, plays a role of fundamental importance since sampling from this distribution allows obtaining random variable sample obeying any other distribution.

Let continuous function $F(x)$ be the cumulative distribution function of a random variable x and $r = F(x)$, then random variable r is uniform distributed in the interval $[0,1)$.

Let $F^{-1}(r)$ be the inverse transform of $F(x)$, we can obtain the value $x = F^{-1}(r)$ which obeys the cumulative distribution function $F(x)$, where r is sampled in uniform distribution in the interval $[0,1)$.

Using the sampling method of inverse transform discussed above, the Monte Carlo Simulation of reliability model of the overhead contact line system discussed in 1.2 is performed as follows:

- (1) Assume mission time of the simulation is T at the beginning the simulation time $t = 0$, the overhead contact system is in operational state
- (2) Sample the time at which a failure occurs for 4 components of the system obeying Weibull distribution; Let the minimal value of the 4 samples be the time t_f at which the system transits to failure state; determine the failed component i ; simulation time t increased by t_f .
- (3) Sample the multinomial distribution of the failure modes of the failed component i , determine the specific failure mode and corresponding repair time t_r , simulation time t increased by t_f .
- (4) Repeat step (2) and (3) until simulation time is reached to T .

4 Simulation Data and Result

4.1 Simulation Data

Liu et al. [8] summarizes the failure data of the overhead contact line system of 2009–2011 in Changsha.

The parameters of the Weibull distribution of four components of overhead contact line system is calculated according to failure data in Table 1. The results are listed in Table 2.

According to failure data in Table 1, we calculate the frequency of the failure occurrence. Then the multinomial distribution of the failure modes of the system components is calculated and corresponding repair time of the failure mode is allocated. The results are listed in Table 3.

The reliability of four components of overhead contact line system can be calculated by formula (3), as shown in Fig. 2.

4.2 Simulation Result

Monte Carlo simulation was performed in a mission time of 8760 h for 2000 times. The reliability indices of overhead contact line system obtained in simulation are shown in Table 4.

Table 2 Fitting parameters of Weibull distribution

System component	Shape parameter m	Scale parameter η
Pole and foundation	1.1984	2786.2
Support device	1.0943	942.93
Positioning device	1.2803	4205.9
suspension	1.019	1445.8

Table 3 Probability rate of failure and repair time

System component	Multinomial distribution	Repair time (h)
Pole and foundation	(1) (only one failure mode)	(8)
Support device	(0.71, 0.81, 0.9, 1)	(3, 2, 3, 3)
Positioning device	(0.82, 1)	(1, 0.5)
Suspension	(0.54, 0.71, 0.83, 1)	(6, 4, 0.5, 0.5)

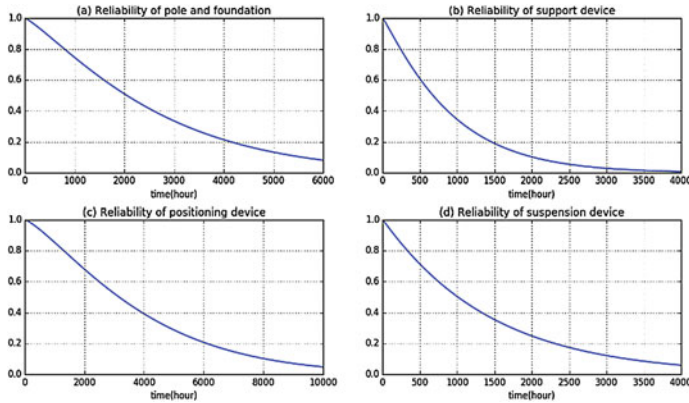


Fig. 2 Reliability of overhead contact line system components

Table 4 Reliability indices of overhead contact line system

Reliability index	Simulation result	Std. of result
Avg. availability	0.992	
Avg. failure times	19	4
Avg. unavailable time (h)	70	17

The average availability of the considered overhead contact line system is 0.992 which is comparable to the result obtained from Markov analysis method in [5, 6].

The system availability can be improved by optimizing the maintenance task and reducing repair time. For example, if the repair time of failure modes of system components is reduced as following: Pole and foundation: (6); Support device: (2, 1, 2, 2); Positioning device: (1, 0.5); Suspension: (4, 3, 0.5, 0.5), the simulation result shows that system availability is improved to 0.994.

The estimation of the instantaneous failure rate of the overhead contact line system is shown in Fig. 3, which is calculated by the frequency of the failure occurs within 24 h.

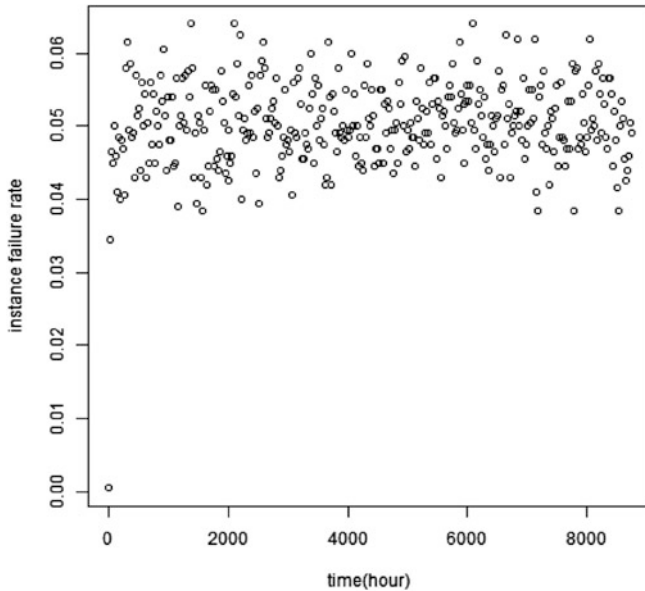


Fig. 3 Instantaneous failure rate of the system

5 Conclusion

In this paper, we proposed an application of Monte Carlo Simulation (MCS) method for reliability evaluation of the railway overhead contact line system. Comparing to Markov analysis method, MCS method does not need some simplifying assumptions of the system model, allows the adoption of the probability distribution of the failure rate or repair rate other than exponential distribution, thus can be applied in the modeling of complex engineering systems.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant (U1534209) and the China Railway Corporation Technology Research and Development Program under Grant (2015J008-A).

References

1. Wu JQ (2010) Pantograph and overhead contact line system. Southwest Jiaotong University Press, Chengdu (In Chinese)
2. Duque O, Zorita AL, García-Escudero LA, Fernández MA (2009) Criticality determination based on failure records for decision-making in the overhead contact line system. Proc Inst Mech Eng Part F J Rail Rapid Transit 223(5):485–494
3. Wan Y, Deng B, Li HJ et al (2005) Research on reliability of the catenary system based on FTA. J Railway Eng Soc 90(6):56–59 (In Chinese)

4. Yang Y, Wu JY, Wu Y et al (2008) Fuzzy reliability evaluation of electrified railway catenary system based on credibility theory. *J China Railway Soc* 30(6):115–119 (In Chinese)
5. Wan Y, Deng B, Li HJ et al (2006) Analysis of railway catenary system based on Markov theory. *J Appl Sci* 24(6):633–636 (In Chinese)
6. Wang SH, Zhao Q, Ma JD et al (2013) Markov model for analyzing catenary system reliability. *China Safety Sci J* 23(9):39–44 (In Chinese)
7. O'Connor P, Kleyner A (2012) *Practical reliability engineering*, 5th edn. Wiley, Hoboken
8. Liu H, Liu ZG, Jiang J (2016) Optimization research of preventive maintenance strategy on the catenary system based on reliability constraints. *J Mech Strength* 38(1):74–79 (In Chinese)

Optimization Study on Operation Plan of Intercity Trains Based on Minimization of Total Social Cost in Intercity Passenger Corridor

Yu-ling Ye, Yun-fei Zhou and Lu-qi Yang

Abstract By taking operating cost and transportation external cost into consideration, a bi-level programming model is established, with which the upper level aims at the minimum social cost and the lower level is modeled by traffic flow assignment in intercity passenger corridor. A Simulated Annealing (SA) algorithm is used in the solution of the bi-level model to figure out sharing rate of all kinds of transportation modes and operating plans of trains. With an example in the end, this paper contrasts and analyzes the variation trend of the minimum total social cost under different operating frequencies of intercity trains to conduct a sensitivity analysis on the operating plan.

Keywords Bi-level programming model · Operation plan · Total social cost

1 Introduction

Many scholars did researches on the railway operating plan. To explore how the factors affect operating plan, Iljoon [1] has studied the competitive relationship among multiple transport modes based on network model and verified it by an example of NW–SW corridor in Korea. Huang and Peng [2] established the allocation model of passenger trains flow for the condition of elastic demand using Logit model.

To make the benefits of enterprise maximized, Shi et al. [3] deemed that passenger flow would be affected by train's operating plan. Aimed at the minimized cost, Alberto et al. [4] established a mixed integer programming model through passenger maximization and operating cost minimization to study passenger train's operating plan. Goossens et al. [5] studied the problem of formulating operating plans for trains which had different dwell plans. From another angle, Li et al. [6]

Y. Ye (✉) · Y. Zhou · L. Yang

Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
Tongji University, NO. 4800 Cao An Road Jiading District, Shanghai, China
e-mail: yuling71@163.com

designed a series of evaluation indicators of transport efficiency, quality and service level, and then built a fuzzy optimization model of the multi-objective decision-making for the CTC (comprehensive transportation corridor) combined capacity.

At present, the study of intercity railway's operation plan and operation frequency chooses the maximum benefits as its objectives. So, in this paper, we choose the minimum social total cost as optimal objectives and then establishes a bi-level programming model to research the influences of the operation frequency of intercity railway on passenger flow distribution in intercity transport corridors.

2 The Passenger Volume Distribution Model

2.1 Virtual Transportation Network

The path of traffic modes in the transportation corridor is regarded as a transportation corridor virtual network which consisted of two basic elements, the nodes and the links section. In the network, the nodes represent the cities, and the links represent the corresponding modes of transportation between the cities.

Let $G = (N, A)$ be a transportation network, the symbols are defined as follows:

- N the set of network nodes which includes station nodes, trains departure nodes, and stop nodes;
- A the set of network arcs which includes Getting-on-the-Train arcs, Getting-off-the-Train arcs, Train-Running arcs, and Dwell arcs;
- a a link in the network, $a \in A$;
- W the set of all OD pairs in network;
- w OD pairs in the network, $w \in W$;

2.2 Traffic Assignment Model Based

Different transport modes in the corridor have their own routes, a route should be chosen when we choose a mode, we take travel cost function as power function form as follows:

$$c_k^w = f(q_k^w) = g(q_k^w) - V_k^w, \quad (1)$$

$$V_k^w = \theta_1 P_k^w + \theta_2 T_k^w + \theta_3 TW_k^w + \theta_4 SS_k^w + \theta_5 S_k^w. \quad (2)$$

The assumptions of the model are as follows:

- Do not consider the situation of transfer;

- Do not consider the capacity of intercity bus temporarily, also, it can adjust operation plan timely according to dynamic volume of passenger flow;
- Suppose that ordinary railway takes the “stop by stop” operation plan, and its capacity is able to meet passengers’ travel demand.

Based on user equilibrium principle, we establish the traffic assignment model of transport corridor as follows:

$$\min(q_k^w) = \sum_{w \in W} \sum_{k \in K} \int_0^{q_k^w} C_k^w dx \tag{3}$$

s.t.

$$\sum_k q_k^w = d^w \quad w \in W, \tag{4}$$

$$q_k^w \geq 0 \quad k \in K, \quad w \in W. \tag{5}$$

Constraint (4) means passenger flow in OD pairs is fixed. And in OD pairs, the total passenger flow of all kinds of transportation mode is equal to the passenger’s demand. Constraint (5) means passenger flow is non-negative.

2.3 The Impedance Functions

The impedance function of getting-on links includes the time of passengers entering the station and waiting for the train. The waiting time is related to train departure frequency, so we assume it is half the train departure interval.

$$T_{a1} = \alpha(t_{on}^a + wt). \tag{6}$$

The impedance function of train-running links includes the tickets price (depending on tariff rates and travel distance), running time, congestion tolls, and stop impedance.

$$T_{a2} = D_a P_l^a + \omega t_{run}^a (q_a / U_a)^\theta + \alpha(t_{run}^a + t_{stop}^a). \tag{7}$$

The impedance function of getting-off links includes the average departure time (including the time of lining up and out of the station).

$$T_{a2} = \alpha t_{off}^a \tag{8}$$

3 Bi-level Programming Model of Operation Plan

3.1 The Establishment of the Bi-level Programming Model

As the objective function is to pursue the minimum total social cost, the upper level programming model is shown as the follows:

$$\begin{aligned} \min Y = R_r + R_b + R_c + R_e = & \sum_{\beta} \sum_l f_l^{\beta} s_l^{\beta} b_l^{\beta} C_{rh} + \sum_w C_b s_b^w q_b^w / \delta_b \\ & + \sum_w q_c^w C_c s^w / \delta_c + \sum_w E_c q_c^w s^w / \delta_c + f_r s_r b_r C_r, \end{aligned} \tag{9}$$

s.t.

$$\sum_l \sum_{\beta} f_l^{\beta} \leq N_{hi} \quad i \in n, \tag{10}$$

$$\sum_l \sum_{\beta} f_l^{\beta} \leq N_{ei} \quad i \in n, \tag{11}$$

$$q_{e_i} = \sum_{e_i \in w} q_G^w \quad i \in m, \tag{12}$$

$$q_{e_i} \leq \sum_{\beta} \sum_{l \in L} f_l^{\beta} A_l^{\beta} \quad i \in m, \tag{13}$$

$$q_{\beta l}^{e_i} \leq f_l^{\beta} A_l^{\beta} \quad i \in m, \tag{14}$$

$$f_l^{\beta} \geq 0. \tag{15}$$

In this model,

Constraint (10) means the amount of trains departing from a station cannot exceed its departure capability N_{hi} .

Constraint (11) means the amount of trains passing a section cannot exceed its traffic capability N_{ei} .

Constraint (12) means the passenger flow of section e_i equals the sum of passenger flow between OD points which passes this section.

Constraint (13) means the sum of each section's passenger flow cannot be more than the total rated seats of trains passing this section.

Constraint (14) means the number of passengers of the intercity train which is of rank β and in route l cannot exceed its carrying capacity.

Constraint (15) is a non-negativity restriction of operating frequency.

The lower level model is passenger flow distribution model, the first model is an allocation model based on user equilibrium. It is used to figure out passenger

sharing rate of each traffic mode. The second one is used to divide passenger flow into trains of different ranks in different routes

3.2 *The Simulated Annealing (SA) Algorithm*

SA simulates annealing process of solid [7]. An annealing schedule is given to control algorithm process. When control parameter anneals slowly to almost 0, a relative global optimal solution of combination optimization problem comes into being. The SA algorithm of the model is designed as follows:

1. **Generation of solution space**

In the model, the decision variables are operating route $l(l \in L)$ and frequency f_l^β of β rank trains. l can be obtained by the set of starting and terminal stations.

2. **Generation of new solution**

Due to the operating frequency of trains must be a natural number, the new solution can be made up by adding 1 to an element in the previous matrix randomly.

3. **Objective function**

We use the objective function of upper level programming as the goal function.

4. **Choosing initial solution**

SA does not depend much on initial solution, it is better to think of constraint condition when we choose an initial solution.

5. **Acceptance criterion**

We adopt metropolis criterion.

6. **Way to control temperature**

The algorithm will stop if control parameter T declines to the preset value.

4 An Example

4.1 *Case Background and Relevant Data*

There are 4 stations along an intercity railway, A and D are two big cities in the intercity corridor which are the start and terminal point of the line. C is a medium-sized city and B is a small city. There are four traffic modes in the transportation corridor which are intercity trains, common rail, bus, car, among them, there are three intercity train grades. The distance between cities and the

hourly passenger flow between OD pairs are shown in (Fig. 1; Table 1). In the paper, we assume that the distance between cities is the travel length of expressway and railway.

The tariff rates, travel speed, waiting time, safety, and convenience of different transportation are shown in Table 2. (The tariff rates and travel speed are estimated according to the actual value of Shanghai–Hangzhou corridor.)

The relevant time standard and other parameters are shown in Table 3.

4.2 Result Analysis

4.2.1 Optimum Operating Frequency

Based on the data of line condition, operation plan, train ranks, etc., we use the SA to calculate operating frequency of different ranks for intercity trains. The results of the sharing rate of all transport modes is shown in Fig. 1.

Considering the external cost’s internalization, the model takes the external cost of cars into account, so we can see that the users of cars transfer into other transport modes due to the increase of cars’ generalized transport cost.

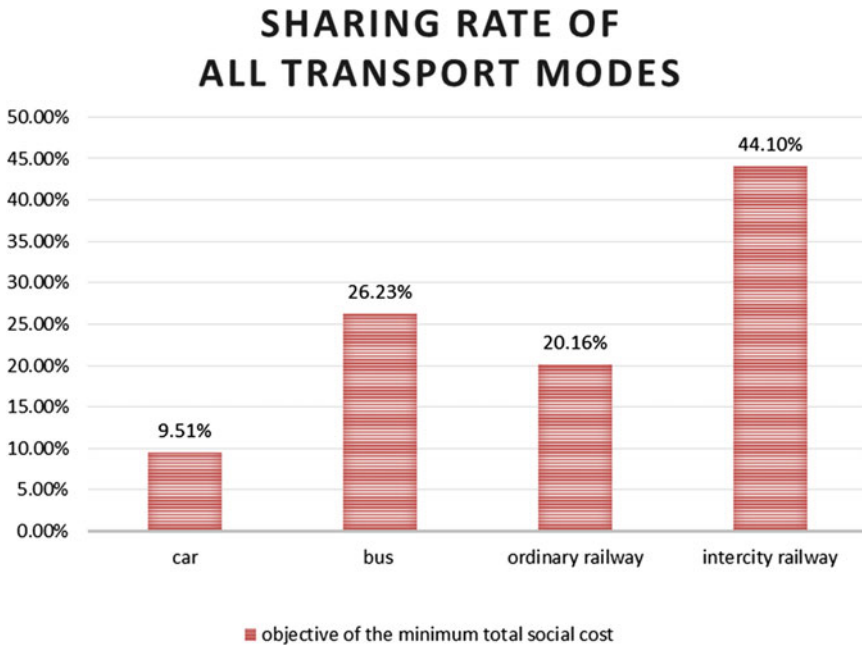


Fig. 1 Comparison between sharing rate of all transport modes in transport corridor

Table 1 The hourly passenger flow between OD pairs (person-time/hour)

	A	B	C	D	
A	0	9000	15,000	24,000	48,000
B		0	3000	6000	9000
C			0	15,000	15,000
D				0	0
		9000	48,000	45,000	72,000

Table 2 Relevant parameters of some transport modes

Transport modes	Basic tariff rate (RMB/km)	Travel speed (km/h)	Total time of entering and exiting station (min)	Waiting time (min)	Safety	Convenience
Intercity railway	0.6	200	Table 3	Formula (12)	1	1
Ordinary railway	0.4	60	40	15	0.9	0.36
Bus	0.2	80	30	10	0.7	0.43
Car	1	80	0	0	0.8	0.6

Table 3 The relevant time standard

Station	Entering time (min)	Exiting time (min)	Dwelling time (min)
A	25	22	5
B	10	10	3
C	15	12	3
D	25	22	5

When the objective is the minimum total social cost, the optimal operating frequency of the through trains, skip-stop trains, and stop by stop trains should be 3, 4, and 3 trains per hour.

4.2.2 The Sensitivity Analysis of Operating Frequency

To simplify the model, we just take the trains as studying objects to analyze the relations between intercity train operating frequency and total social cost. The concrete situations are shown in Fig. 2.

It can be seen from Fig. 2 that when the social cost is minimum, the operating frequency of intercity trains is 5 trains per hour. But the best operating frequency of intercity trains is 4 trains per hour when railway enterprises want the most benefits,

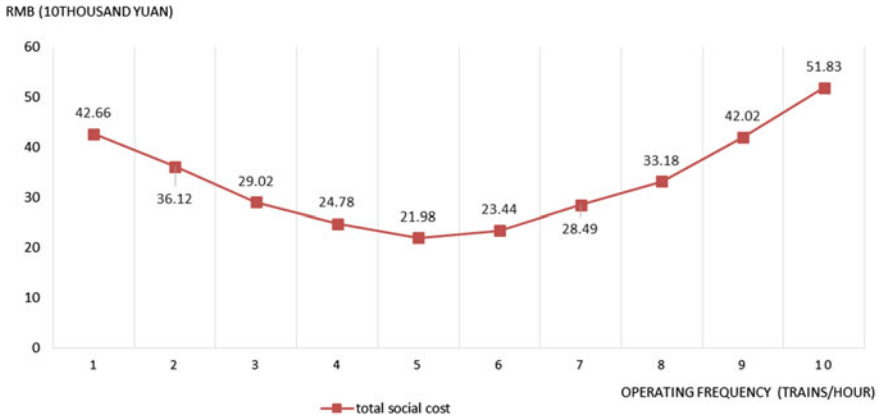


Fig. 2 Relationship between intercity railway operating frequency and total social cost

so it is one more train per hour under the target of the maximum railway enterprise benefits.

The following conclusions can be drawn from the above results:

- When the number of service frequency of intercity railway is less than 5, the total social cost decreases as the frequency increases. When the number of service frequency of intercity railway is more than 5, the total social cost increases as the frequency increases.
- While the total social cost is minimal, the benefits of railway enterprise can not be maximal. When the number of service frequency is 5, railway enterprise benefits will decrease.
- Hence, in order for that two objectives to achieve the optimal value at the same time, the government departments need to push appropriate preferential policies and make up for the sacrifices made by railway enterprise to minimize the total social cost.

5 Conclusions

In this paper, we establish a bi-level programming model. Its upper level model targets at the minimal total social cost. Its lower level is a model of equilibrium passenger flow of different transport modes in transport corridor. We adopt Simulated Annealing Algorithm to obtain the sharing rate of different transport mode, the passenger flow and operating frequency of different rank trains. Verified by an example, we compare and analyze the variation trend of the minimum total social cost under different operating frequencies of intercity trains. Through analysis, we can know that reasonable intercity railway frequency can effectively reduce

the external cost of the whole transport corridor. Therefore, at the decision point, government departments should slant toward intercity railway which has more social benefits, and provide certain subsidies and preferential policies to make up for the sacrificial benefits of the railway enterprise.

Acknowledgments This work is supported by Ministry of Science and Technology of China (2016YBF1200401).

References

1. Iljoon C (2001) A network-based model for market share estimation among competing transportation models in a regional corridor. University of Maryland
2. Huang J, Peng Q (2009) The distribution of passenger trains flow based on elastic demand. *Railway Transp Econ* 31(8):78–80 (in Chinese)
3. Shi F, Zhou W, Chen Y, Deng L (2008) Optimization study on passenger train plans with elastic demands. *J China Railway Soc* 30(3):1–6 (in Chinese)
4. Alberto C, Barnhart C, Laporte C (eds) (2007) *Handbook in OR&MS*, vol 14. Chapter 3: Passenger railway optimization
5. Goossens J-W, van Hoesel S, Kroon L (2001) A branch-and-cut approach for solving line planning problems. Netherland
6. Li Y, Yuan Z, Chen J, Tian M (2009) Combined capacity optimization of comprehensive transportation corridor based on multi-objective decision making. *J Jilin Univ (Eng Technol Edn)* 39(6):1480–1485 (in Chinese)
7. Pang F (2006) The principle of SA algorithm and algorithm's application on optimization problem. Jilin University, Jilin (in Chinese)

High-Speed Railway Timetable Rescheduling Method: A Bi-level Integrated Programming Approach

Huiru Zhang, Limin Jia, Li Wang, Yong Qin and Min An

Abstract China's railway has experienced a large-scale development in the recent years. Making up for delay time considering the energy efficiency when the train is delayed, which can satisfy the travel demand for passengers and save rail energy costs at the same time, will become the focus of future research on the railway. A bi-level programming optimization model is proposed in this paper. In the upper layer, the high-speed railway timetable is adjusted under unexpected interferences, and then the energy saving is optimized in the lower layer. A real-world case study is presented to illustrate the validity of the model and algorithm.

Keywords Railway transportation · Integrated timetable · Energy saving

1 Introduction

High-speed lines are attracting more and more passenger flow because of the high train speed and frequency and better service. However, the risk exists at all times and passengers have to face the possibility of different degrees of trains delay. With the development of train control technology, it is more and more practical to consider the energy consumption while rescheduling the timetable.

H. Zhang · L. Jia · L. Wang (✉) · Y. Qin · M. An
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: wangli@bjtu.edu.cn

H. Zhang · L. Wang
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

M. An
School of Built Environment, University of Salford, Salford M54WT, UK

L. Jia
Beijing Engineering Research Center of Urban Traffic Information Intelligent
Sensing and Service Technologies, Beijing 100044, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_65

Train operation adjustment is a large-scale combination of optimization, which needs to consider a lot of factors. A linear programming model has been designed to determine the optimal avoidance station and the cross station [1]. The minimum difference between the actual timetable and the original timetable has been taken as the optimization target [2]. In general, the factors that are considered for the train operation adjustment include the total late time, the number of late trains, the late rate, etc [3, 4].

In recent years, some scholars began to explore the design of energy-saving timetable. Energy-saving timetable can be achieved by optimizing the stopping time at the station and the coasting time in the interval [5]. A comprehensive timetable can be formed through the optimization of timetable and the speed curve between adjacent stations [6]. The current energy-saving timetable is mainly aimed at urban rail transit, while scholars also have proved the effectiveness of energy-saving timetable in high-speed railway applications [7–9].

Currently, the primary goal of the late train operation adjustment is to reduce the delayed time, so the energy consumed is almost neglected in this process. In this paper, a bi-level model based on high-speed railway timetable optimization is proposed under the delayed condition. In the rest of this paper, the problem of the timetable rescheduling of the upper layer model and the energy saving of the lower layer model will be described in the bi-level model. In the following section, the solving algorithm will be introduced. A case study of Wuhan to Guangzhou high-speed railway will be illustrated in the section of case study and the conclusion of the results are in the final section.

2 The Bi-level Model

2.1 Timetable Rescheduling Problem

The purpose of this part is to adjust the train movements to be consistent with the original timetable as much as possible within a certain range of constraints.

(1) The Decision Variables

$a_{i,j}$ is the new arrival time of train i at station j after adjustment. $d_{i,j}$ is the new departure time. $\text{delay}_{i,j}$ is the delay, that is, the difference in the arrival time between the adjusted and the original one. If the departure time of train i is earlier than train i' in the original timetable, then $T_{i',j}^{\text{original}} = 1$, otherwise $T_{i',j}^{\text{original}} = 0$. If the departure time of train i is changed later, then $T_{i',j}^{\text{reschedule}} = 1$, otherwise $T_{i',j}^{\text{reschedule}} = 0$. If train i occupies track l , then $\lambda_{i,j}^l = 1$, otherwise $\lambda_{i,j}^l = 0$.

(2) The optimization objective

Take a line with m trains and n stations as an example, and the goal of the model is to minimize the total delay time:

$$\min z_d = \sum_{i=1}^m \left(f_i \cdot \sum_{j=1}^n \text{delay}_{i,j} \right). \quad (1)$$

z_d is the total delay time and f_i is the cost per unit time delay for train i .

(3) The Constraints

$\text{sec}_{i,j}$ is the minimum running time of train i on section $[j, j + 1]$. $\text{sta}_{i,j}$ is the minimum dwell time at station j . N is a large enough natural number. τ_j^{sec} is the headway of section $[j, j + 1]$. τ_j^{sta} is the headway between two adjacent trains at station j . ω_j is the track number of station j .

(a) *Section running time restrictions:*

$$a_{i,j+1} - d_{i,j} \geq \text{sec}_{i,j} \quad (2)$$

(b) *Station dwell time restrictions:*

$$d_{i,j} - a_{i,j} \geq \text{sta}_{i,j} \quad (3)$$

(c) *Section headway restrictions:*

$$(d_{i',j} - d_{i,j}) + N(1 - T_{i,i',j}^{\text{original}}) \geq \tau_j^{\text{sec}} \cdot T_{i,i',j}^{\text{original}} \quad (4)$$

$$(d_{i,j} - d_{i',j}) + N(1 - T_{i,i',j}^{\text{reschedule}}) \geq \tau_j^{\text{sec}} \cdot T_{i,i',j}^{\text{reschedule}} \quad (5)$$

(d) *Station headway restrictions:*

$$\lambda_{i,j}^l + \lambda_{i',j}^l - (T_{i,i',j}^{\text{original}} + T_{i,i',j}^{\text{reschedule}}) < 1 \quad (6)$$

$$(a_{i',j} - a_{i,j}) + N(1 - T_{i,i',j}^{\text{original}}) \geq \tau_j^{\text{sta}} \cdot T_{i,i',j}^{\text{original}} \quad (7)$$

$$(a_{i,j} - a_{i',j}) + N(1 - T_{i,i',j}^{\text{reschedule}}) \geq \tau_j^{\text{sta}} \cdot T_{i,i',j}^{\text{reschedule}} \quad (8)$$

(e) *Track restrictions:*

$$\sum_{l=1}^{w_j} \lambda_{i,j}^l = 1 \quad (9)$$

2.2 Energy-Saving Problem

The purpose of this part is to optimize the energy efficiency of the train timetable. The energy consumption is taken as the optimization target.

(1) The Decision Variables

t_k is the optimal operation time of train in section k . v_y^k is train constant speed. t_d^k is train coasting time. S_f^k , S_y^k , S_d^k , and S_b^k are train acceleration, cruising, coasting and braking distance in section k respectively. $E_i(t_k)$ is the energy consumption.

(2) The optimization objective

$$\min z_e = \min \sum_{i \in I} \sum_{k \in K} E_i(t_k), \quad (10)$$

$$E_i(t_k) = \int_0^{t_k} k_f(t) \cdot v(t) \cdot F(t) dt. \quad (11)$$

k is running section and K is running section set. i is train and I is train set. z_e is energy consumption. $k_f(t)$ is traction coefficient, $k_f \in [0, 1]$. $v(t)$ is speed.

(3) The Constraints

k_b is braking force coefficient. $k_b \in \{0, 1\}$. M is train quality. W is running resistance, here we only consider the basic resistance. t_f^k , t_y^k , and t_b^k is train acceleration, cruising, and braking time respectively. S is the total distance. V_0 is initial speed. V_T is final speed and V_{\max} is maximum speed allowed. t_k^{\min} and t_k^{\max} are the minimum and maximum limits. T is the total running time after the upper model is adjusted.

(a) *Force restrictions:*

$$k_f F - k_b B - W = M \cdot (dv(t)/dt) \tag{12}$$

(b) *Distance restrictions:*

$$\begin{cases} S = \sum_{i^*} S^*, \\ S^* = \int_0^{t^*} v(t) dt. \end{cases} \tag{13}$$

When $S^* = S_f^k$, then $t^* = t_f^k$; similarly, when $S^* = S_d^k$, $t^* = t_d^k$; when $S^* = S_b^k$, $t^* = t_b^k$; when $S^* = S_y^k$, $t^* = t_y^k$, and the formula is equivalent to $S_y^k = v_y^k \cdot t_y^k$.

(c) *Speed restrictions:*

$$\begin{cases} V_0 = 0 \\ V_T = 0 \\ v \leq V_{\max}. \end{cases} \tag{14}$$

(d) *Time restrictions:*

$$\begin{cases} t_k = \sum t_*^k, \quad t_*^k \in [t_f^k, t_y^k, t_d^k, t_b^k] \\ t_k^{\min} \leq t_k \leq t_k^{\max} \\ \sum_{k \in K} t_k = T. \end{cases} \tag{15}$$

(e) *Tracking interval restrictions:*

The purpose of the restrictions is to ensure that the departure and arrival time of adjacent train i and train i' satisfies the minimum tracking interval in the station, and the running time satisfies the minimum tracking interval in section. The specific expression of the tracking interval is the same as the upper layer.

3 Solving Algorithm

An iterative loops algorithm is designed to solve the model. First, the arrival and departure time is adjusted by the upper layer. Second, the time is optimized by lower layer. Third, the result is returned back to the upper layer to check the constraints. So it is iterated, and is expected to converge to an optimized solution.

Step 1: The adjustment of the arrival, departure, and stop time.

Step 1.1: Read the file, and get train, station, line, and late information.

Step 1.2: Select the Wuhan to Guangzhou railway data, and sort the train.

Step 1.3: Adjust the timetable according to the constraints using Ilog Cplex.

Step 1.4: Get the arrival and departure time after the train is adjusted.

Step 2: Energy-saving optimization

Step 2.1: Calculate the shortest running time of all trains.

Step 2.2: Calculate the spare time between the adjusted and shortest time.

Step 2.3: Distribute the spare time little by little using MATLAB.

Step 2.3.1: Divide the total spare time into small portions.

Step 2.3.2: Each section adds the same amount of spare time.

Step 2.3.3: Select the section with the most energy reduction to add the small amount of spare time, and time in other section is unchanged.

Step 2.4: Get optimized train arrival and departure time.

Step 3: Check whether the result meets the constraints of the upper layer.

Step 3.1: If all the constraints are met, go to step 4.

Step 3.2: Otherwise, go to step 1.3.

Step 4: Output the integrated timetable.

Step 4: Output the integrated timetable.

4 Case Study

The total length of the railway line is 1069 km, and includes 16 stations. Select the data from 9 to 11 o'clock, and there are 11 trains. Sorting them by start time is G77, G1109, G1111, G1143, G1007, G551, G93, G1113, G541, G1011, and G1115. Since the high-speed line is a double track, we only consider the direction from Wuhan to Guangzhou without loss of generality.

In this paper, a very short delay is studied. Because the timetable only contains 2 hours, so more serious delay is not considered.

4.1 The Delay Adjustment

In this section, the delay time is set to within 5 min. To simplify the analysis process, the delay of G1115 is not considered, because it is the last train.

Observe the train timetable: (a) The average departure time interval is 4 min, so the delay is easy to spread. (b) After three times overtaking, G93 became the fifth

Table 1 Train delay after optimization

The nature, search time, and value of z_d	Delay on terminal stations (s)	Number of affected intermediate stations	Delay on intermediate stations (s)
Optimal, 0.17 s, 972.0	0	5	652

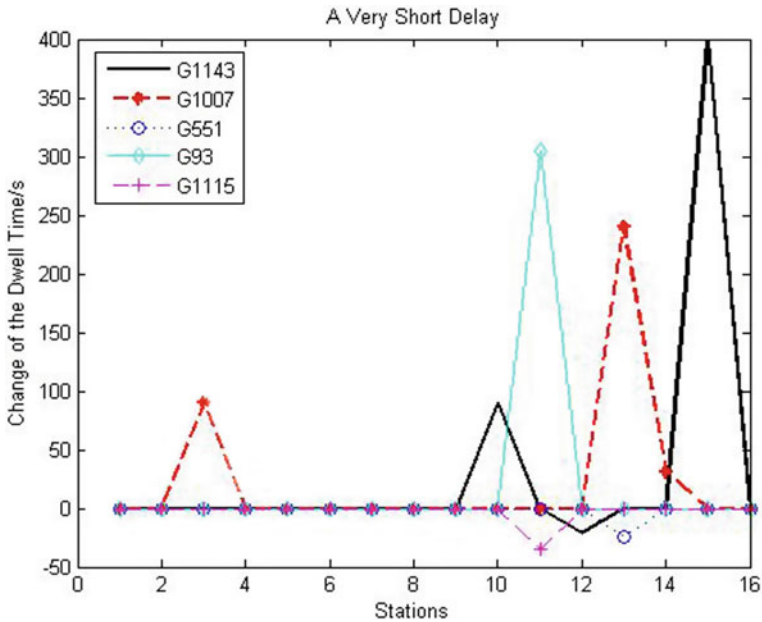


Fig. 1 Change of the dwell time after a very short delay

train to reach the terminal station, so a small delay in G93 will have a great impact on other trains. In the case study, the delay time of G77, G1143, and G93 is set 3, 5, and 2 min, respectively.

All the trains are in accordance with the original timetable to reach the terminal station, of which arrival and departure time of 5 intermediate stations has changed, as shown in Fig. 1, and the total time delayed on intermediate stations is 652 s. The search time of the algorithm is 0.17 s, which shows the feasibility of the algorithm (Table 1).

The impact of the delay on G1143 is the largest, and G1143 stops in Guangzhou North for nearly 400 s. The increase in dwelling time is to make G93 to overtake. Then, the result is input to the lower layer to further optimize.

Table 2 Train energy consumption comparison table

Train number	*EECOT (kw h)	*ECEST (kw h)	Train number	*EECOT (kw h)	*ECEST (kw h)
G77	8018	7948.141	G93	8018	7835.567
G1109	8018	7847.609	G1113	8018	7507.501
G1111	8018	7703.140	G541	8018	7912.983
G1143	8018	7803.907	G1011	8018	7754.668
G1007	8018	7871.933	G1115	8018	7658.879
G551	8018	7609.463	–	–	–
*TEC(kw h): *EECOT-88198, *ECEST-85453.790			*ESPCEV: 3.1%		

EECOT Experienced Energy Consumption of the Original Timetable (kw h), *ECEST* Energy Consumption of the Energy-Saving Timetable (kw h), *TEC* Total Energy Consumption (kw h), *ESPCEV* Energy-Saving Percentage Compared to Experience Value

4.2 Energy-Saving Optimization

Table 2 shows the energy consumption value of each train under different time conditions. Here we take the train drive system energy conversion efficiency as $\eta = 0.9$. Experience has indicated that the eight groups of CRH series EMU trains running 1069 km will consume 8018 kw h electric energy. Based on this data, after energy-saving optimization, the whole line can save energy 2744.210 kw h, accounting for 3.1% of the total energy, and the energy-saving effect is obvious.

5 Conclusion

This paper presents a bi-level integrated programming approach on high-speed railway timetable rescheduling, which is designed to optimize the timetable under the delayed condition. The upper layer reschedules the timetable by considering the minimum delay time and the constraints, so that the train movement can be consistent with the original schedule as much as possible. In the lower layer, the energy efficiency is optimized under the constraints by the small-scale adjustment of the arrival and departure time at intermediate stations. Taking the actual operation data of Wuhan to Guangzhou high-speed railway as an example, the case study is carried out on a different degree of time delay, and the results verify the feasibility of the model and algorithm.

Acknowledgments This study is funded by the National Key Research and Development Programme of China (2016YFB1200401), National Natural Science Foundation of China (71701010), the Fundamental Research Funds for the Central Universities (2017JBM033), and Beijing Jiaotong University State Key Laboratory of Rail Traffic and Control and Safety (RCS2017K001 & RCS2016ZT016).

References

1. Szpigel B (1973) Optimal train scheduling on a single track railway. *Oper Res* 72:343–352
2. Zhang Y, Jin W (2005) Model and algorithm for train operation adjustment on single-track railways based on genetic algorithm. *J Southwest Jiaotong Univ* 40(2):147–152 (in Chinese)
3. Corman F, D'Ariano A, Pacciarelli D, Pranzo M (2012) Bi-objective conflict detection and resolution in railway traffic management. *Transp Res Part C* 20(1):79–94
4. Almodóvar M, García-Ródenas R (2013) On-line reschedule optimization for passenger railways in case of emergencies. *Comput Oper Res* 40(3):725–736
5. Wong K, Ho T (2007) Dwell time and run-time Control for DC mass rapid transit railways. *IET Electr Power App* 1(6):956–966
6. Su S, Li S, Tang T et al (2013) A subway train timetable optimization approach based on energy-efficient operation strategy. *IEEE Trans Intell Transp Syst* 14(2):883–893
7. Wang T, Chen F, Zhang Q (2014) Method of energy-saving operation diagram compiling for train operation dispatch of high-speed railway. *Railway Comput Appl* 23(12):5–8 (in Chinese)
8. Li X, Hong K (2014) An energy-efficient scheduling and speed control approach for metro rail operations. *Transp Res Part B* 64:73–89
9. Yang X, Anthony C, Bin N et al (2016) A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass. *Transp Res Part B* 91:424–445

Characteristic Analysis of High-Speed Railway Network in China

Dan Guo, Yong Qin, Li Wang and Min An

Abstract The high-speed railway has been developed rapidly, and the topological structure and transportation organization of the railway network have changed greatly. In this paper, China's high-speed railway is taken as the research object and train flow network is established. Indexes are calculated from the perspective of the network and transportation organization. The comprehensive evaluation index is determined with TOPSIS method to identify key indicators of the station. The example identifies not only the provincial capital stations, but also the stations with small scale but critical position affecting the whole network transportation service capacity. The research of high-speed railway network property contributes to forming the transport organization strategy, which makes the network service feature more robust and helps for the optimization of railway network infrastructure.

Keywords High-speed railway · Complex network · Train flow network

D. Guo · Y. Qin
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China

D. Guo · L. Wang (✉)
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing 100044, China
e-mail: wangli@bjtu.edu.cn

Y. Qin
National Engineering Laboratory for System Safety and Operation
Assurance of Urban Rail Transit, Guangzhou 510330, China

L. Wang
Beijing Engineering Research Center of Urban Traffic Information Intelligent
Sensing and Service Technologies, Beijing 100044, China

M. An
School of Built Environment, University of Salford, Salford M5 4WT, UK

1 Introduction

With the implementation of the “medium and long-term railway network planning”, the development of railway in China has made remarkable achievements, which has played an important role in promoting economic and social development. How to improve the reliability and survivability of the high-speed railway network becomes the key problem in the field of high-speed railway research.

In recent years, more and more research on complex network theory has been made. With the discovery of small-world network characteristics and the scale-free network characteristics, the research and application of complex network theory in various fields have become more extensive. Wang [1] put forward the method of calibrating the important nodes and analyzed the survivability of the railway network [2]. Jiang [3] and Zhang [4] established the railway geography network model and the railway transportation network model. They both used the network global efficiency as the reliability index of the railway network, and simulated the emergencies to verify the network reliability. Wang et al. [5] proposed the catastrophe model to describe the dynamic change process of the railway system security, so as to establish the system risk structure and propose recommendations for the safety of rail systems. Guidotti et al. [6] proposed a method to quantify the network reliability probability based on the existing (diameter and efficiency) and new (eccentric and heterogeneous) connectivity methods, considering the weight of the edges and nodes and the auxiliary nodes. Zhou et al. [7] introduced a method of building direct weighted accident causal network, which was based on the complex network and accident event chain. Dunn et al. [8] described a new method for quantifying the reliability of complex systems using network theory. Zhang et al. [9] proposed a method to extract the central network in urban rail transit, analyzed its characteristics and verified the feasibility and effectiveness of the method. Deng [10] proposed a new framework based on network theory and FMECA method. The vulnerability of the subway system was studied by using the network theory and the FMECA risk matrix method.

Most of the existing research focuses on urban rail transit or road traffic, but the results on high-speed railway system are less. Based on the theory of complex network, this paper established a high-speed railway train flow network from the perspective of geography accessibility and transportation organization. We analyzed the topological properties of the high-speed railway network in China by calculating indexes such as the degree, clustering coefficient, and average network distance which represent the static statistical characteristics of the network. And the centrality metric is used to identify the key stations in the traffic flow network, so as to provide a reference for improving the reliability of high-speed railway operation network.

2 Construction of High-Speed Railway Train Flow Network

We establish the train network with the specific method as follows: the stations that a train originates, stops, or finally arrives at are regarded as a “node”. There are connections between two nodes as long as the same train stops at the two stations, that is, we take the train flow between two stations as the lines in the network, and the number of trains which stop at the two stations as weight. The network can be represented by an undirected weighted graph. An example is as follows:

Take A, B, C, D as four sites on a railway line, and train D1001 stops at A, B, C, D, D1002 stops at A, C, D. Line AB, AC, AD, BC, BD, and CD are connected by train D1001, and line AC, AD, CD are connected by the train D1002. Duplicate edges are not drawn anymore, but the weight of the line plus one. High-speed rail train flow network reflects the traffic flow connection between two stations (Fig. 1).

Train flow network is based on physical connection, and the station has specific geographical coordinates. Traffic connections between the two stations are premised on the existence of actual tracks between them, and train traffic frequency, travel interval, train type refer to the statistics of passenger flow volume and freight volume. Therefore, the traffic flow network is a network which embodies the physical connection, OD demand, and transportation organization strategy.

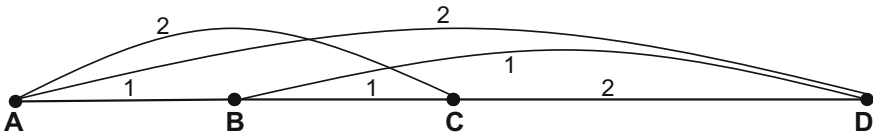


Fig. 1 Sketch map of train flow network

Fig. 2 Sketch map of high-speed railway geography network structure in China

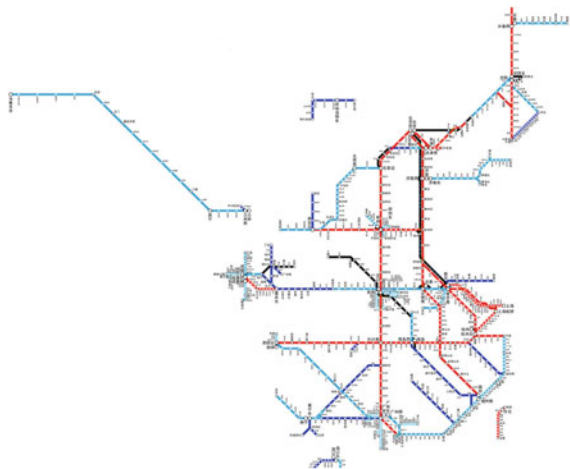
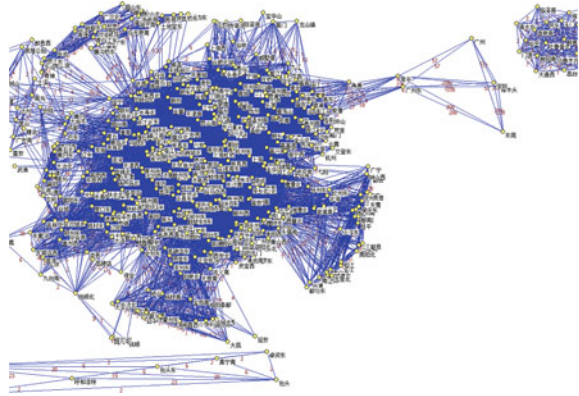


Fig. 3 Sketch map of high-speed railway train flow network structure



The “long-term railway network plan (2016)” proposed that China has formed a “four vertical and four horizontal” high-speed railway network (Fig. 2). On this foundation, expansion of high-speed railway lines that are supported by passenger flow, standard compliant and accord with development can form a high-speed railway network with the “eight vertical and eight horizontal” as the backbone.

China high-speed railway is defined as follows: it is a passenger-dedicated railway, and CRH trains are newly designed to run at a speed of 250 km/h (including reserved), its initial operating speed is not less than 200 km/h. Therefore, this paper mainly considers the following three trains: high-speed train (G), intercity train (C), and CRH trains that run on the existing lines (D). The train is running in pairs in the high-speed railway transport organization, so the high-speed railway network is simplified as an undirected network on the basis of not affecting the network characteristics. In this section, we use the timetable data of October 10, 2015 to extract the high-speed railway station and train traffic flow information. An undirected weighted network with 487 nodes and 14,194 edges is obtained according to the traffic flow network construction method mentioned above as shown in Fig. 3.

3 Characteristics Analysis of Traffic Flow Network of High-Speed Railway

3.1 Topology Characteristics of Network

In this paper, Pajek and R are used to realize the calculation. The number of unreachable pairs in China traffic flow network is 26,872, and the average distance between reachable pairs is 2.28338. The most distant vertices are Dongguan and Pixian, and distance is 6. The Watts–Strogatz clustering coefficient is 0.72677851 and network clustering coefficient (transitivity) is 0.53988206.

3.2 Characteristics Analysis of Stations

(1) Parameter definition

In order to clarify the key station identification indicators more clearly, the parameters involved are defined as follows:

V is the set of nodes in the network, which is the collection of all stations that trains originate, stop, or finally arrive at in the high-speed railway traffic flow network. E is the set of edges in the network, which is the collection of all traffic flow edges in the high-speed railway traffic flow network.

w_{ij} is the weight coefficient of traffic flow, which represents the number of trains passing through the station v_i and v_j the station at the same time.

O is the set of originating stations and D is the set of final arrival stations.

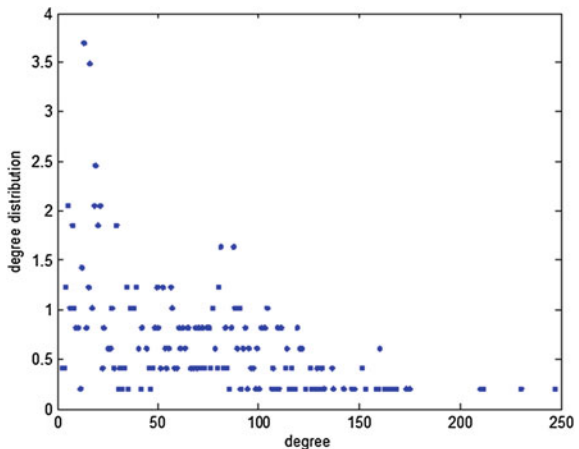
R is a collection of all feasible paths in the network. R_{od} is a set of feasible paths from a node o to d in the network. $o \in O, d \in D$, and $R = \cup R_{od}$. r_{od} is one of the feasible paths in the network from node o to node d .

SR is a collection of all shortest paths in the network. SR_{od} is a set of shortest paths from a node o to d in the network. $o \in O, d \in D$ and $SR = \cup SR_{od}$. sr_{od} is one of the shortest paths in the network from node o to node d .

(2) Degree distribution of stations

By calculating the station degree of the traffic flow network model of high-speed railway, the station degree distribution is plotted as shown in Fig. 4. In 2015, the highest degree of a node in the high-speed railway network was 247, and the average degree is 58.29158111, indicating that an average station can be transferred to 58 stations without changing. Only 17.5% of stations whose degree is greater than 100 and only 3.7% of the stations whose degree is greater than 150. This

Fig. 4 Degree distribution for high-speed railway traffic flow network



shows that most stations in the high-speed railway network have a low degree of distribution and relatively centralized.

(3) Cluster coefficient distribution of stations

The average cluster coefficient of high-speed railway traffic flow network is 0.723463 throughout the calculation, and Fig. 5 shows the distribution of cluster coefficient for high-speed railway traffic flow network. Traffic flow network shows high aggregation characteristics. The relationship between cluster coefficient and degree of each node was shown in Fig. 6. The data reflect that the connection degree of adjacent stations is high, and most stations whose clustering coefficient is 1. The general trend showed that stations with a lower degree have a smaller clustering coefficient, and the two are of negative correlation.

(4) Traffic intensity

Fig. 5 Distribution of cluster coefficient

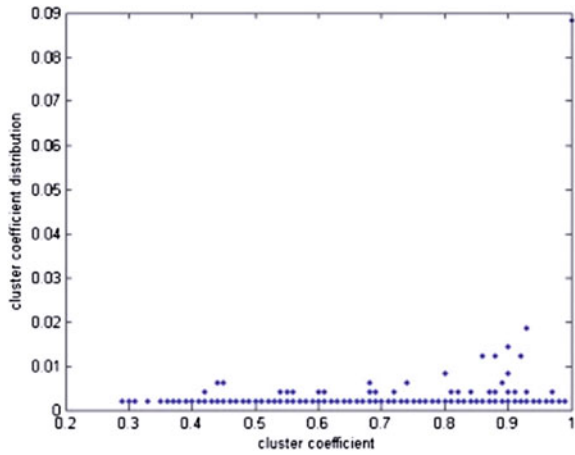
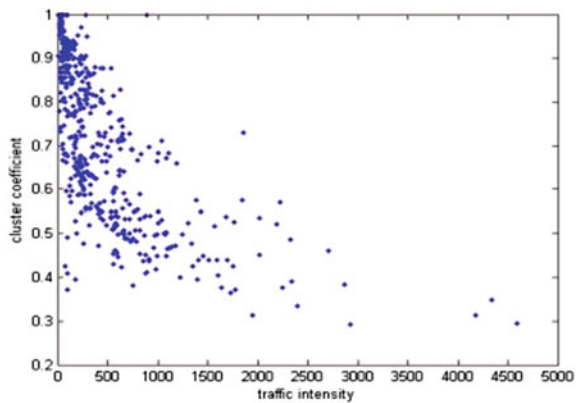


Fig. 6 Distribution of cluster coefficient versus traffic intensity



The concept of traffic intensity is presented as formula (1) which is as follows:

$$S_i = \sum_j w_{ij}. \tag{1}$$

w_{ij} is the number of trains passing through the station v_i and the station v_j at the same time, it is also the weight of the edge e_{ij} . Since the high-speed railway traffic flow network is undirected, we only consider one-way traffic intensity, which reflects the number of stations that a passenger can reach directly from the station without transferring to other stations and the service capacity and frequency of the station. The greater the traffic intensity, the better the accessibility of the station. It has a wider range of path options. The traffic intensity can be understood as the degree centrality considering the train flow.

(5) Traffic closeness centrality

We define traffic closeness centrality as formula (2) which is as follows:

$$CF(i) = \frac{\sum_{d \in D, d \neq i} \sum_{sr_{id} \in SR_{id}} f(sr_{id})}{N - 1}. \tag{2}$$

$f(sr_{id})$ represents the traffic flow of shortest path sr_{id} . The shortest path is usually defined as the path which includes the least nodes between the two nodes, but in the high-speed railway traffic flow network, the traffic flow should be taken into account. Therefore, a new definition of the shortest path is proposed as shown in (3) and (4).

$$SR_{od} = \left\{ r_{od} \mid \max \frac{\sum_{k=1}^{r_{od}} w_k}{|r_{od}|} \right\}. \tag{3}$$

$$f(sr_{id}) = \frac{\sum_{k=1}^{|sr_{id}|} w_k}{|sr_{id}|}. \tag{4}$$

w_k represents the traffic flow in a section of the path, that is, the weight of an edge in the path. $|r_{od}|$ indicates the actual topology path length of the path r_{od} . Similarly, $|sr_{id}|$ indicates the actual topology path length of the path sr_{id} .

In the high-speed railway traffic flow network, the traffic closeness centrality of stations reflects the speed of transmission for failure or emergency resources when the station is the origin of faults or emergency resources. It also describes the average flow that it can transmit from the central station to other stations (not only

adjacent stations). The higher the traffic closeness centrality of the station, the spread of information or resources will be faster, and the station is more important.

(6) Traffic betweenness centrality

We define traffic betweenness centrality as shown in formula (5) which is as follows:

$$BF(i) = \frac{\sum_{(o,d) \in (O,D)} \sum_{sr_{od} \in SR_{od}} [f(sr_{od}) \bullet \varphi_i(sr_{od})]}{\sum_{(o,d) \in (O,D)} \sum_{sr_{od} \in SR_{od}} f(sr_{od})}. \tag{5}$$

When the shortest path sr_{od} passes through the station i , $\varphi_i(sr_{od}) = 1$, otherwise $\varphi_i(sr_{od}) = 0$. Define $f(sr_{od})$ in the same way as in traffic closeness centrality, it represents the traffic flow of sr_{od} .

Traffic betweenness centrality reflects the size of the impact of the station. In the high-speed railway traffic flow network, higher traffic betweenness centrality indicates that the station can control more information or resources in network transmission, and the station is more important. Controlling the station whose traffic betweenness centrality is larger in the event of an emergency in the network can reduce the range of failure.

(7) Eigenvector centrality

Eigenvector centrality is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. We use the adjacency matrix to find eigenvector centrality.

For a given graph $G = (V, E)$ let $A = (a_{i,j})$ be the adjacency matrix, i.e., $a_{i,j} = 1$ if the vertex i is linked to vertex j , and $a_{i,j} = 0$ otherwise. The centrality score of the vertex i can be defined as

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j = \frac{1}{\lambda} \sum_{j \in G} a_{i,j} x_j, \tag{6}$$

where $M(v)$ is a set of the neighbors of i and λ is a constant. With a small rearrangement, this can be rewritten in vector notation as the eigenvector equation $Ax = \lambda x$.

The four indexes of each station in the high-speed railway train flow network are calculated. In order to integrate the four indexes, the TOPSIS method was used to evaluate each station and the top 10 nodes of each index drawn in Table 1 represents ranking. B represents traffic intensity ranking. C represents traffic intensity. D represents traffic closeness centrality Ranking. E represents traffic closeness centrality. F represents traffic betweenness centrality ranking. G represents traffic betweenness centrality. H represents eigenvector centrality ranking. I represents

Table 1 Comparison of top 10 stations with indicators

A	B	C	D	E	F	G	H	I	J	K
1	SHHQ	4591	ZZ	0.0000634000	CDD	7411.719154	SHHQ	0.323824	HK	0.636122
2	HZD	4339	HND	0.0000634000	XJ	6523.408333	HZD	0.31608	CDD	0.599192
3	NJN	4182	CBB	0.0000634000	HK	6300.122776	NJN	0.284737	XJ	0.527024
4	WH	2928	ZZX	0.0000633000	SN	3822.89001	NB	0.231931	NJN	0.482428
5	CSN	2863	TA	0.0000633000	HND	3090.983467	SXB	0.202942	SHHQ	0.467614
6	NB	2705	LY	0.0000633000	LBB	3042.739579	WZN	0.188098	HZD	0.440011
7	GZN	2402	WHN	0.0000633000	ZZ	2882.366715	JXN	0.172485	WH	0.375295
8	JNX	2344	FD	0.0000632000	TA	2747.966823	TZ	0.162087	SN	0.369873
9	SXB	2322	NCX	0.0000632000	CD	2727.367856	YYB	0.147902	KSN	0.356154
10	XZD	2248	HF	0.0000632000	HG	2656.238242	XZD	0.144192	NCX	0.336264

eigenvector centrality. J represents comprehensive index ranking. K represents comprehensive index. The station name is abbreviated.

The first three stations were identified as Hankou Station, Chengdu East Station, and Xinjin Station.

Most of the stations in Table 1 are of high traffic intensity, that is, the stations that are municipalities and directly under the Central Government or provincial capital. There are also small stations that are usually overlooked. Large cities generally have several high-speed rail stations, such as Beijing West Railway Station, Beijing South Railway Station, and Beijing Railway Station. The comprehensive index values were 0.185836664, 0.248281523, and 0.088421432 respectively. If they merge into “Beijing”, the structure of the high-speed railway train flow network will change and the ranking of evaluation indexes will change too. HanKou, Chengdu East, and Xinjin Railway Station are located in the south-central part of China. North–south line and east–west line pass through them, so the traffic centrality index value is large. The comprehensive evaluation index takes a number of factors into account, which can effectively identify some stations which may be ignored, and it can provide an important reference for the future operation and maintenance and the new line planning.

Traffic intensity, traffic closeness centrality, traffic betweenness centrality, and eigenvector centrality define the importance of stations from different aspects, which provide a reference for the different target network adjustment. Stations with low eigenvector centrality and high centrality are important contacts, while stations with high eigenvector centrality and low centrality are associated with important stations. The daily maintenance of stations should focus on the station whose traffic intensity is high, and stations whose traffic closeness centrality is high should be chosen as originating stations in the unexpected events. Simultaneously, we also should pay attention to the daily maintenance and construction of stations whose comprehensive index are high.

4 Conclusion

In this paper, we used the October 10, 2015 China high-speed railway timetable to make an empirical analysis. Some stations with a low degree have high comprehensive value, indicating that these stations have higher closeness centrality, betweenness centrality, and eigenvector centrality, they may play an important role in the reliability of road network in the future, so we need to focus on the maintenance of these stations. In the normal operation condition, using the comprehensive evaluation index to evaluate the key station can provide a reference for key maintenance; in the network paralysis fault conditions, using a single index to evaluate the key station can provide a reference for the rapid recovery of network operation.

Acknowledgements This study is funded by the National Key Research and Development Programme of China (2016YFB1200401), National Natural Science Foundation of China (71701010 & 61374157), the Fundamental Research Funds for the Central Universities (2017JBM033), and Beijing Jiaotong University State Key Laboratory of Rail Traffic and Control and Safety (RCS2017K001 & RCS2016ZT016).

References

1. Wang W, Liu J et al (2010) Topology properties on Chinese railway network. *J Beijing Jiaotong Univ* 04:148–152 (in Chinese)
2. Wang W, Liu J et al (2010) Survivability analysis of railway network. *J China Railway Soc* 04:18–22 (in Chinese)
3. Jiang YC (2011) Study on reliability of railway network based on complex network theory. Southwest Jiaotong University (in Chinese)
4. Zhang YQ (2010) Study on the performance and spread of railway network in the event of emergency. Beijing Jiaotong University (in Chinese)
5. Wang Y, Weidmann UA, Wang H (2017) Using catastrophe theory to describe railway system safety and discuss system risk concept. *Saf Sci* 91:269–285
6. Guidotti R, Gardoni P, Chen Y (2017) Network reliability analysis with link and nodal weights and auxiliary nodes. *Struct Saf* 65:12–26
7. Zhou J, Xu W, Guo X et al (2015) A method for modeling and analysis of directed weighted accident causation network (dwacn). *Physica A* 437:263–277
8. Dunn S, Wilkinson S (2017) Hazard tolerance of spatially distributed complex networks. *Reliab Eng Syst Saf* 157:1–12
9. Zhang J, Wang S, Zhang Z et al (2016) Characteristics on hub networks of urban rail transit networks. *Physica A* 447:502–507
10. Deng Y, Li Q, Lu Y (2015) A research on subway physical vulnerability based on network theory and fmeca. *Saf Sci* 80:127–134

Coordination of Transport Organization and Building Maintenance for Jixi Marshalling Station

Jianxiang Li and Xiaoning Zhu

Abstract The coordination of transport organization and building maintenance is an important method by which to improve the efficiency of marshalling yard building maintenance. This study takes Jixi station as an example, and analyzes currently existing problems affecting building maintenance. It proposes a coordinated approach to the maintenance of buildings and the organization of marshalling station transportation. It uses network planning methods to analyze the shortest or longest times that and construction and reparation tasks require. Traffic data are analysed in order to determine the methods for the maintenance schedule of a building's skylights. A coordinated approach to the original station Jixi weekly skylight plan is proposed in order to optimize coordination, and ultimately to reduce the impact of construction on the repair work of marshalling station.

Keywords Jixi station · Marshalling station · Transportation organization
Maintenance construction · Coordination

1 Introduction

Safety is key in railway transportation. With the enlargement of the scale of the network, and with the constant improvement to train speed and traffic density, transportation equipment can break down, the maintenance of equipment is not carried out in a timely fashion, maintenance units can disagree with the station etc., and such phenomena are widespread, being more prevalent in marshalling stations. The proper organization of transport and building maintenance can partly improve transportation conditions and eliminate safety problems. An appropriate construction schedule can reduce the need for equipment maintenance or accident repairs and improve the organizational efficiency of the marshalling station. It can also play an important role in the organization of daily traffic operations in the marshalling

J. Li (✉) · X. Zhu

School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
e-mail: 16120830@bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_67

661

station [1, 2]. Many studies have prepared the M&R (Maintenance and Renew) program for line equipment by predicting the state of the equipment itself [4]. Based on the optimal allocation of maintenance resources [3], an adjustment to the relationship between transportation and maintenance is possible, allowing the formulation of the appropriate coordination of transport organization and building maintenance, in order to realize the win-win situation of economic benefit and transportation safety [5].

2 Current Situation and Existing Problem

2.1 Analysis of Transport Organization of Jixi Station

Jixi marshalling station runs a length of 5–6 km in a north-south direction. It is one of six secondary freight marshalling yards, and is one of the top ten network marshalling yards across the country. Jixi station upline traffic goes in four directions: to Dezhou, Handan, Jinan and Zibo. Jixi station exchanges traffic in the downline direction with Xuzhou, and this is at a lower level than the up direction. This should be considered when drawing up the skylight maintenance plan.

2.2 Jixi Marshalling Station Interlocking Area

The maintenance of “skylights” in marshalling yards is an extension of the line maintenance of skylight construction. Jixi station is divided into 36 interlocking areas at present, and these interlocking areas are differentiated as to their specific situation as follows.

Transportation equipment is different in the differently numbered interlocking areas. The amount of work decides the single skylight maintenance time length (Fig. 1).

2.3 Problems with the Marshalling Station Skylight Maintenance Plan

1. The skylight maintenance plan is not established on a scientific basis. The actual construction units in the formulation of the building maintenance scheme are not derived from in-depth field investigation and study and they cannot combine the actual situation of line equipment, driving method, density and so on in the preparation of the plan.

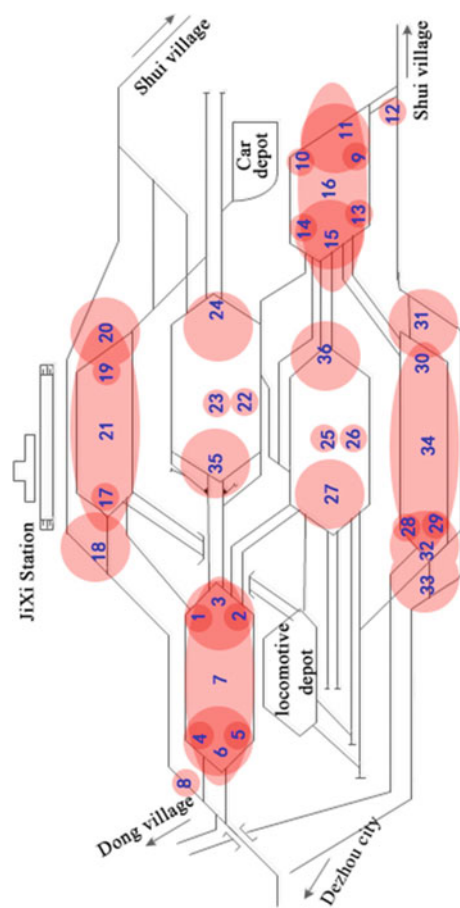


Fig. 1 Jixi station interlocking areas

2. Jixi station’s skylight maintenance time is fixed. As a result, the skylight maintenance time cannot be changed according to the uneven departure column flow situation very good together.
3. The plan is not appropriate. The station construction period often requires four to five maintenance operations at the same time. This increases the level of difficulty and also cannot guarantee the effective safety of labor.

3 Method of Coordinating Transport and Maintenance

3.1 Jixi Station Traffic Data Analysis

When we analyze traffic data, we should eliminate abnormal cases and abnormal data, extracting representative data only under normal circumstances without interference. Traffic during the Spring Festival and summer vacation periods is volatile, so these two periods of time should be avoided altogether. Traffic data were researched from the middle of March to June and from October to November. The original data were obtained from TDCS (18:00 on March 8, 2016 to 18:00 on March 11, 2016). According to the relevant data, statistics the arrival time and time interval of each direction train.

Due to the imbalance of the column flow in different periods of the same yard, the appearance of column flow intensity is exactly the opposite [6].

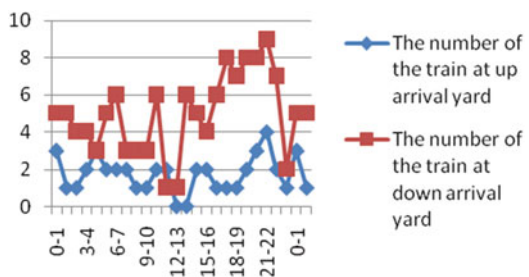
Through the above analysis, the degree of imbalance was calculated according to the number characteristics of train interval (Fig. 2).

Freight train imbalance caused by uneven idling is not an accidental phenomenon and appears in daily operations. When compiling the maintenance plan, this phenomenon should be fully considered.

According to the actual situation regarding west station column flow, one day can be divided into five different stages:

- 20–23, Peak hours. In this period, 33 trains arrived, comprising 21% of the total number of trains arriving per day, while the time occupied only 1/8 of the whole day. All the equipments in the station were in a busy state.
- 23–24, Idle time. In this time period there were very few trains, and only four columns.

Fig. 2 Train arrival frequency period at the down arrival yard



- 0–7, Ordinary time. A total of 46 columns arrived, with an average of 6.6 per hour.
 - 7–10, Idle time. 13 columns, with trains arriving at a rate of 4.3 per hour.
 - 10–20, Ordinary time. 65 columns, with trains arriving at a rate of 6.5 per hour.
- Skylight maintenance time should be set in general time and leisure time.

3.2 Method to Determine the Length of Time of Building Maintenance Mission Requirements

1. Method to determine the length of time of building maintenance mission requirements [7].
2. Network planning technique.

Construction of any node in the network graph k , using k as the starting point of the ES (k) is the earliest start time, and its meaning is to point to: node k to start work procedure are must wait before it's tight to start after completion, it expected the first time in the whole project construction ES (k) days later (Fig. 3).

In the figure, the process finishes at k . The length of the process is $D(I, k)$, and is longest from node 1 to node i . I dashed lines correspond to the longest lines.

We obtain the following recurrence formula:

$$\begin{cases} ES(1) = 0 \\ ES(k) = \max\{ES(i) + D(i, k) | i < k, (i, k) \in E\} \end{cases} \quad (1)$$

On the type of collection E is the collection of all the construction process.

Set the destination node n , then the total duration is the earliest time $Es(n)$ of the node n , sign as the Tc .

Set node k at the latest time of $LF(k)$, its meaning is to point to in the Tc don't delay the total duration of the premise, for the start node to node k process should be in the whole project starts the latest $LF(k)$ days later, with node k for the completion of the late completion time for the $LF(k)$ days (Fig. 4).

Fig. 3 Construction network diagram

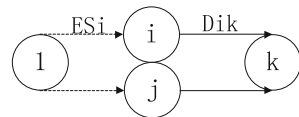
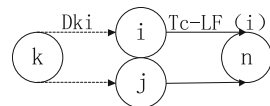


Fig. 4 Construction network diagram



In the figure, from k to ki represents the starting point of the process. Dashed lines correspond to the length of the longest line between the nodes I and n . The recursive formula is as follows:

$$\begin{cases} \text{LF}(n) = Tc \\ \text{LF}(k) = Tc - \max\{Tc - \text{LF}(i) + D(k, i) | i > k, (k, i) \in E\} \end{cases} \quad (2)$$

$$\begin{aligned} & Tc - \max\{Tc - \text{LF}(i) + D(k, i) | i > k, (k, i) \in E\} \\ & = \min\{Tc - [Tc - \text{LF}(i) + D(k, i)] | i > k, (k, i) \in E\} \\ & = \min\{\text{LF}(i) - D(k, i) | i > k, (k, i) \in E\} \end{aligned}$$

So type into the following form:

$$\begin{cases} \text{LF}(n) = Tc \\ \text{LF}(k) = \min\{\text{LF}(i) - D(k, i) | i > k, (k, i) \in E\} \end{cases} \quad (3)$$

$\text{ES}(i, k) - \text{ES}(i)$: the earliest start time of working i - k procedure;

$\text{ES}(i, k) - \text{ES}(i) + D(i, k)$: the earliest completion time of working i - k procedure.

$\text{LS}(i, k) - \text{LS}(j)$: the latest completion time of working i - k procedure.

$\text{LS}(i, k) - \text{LS}(j) - D(i, k)$: the latest starting time of working i - k procedure.

$\text{TF}(i, k) - \text{LS}(k) - \text{ES}(i) - D(i, k)$: total time of process of working i - k procedure.

$\text{FF}(i, k) - \text{ES}(k) - \text{ES}(i) - D(i, k)$: the free time of process of working i - k procedure.

The image below for the relationship between the process parameters of each time:

By using the network planning technique, it is possible to find the shortest construction time of comprehensive maintenance, electrically integrated maintenance tasks, weekly inspection of TFDS equipment, monthly check-ups of tamping THDS equipment, diversions of traffic, flaw detection and so on [8].

4 Optimization of Skylight Plan Coordination

Analysis and optimization for skylight maintenance of each yard in Jixi station from March 7–11, 2016 was undertaken and a skylight maintenance plan was devised. This plan for Jixi station from March 7–11, 2016 is shown in Table 1.

Table 1.1 March 7–11, 2016
Jixi station “skylight” plan

Date	Time	Interlocking area number
8,10	8:00–8:50	3
7,9,11	8:00–8:40	6
11	12:00–13:00	21
7,8,9,10,11	12:20–13:00	24
7,8,10,11	8:00–8:50	35
9	8:00–9:00	35
7,8,9,10,11	8:00–8:40	27
7,8,10,11	8:00–8:50	36
9	8:00–9:00	36
7,9,11	8:00–8:50	15
8,10	8:00–8:40	16
8,10	11:05–12:05	33

Use interlocking areas 4 and 5 blocked instead of interlocking area 6 blocked maintenance. On the one hand, if 3–8 are blocked, 9–13 can still pick up train operations. This does not affect the train arriving during the 8:00–9:00 a.m. interval. On the other hand, interlocking area 6 in the ‘skylight time’ electricity maintenance. Although interlocking areas 4 and 5 don’t entirely belong to different power supply section, but they can still adopt the way of the whole yard power outages maintenance. Comprehensive maintenance in interlocking areas 4 and 5 can use artificial receiving or departure trains. In the case of power failure, transport organization and building maintenance will not be interrupted.

Use interlocking areas 1 and 2 blocked instead of interlocking area 3 blocked maintenance. On the one hand, if lines 3–7 are blocked, then lines 10–13 can still go on receiving operations. This does not affect the trains arriving during the 8:00–9:00 a.m. interval. On the other hand, interlocking area 3 in the skylight time can undergo electricity integrated maintenance, since interlocking areas 1 and 2 belong to a different power supply section, and can be turned off respectively for maintenance. Interlocking areas 13 and 14 can be blocked respectively instead of interlocking area 15 being blocked during maintenance for the same reason.

The number of down arrival break up trains is 74, and the number of up arrival break up trains is 26. It is obvious that the downside hump capacity is tense. Because of the shunting operation in the whole yard is stop during the shift time, it should be used to maintain the hump equipment. Maintain the down direction hump as a priority, which is a defect of capability at the succession time (8:10–8:30). Maintain the tail of Yard IV during the period 8:00–8:40. Maintain the tail of Yard V during the period 12:20–13:00.

The flow density is 1 from 11:00 to 13:00. So skylight maintenance in the up arrival yard should be set as a priority between 11 and 13 points and set secondarily between 7 and 10 points. Skylight maintenance of the down arrival yard should be set between 12 and 14 points as a priority because there is no train between 12:00 and 14:00. It should be secondarily set between 8:00 and 10:00 or 16:00 and 19:00 (Table 2).

Table 2 March 7–11, 2016
Jixi station new “skylight”
plan

Date	Time	Interlocking area number
8	8:00–8:50	1
10	8:00–8:50	2
7	12:00–13:00	4
9	12:00–13:00	5
11	12:00–13:00	21
7,8,9,10,11	8:00–8:40	24
7,8,10,11	8:00–8:50	35
9	8:00–9:00	35
7,8,9,10,11	12:20–13:00	27
7,8,10,11	8:00–8:50	36
9	8:00–9:00	36
7,9,11	8:00–8:50	15
8	15:55–16:45	9,13
10	15:55–16:45	10,14
8,10	11:05–12:05	33
7,9,11	10:55–12:55	29

In the weekly station comprehensive maintenance “skylight” plan ensures that the maintenance time of each yard coincides with the yard’s connecting line maintenance time which traffic flow is the largest. Due to the lack of maintenance personnel, it is necessary to consider what is reasonable in terms of construction work. According to the traffic density to adjust the skylight time, in order to meet the requirements of Jixi station maintenance plan. Each yard is divided into two comprehensive maintenance areas. The plan ensures the departure (arrival) yard can keep operating throughout any 24 h period. Due to the lack of maintenance personnel, it is necessary to consider what is reasonable in terms of construction work, according to the traffic density on the skylight set period of time adjustment, in order to meet the requirements of Jixi station construction organization.

5 Conclusion

This paper researches the coordination of transport organization and building maintenance of marshalling station, analyses the condition of Jixi station, transport organization characteristic and maintenance work characteristic. This paper proposes a method of transport organization and building maintenance of Jixi marshalling station. An optimized and coordinated weekly skylight plan for Jixi station in March is proposed.

6 Acknowledgments

This work was supported by the Major Program of the National Natural Science Foundation of China (grant number 71390332) and the Specialized Research Fund for the Doctoral Program of Higher Education (grant number 20130009110001). This paper was influenced by my supervisor and as such I am very grateful for his patient instructions and correction of my mistakes. He initially inspired me to think about this problem and to improve my writing skills.

References

1. Mobley RK (2002) An introduction to predictive maintenance, 2nd edn. Butterworth-Heinemann, Oxford, pp 413–416
2. Mobley RK (2011) Maintenance fundamentals, 2nd edn. Butterworth-Heinemann, Oxford, pp 8–10
3. Budai G, Huisman D, Dekker R (2006) Scheduling preventive railway maintenance activities. *J Oper Res Soc* 57:1035–1044
4. Grimes GA, Barkan CPL (2006) Cost-Effectiveness of railway infrastructure renewal maintenance. *J Transp Eng* 132(8):601–608
5. Zhang T, Andrews J, Wang R (2012) Optimal scheduling of track maintenance on a railway network
6. Zeng M (2009) Optimization of railway transport organization work of the marshalling station. *China Railway* 12:29–31 (in Chinese)
7. Guo R (2015) Railway track maintenance maintenance planning theory and method. Beijing Jiaotong University, Beijing, pp 41–43 (in Chinese)
8. Chen H (2005) Mainly talks about the monorail sunroof open and transportation organization. *Production safety science and technology of China*, p 4 (in Chinese)

A Method to Forecast and Estimate Traffic Demand of Airport Moving Sidewalk

Yuchun Ren and Xiaoning Zhu

Abstract Large airports have the problem of long passengers' walking distance. In order to enhance the level of service and ensure passengers' experience, developing airport automated people mover (APM) system and moving sidewalk systems are the important ways to solve the problem of inter-traffic in large terminals. Moving sidewalk is often used within a terminal while APM system is often used between two terminals or among all terminals. This paper proposes a method to forecast and estimate traffic demand which is suitable for airport moving sidewalk system based on the traditional four-step method. Experimental results prove that the proposed method is reasonable and effective.

Keywords Transportation planning theory · Four-step method
Passengers flow prediction · Airport moving sidewalk system

1 Introduction

Airport automated people mover system and airport sidewalk system are essential ways to improve airport operation efficiency.

Although airport APM system has become a mature way in foreign countries, it is still at a preliminary and exploratory stage in China. In 2007, Yuan Wan proposed a planning method of airport APM system [1]. Reference [2] predicts the passenger flow in the terminal of Hongqiao International Airport. The relevant research results are applied in the internal passenger rapid transportation system planning program.

At present, four-stage predict method is still the universal theory when forecasting the transport demand [3, 4]. Transport demand forecasting models can be generally categorized into four steps including trip generation, trip distribution,

Y. Ren (✉) · X. Zhu
School of Traffic and Transportation, Beijing Jiaotong University,
No. 3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 16120875@bjtu.edu.cn

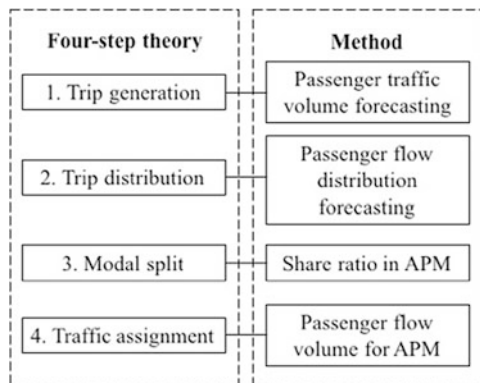
modal split, and traffic assignment. Reference [5] comprehensively explores and productively overviews the growing research field of demand forecasting in transport. The method that this paper proposed is based on the traditional four-step transportation planning theory. Lin Huang establishes a scientific and objective four stages forecast method system to forecast the rail transit’s passenger flow volume [6]. Aidi Zhou introduces some practical improved Logit models and applies these models to Ganzhou’s residents’ OD modal-split program [7]. Xiaoqiang Luo proposes a method to forecast country city urban rail transit passenger flow [8].

2 Overview of Forecasting Procedure

This paper focuses on passenger flow forecasting method for airport moving sidewalk system. Due to the different travel paths of the passengers in the terminal, the layout of airport moving sidewalk system and the distribution of the terminal passenger flow are closely related. As shown in Fig. 1, the forecasting procedure is based on traditional four steps. The details are as follows:

1. Passenger traffic volume forecasting: In this stage, this paper uses the double moving average method, the exponential smoothing method and single regression linear return to forecast method to calculate the airport passenger throughput. The results can be used as basic data.
2. Passenger flow distribution forecasting: According to the plan of airport layout, total passenger flow in every airport gallery or traffic zone is predicted.
3. Modal split: The means of transportation in the terminal can be divided into two ways—on foot or by airport moving sidewalk system. In this step, we calculate the passenger flow of those who use airport moving sidewalk system.
4. Traffic assignment: Assign rush-hour passenger flow to the different floors of airport terminal according to certain rules.

Fig. 1 Overview of the predicting procedure



3 The Models and Methods for Forecasting

3.1 Passenger Traffic Volume Forecasting

The size of the airport moving walkway system is closely linked with the passenger traffic volume. Thus, the first step of the planning of airport moving sidewalk system is to forecast airport traffic volume. This section puts forward three kinds of forecasting methods. We choose airport passenger throughput as the predictive index.

- Double moving average method: The prediction method of forecasting model is as follows:

$$\hat{T} = \left(2\bar{T}_t^{(1)} - \bar{T}_t^{(2)}\right) + \frac{2}{n-1} \left(\bar{T}_t^{(1)} - \bar{T}_t^{(2)}\right)k, \quad (1)$$

where \hat{T} is airport passenger throughput of predicted year, $\bar{T}_t^{(1)}$ is a single moving average of the t 's value, $\bar{T}_t^{(2)}$ is a double moving average of the t 's value, and n is the number of data.

- The exponential smoothing method: the prediction method of forecasting model is as follows:

$$\hat{T} = \left(2S_t^{(1)} - S_t^{(2)}\right) + \frac{\alpha}{1-\alpha} \left(S_t^{(1)} - S_t^{(2)}\right)k, \quad (2)$$

where \hat{T} is airport passenger throughput of predicted year, $S_t^{(1)}$ is exponential smoothing coefficient of t 's value, $S_t^{(2)}$ is two exponential smoothing coefficient of t 's value, α is weight coefficient, and k is the extrapolated years.

- Single regression linear return forecast method: The prediction method of forecasting model is as follows:

$$y_i = a + bx_i, \quad (3)$$

where y_i is airport passenger throughput of i 's year, x_i is the i 's year, a , and b are the coefficients.

We take the average of three kinds of the method as predicted airport passenger throughput.

3.2 Passenger Traffic Volume Forecasting

We usually divided one terminal into several traffic zones according to the layout of airport gallery. In this step, we predict the distribution of passenger flow of every traffic zone based on the model of airport gate capacity which is as follows:

$$C_{gk} = \sum_i W_i \cdot \frac{60}{t_i}, \quad (4)$$

where C_{gk} is the gate capacity per hour of zone k, W_i is the number of i gate, and t_i is gate's average occupancy time of i plane.

The types of airport gate are classified by the types of the aircraft. Gate C is suitable for aircraft C including B737, A320, A321, etc. The average seat number of aircraft C is 185. Gate D is suitable for aircraft D including B757, B767, A300, A330, etc. The average seat number of aircraft D is 386. Gate E is suitable for aircraft E including B747, A340, etc. The average seat number of aircraft E is 416. Gate F is suitable for aircraft F including B787, A380, etc. The average seat number of aircraft E is 550. Gate's average occupancy time is in accordance with the least standing time in annex 2 of "Civil Aviation Flight Normal Statistical Method". We assume the passenger load factor is 85%. The computation formula of passenger flow volume per rush-hour of each traffic zone is as follows:

$$T_k = \sum_i C_i \times N_i \times M_i \times 85\%, \quad (5)$$

where T_k is passenger flow volume per rush-hour of traffic zone k, C_i is the gate capacity per rush-hour of gate i, M_i is the full seats of plane i and N_i is the gate amount of gate i.

3.3 Modal Split

Passengers' choice of different mode is closely related to the passenger characteristics, travel characters and characteristics of transport mode. Passenger characteristics include passenger's age, gender, health, etc. Travel characters include walking distance, passenger's luggage amount, etc. Characteristics of transport mode refer to the convenience, security, punctuality, reliability, and operational speed of airport moving walkway.

The model is as follows:

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}, \quad (6)$$

where P is the probability of choosing moving walkway, $\beta_0, \beta_1, \dots, \beta_k$ is the coefficient, and x_1, \dots, x_k is the influence factors of two transport modes.

The paper uses SPSS to demarcate coefficients based on SP intention survey method. The structure of questionnaire is as follows: x1—gender; x2—age; x3—the amount of carrying luggage; x4—walking distance; x5—the degree of fatigue; y ucode Type="General_Pun">—whether passenger chooses moving walkway or not.

The paper uses SPSS to demarcate coefficients based on SP intention survey method. The structure of questionnaire is as follows: x1—gender; x2—age; x3—the amount of carrying luggage; x4—walking distance; x5—the degree of fatigue; y ucode Type="General_Pun">—whether passenger chooses moving walkway or not.

3.4 Traffic Assignment

We assume passengers include departing passengers and arriving passengers, ignoring transfer passengers. Meanwhile, the terminal is divided into departure floor and arriving floor. In this step, we assign passenger flow into departure floor layer and arriving layer. The formula is as follows:

$$T_{i,k} = T_k \cdot \omega_i \cdot \theta \quad i = 1, 2, \quad (7)$$

where $T_{i,k}$ is passenger flow volume for moving walkway of traffic zone k in the i floor, T_k is passenger flow volume for moving walkway of traffic zone k , ω_1 is the percentage of departing passengers, ω_2 is the percentage of arriving passengers and θ is the share rate of moving walkway.

4 Case Study Based on Shenzhen Bao'an International Airport

4.1 Introduction of Shenzhen Bao'an International Airport

Shenzhen Bao'an International Airport (IATA: SZX, ICAO: ZGSZ) is located in Bao'an District, Shenzhen, Guangdong, China, 32 km northwest of the city center.

Terminal 3 is divided into three parts: the main building and two large corridors—1050 m long and 650 m wide—that form a cross shape. The main building has four floors above ground and two floors underground. The main building's third floor is the waiting area for domestic and international flights, which depart from the east wing. The second floor is the arriving floor. There, we need to set moving walkway in the second and the third floor.

4.2 Case Study

The airport passenger throughput of Shenzhen Bao’an International Airport from 2007 to 2015 are 20619164, 21400509, 24486406, 26713610, 28245738, 29569725, 32268457, 36272701, 39721619 people/year. Based on these data, we calculate the airport passenger throughput of 2020 by three methods as mentioned before. The results are shown in Table 1.

As it shows in Fig. 2, Terminal 3 consists of 6 traffic zones with different types of gates. The gate capacity per rush-hour is shown in Table 2.

According to the formula (5) mentioned in Sect. 3, we can calculate passenger flow per rush-hour for each traffic zone. The results are shown in Table 3.

According to the formula (6) mentioned in Sect. 3, we can predict that the share rate of airport moving walkway is 0.9. According to “the survey report of the capital airport passenger service demand in 2013,” the percentage of departing passengers is 68% while the percentage of arriving passengers is 32%. We can approximately determine that Shenzhen Bao’an International Airport has the same passenger form. Thus, we can calculate the rush-hour passenger flow in airport moving walkway for each floor in each traffic zone. For the departing floor, the rush-hour passenger flow in airport moving walkway from traffic zone 1 to traffic zone 6, respectively, are

Table 1 Shenzhen Bao’an Airport passenger throughput predicted value in 2020

Predict method	Method 1	Method 2	Method 3	Average
Predicted value	49,746,481	47,789,782	49,727,873	49,088,045

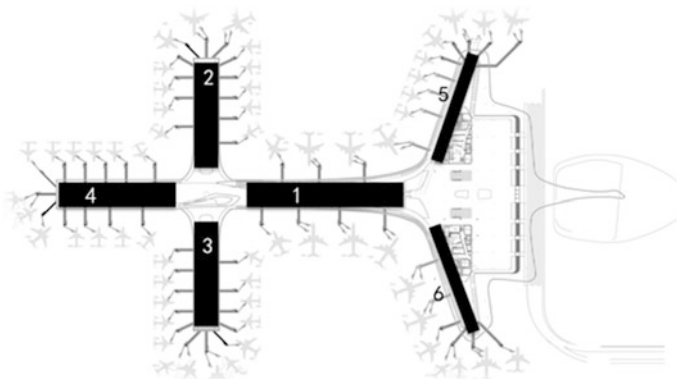


Fig. 2 Shenzhen Airport terminal 3

Table 2 The gate capacity per rush-hour

Traffic zone	Gate type	Gate amount/unit	The least standing time/minutes	Gate capacity per rush-hour/number of flight per hour
1	D	2	65	1.8
	E	5	75	4
2	C	9	65	8.3
	D	1	65	0.9
	E	2	75	1.6
3	C	8	65	7.4
	D	1	65	0.9
	F	2	120	1
4	C	10	65	9.2
	D	1	65	0.9
	F	1	120	0.5
5	C	4	65	3.7
	D	2	65	1.8
	E	3	75	2.4
6	D	2	65	1.8
	E	3	75	2.4
	F	1	120	0.5

Table 3 Rush-hour passenger flow

Traffic zone	1	2	3	4	5	6
Predicted value/person per hour	7895	13,084	10,450	14,906	5696	3603

Table 4 Rush-hour passenger flow for each floor

Traffic zone	Total value	Departing floor	Arriving floor
1	7895	5369	2526
2	13,084	8897	4187
3	10,450	7106	3344
4	14,906	10,136	4770
5	5696	3873	1823
6	3603	2450	1153

4832 persons, 8007 persons, 6395 persons, 9122 persons, 3486 persons, and 2205 persons. For arriving floor, the rush-hour passenger flow in airport moving walkway from traffic zone 1 to traffic zone 6, respectively, are 2350 persons, 3894 persons, 3110 persons, 4436 persons, 1695 persons, and 1072 persons. The results are also shown in Table 4.

5 Conclusion

With the expanding airport scale and the improving service level, the terminal is no longer a simple transfer field between air and ground anymore. It has become a large and complex transportation hub. Amplification of the single terminal area directly brings the problem of passenger's long walking distance. Meanwhile, the emergence of more-than-one-terminal brings another problem-the traffic mode between two terminals. In this background, the planning and operation of airport moving walkway system and airport automated people mover system is urgent and important. This paper proposes a procedure to forecast passenger flow volume for airport moving walkway system based on the traditional four-step predicting theory. The methods can be used by the relative department as a reference and the results can be used as basic data in the planning of airport moving walkway system.

Acknowledgments This work was supported by the Major Program of the National Natural Science Foundation of China (grant number 71390332), the Specialized Research Fund for the Doctoral Program of Higher Education (grant number 20130009110001).

References

1. Wan Y, Zhu Z, Lin C, Bamberg W (2007) A study on planning method for airport automated people mover, pp 20–26 (in Chinese)
2. Lin C, Wan Y, Zhu ZL, Ji Y (2008) The expansion project of the west terminal in Hongqiao International Airport for APM flow distribution passenger prediction research. In: Traffic and transportation, pp 113–116 (in Chinese)
3. Huapu L (2006) Theory and method in transportation planning, 2nd edn. Tsinghua University Press (in Chinese)
4. Guo X, Wang G (2006) Transportation engineering. People's press, Beijing, pp 59–70 (in Chinese)
5. Tsekeris T, Tsekeris C (2011) Demand forecasting in transport: overview and modeling advances, In: Economic research, pp 82–94
6. Huang L (2006) Research on the passenger flow forecast of Guangzhou-Zhuhai intercity rapid rail transit (in Chinese)
7. Zhou A (2003) Study on the method of model-split model. J Lanzhou Railway Univ (Nat Sci) 22(3) (in Chinese)
8. Lu X (2005) Analysis of the urban rail transit passenger traffic volume forecast (in Chinese)

Dynamic Multi-objective Optimization Problem of Container Intermodal Transport: An Empirical Analysis on the Belt and Road Initiative of China

Jue Hou

Abstract The Belt and Road Initiative of China proposes a higher request for the development of container intermodal transport. The traditional intermodal transport organization mode may change: the daily handling capacity of container railway central terminal will increase, speed of train will improve, there will be more and more double-stack container train come into use. These changes will have far-reaching consequences on the container multimodal transport development of China. This paper attempts to evaluate the economic and environmental effect on the potential change of container intermodal transport of China on the Belt and Road Initiative, with a specific focus on energy consumption and emission. Multi-objective optimization model is built to minimize transport time, cost, energy consumption, and also emissions, which is based on simulation model of locomotive, Non-Dominated Sorting Genetic Algorithm II (NSGA-II) is used to solve the multi-objective optimization model. An empirical case regarding container intermodal transport from Shanghai to Urumqi is applied to verify the validity of model in the end. The dynamic multi-objective optimization model allows the logistics enterprise to be more flexible in their decision to plan the scheme of container intermodal transport.

Keywords Intermodal transport · Multi-objective optimization problem
Emissions · Simulation · Belt and Road Initiative

1 Introduction

Railway transport plays an important role in the freight industry and it influences the related industries in China. Over the past decade, with China's domestic and foreign trade continues to expand, container in railway transport increasingly val-

J. Hou (✉)

China Waterborne Transport Research Institute, Ministry of Transport of the People's Republic of China, Xitucheng Road no. 8, Haidian District, Beijing, China
e-mail: mars8510@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_69

679

ued. Railway container transport volume increases 93.90% from 2003 to 2015. Especially in 2015, on the Belt and Road (the Silk Road Economic Belt and the 21st-Century Maritime Silk Road) Initiative, railway container transport volume is more over 20.2% than 2014. An efficient and environmentally friendly container intermodal transport will benefit social and economy, ship, and transport operators, it can reduce transportation time, cost, energy consumption, and also emissions [1–4], promoting the development of the transport system. Therefore, the Belt and Road Initiative proposes a higher request for the development of container intermodal transport of China [5, 6].

The Railway Container Center Terminal (RCCT), as the hub and the regional logistics center, is playing an important role in the modern logistics system in China, located in the important transport hub city. RCCT is modern and comprehensive freight handling space, it has a strong radiation effect on the container transport in the surrounding area. The RCCT provides services including train marshaling, container loading and unloading, storage, transmission, etc. The Railway Container Handling Stations (RCHS) is the regional logistics center, mainly located in the provincial capital city.

This paper attempts to evaluate the economic and environmental effects from the potential change of container intermodal transport of China on the Belt and Road Initiative. With the consideration of energy consumption and emissions of container intermodal. A multi-objective optimization model is built to minimize time, cost, energy consumption, and emissions of container intermodal transport, based on dynamic simulation model of locomotive. An empirical case regarding container intermodal transport from Shanghai to Urumqi is applied to test the model, heuristic algorithm (NSGA-II) is presented to solve the model. The numerical results show that the model with emission consideration in dynamic view can reduce the air pollution emissions and energy consumption in container intermodal transport, and maintain service levels.

2 Locomotive Model

2.1 Static Model

The static locomotive energy consumption model is formulated as follows:

$$Q = Q_y + Q_0 + Q_t - Q_h, \quad (1)$$

where Q is total energy consumption (kWh/kg), Q_y is traction power consumption (kWh/kg), Q_0 is coasting and braking and stop power consumption (kWh), Q_t is entry section, and the way of shunting operation power consumption (kWh/kg), Q_h is the power consumption of braking back to power grid (kWh/kg).

2.2 Dynamic Simulation Model

In order to get more accuracy in calculation of locomotive energy consumption than static model. A dynamic model is presented based on the simulation platform. The electric-locomotive model is built based on HXD_{2B} locomotive’s electric structure and its control algorithms, which is widely used in container intermodal transport of China, including main circuit and auxiliary circuit [7]. Based on MATLAB/Simulink (2015a), this paper built the simulation model of HXD_{2B} electric locomotive. Four HXD_{2B} electric locomotives in the simulation model. Figure 1 shows a dynamic simulation model that includes traction power supply system, traction substation (transformer), traction contact network and control block, etc., which is in accordance with the actual situation [8].

Time–speed curve and time–power consumption curve, as shown in Fig. 2a, b.

3 Dynamic Multi-objective Optimal Model

3.1 Parameter Description

V is set of all nodes (the RCCT or RCHS), except destination node.

V_1 is set of all nodes (the RCCT or RCHS), except origin node.

N is the transport mode, $N = \{1, 2\}$, $N = 1$ is by truck, $N = 2$ is by train.

i, j is index of node (the RCCT or RCHS), $i, j \in V$.

k, l is index of transport mode, $k, l \in V$.

$d_{i,i+1}^k$ is distance from the RCCT or RCHS i to the RCCT or RCHS $i + 1$ by transport mode k .

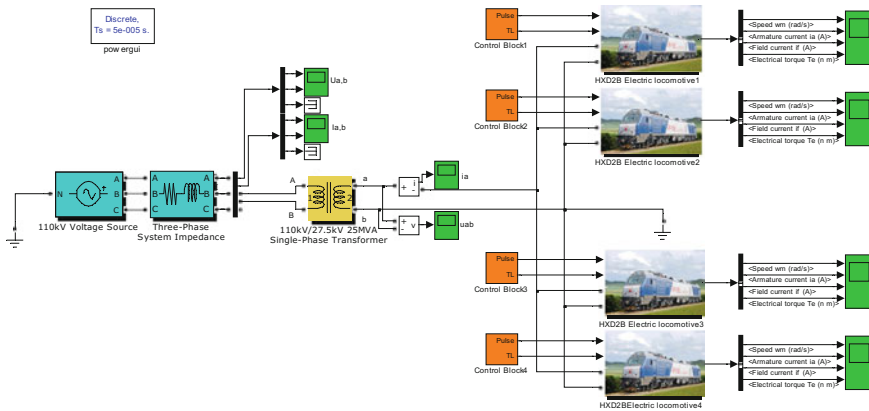


Fig. 1 Diagram of electric-locomotive simulation model

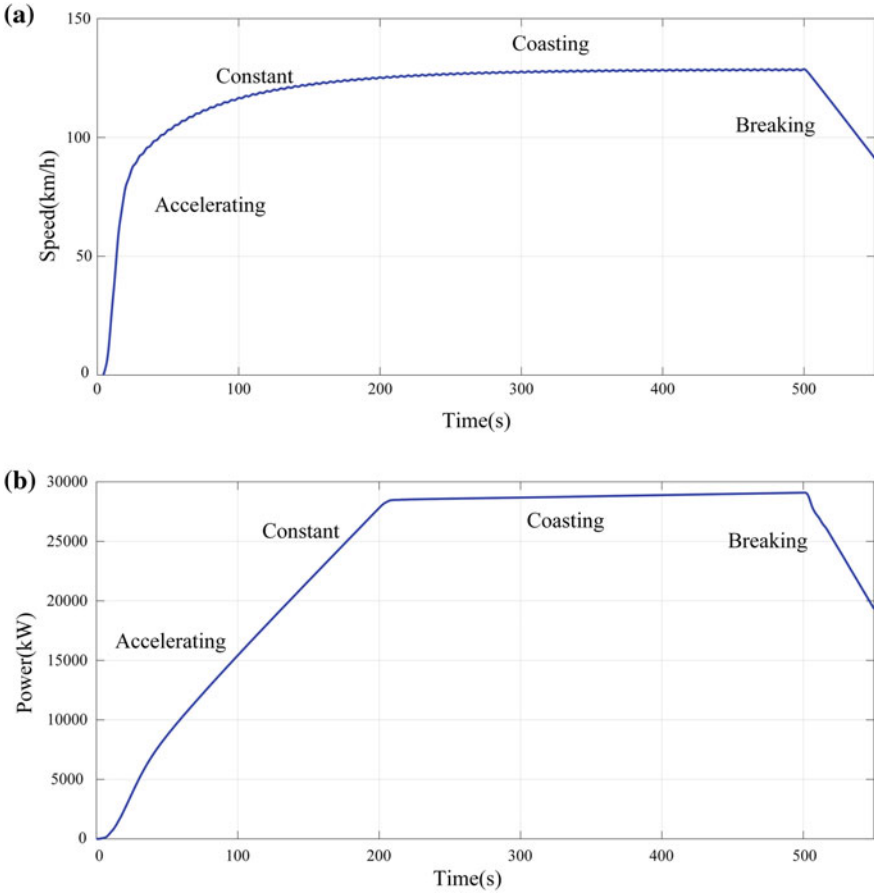


Fig. 2 a Diagram of time–energy consumption curve of electric locomotive. b Diagram of time–speed curve of electric locomotive

c_k is cost by mode k per distance (km) per container (20 ft or 40 ft), $k \in N$.
 t_k is time by mode k per distance (km) per container (20 ft or 40 ft), $k \in N$.
 e_k is energy consumption by mode k per distance (km) per container (20 ft or 40 ft), $k \in N$.
 g_k is emissions by mode k per distance (km) per container (20 ft or 40 ft), $k \in N$.
 $c_i^{k,l}$ is transport cost of transfer the mode k to mode l on the RCCT or RCHS i per container (20 ft or 40 ft).
 $e_i^{k,l}$ is transport energy consumption from the RCCT or RCHS i to the RCCT or RCHS $i + 1$ by transport mode k , in this paper, $e_i^{k,l}$ is decided by simulation model.
 g_k is transport emissions from the RCCT or RCHS i to the RCCT or RCHS $i + 1$ by transport mode k .

$t_{i,i+1}^k$ is transport time from the RCCT or RCHS i to the RCCT or RCHS $i + 1$ by transport mode k .

$s_i^{k,l}$ is transfer time of transfer the mode k to mode l on the RCCT or RCHS i per container (20 ft or 40 ft).

a is lower limit of time (hour), b is upper limit of time (hour), α, β is dynamic factor of ratio.

$$x_{i,i+1}^k = \begin{cases} 1 & \text{1 if train or truck travels from node } i \text{ to node } j, \\ 0 & \text{0 otherwise.} \end{cases}$$

$$r_i^{k,l} = \begin{cases} 1 & \text{transport mode changed from } k \text{ to } l \text{ on node } i, \\ 0 & \text{0 otherwise.} \end{cases}$$

3.2 Modeling

Multi-objective includes three parts: minimization of time, cost, energy consumption, and emission of container transport. The dynamic multi-objective optimization model is formulated as follows:

$$\min f_1 = \sum_{i \in V} \sum_{k \in N} c_k d_{i,i+1}^k x_{i,i+1}^k + \sum_{i \in V_1} \sum_{k \in N} c_i^{k,l} r_i^{k,l} + p(t) \tag{2}$$

$$\min f_2 = \sum_{i \in V} \sum_{k \in N} t_k d_{i,i+1}^k x_{i,i+1}^k + \sum_{i \in V_1} \sum_{k \in N} t_i^{k,l} r_i^{k,l} \tag{3}$$

$$\begin{aligned} \min f_2 = & \alpha \left(\sum_{i \in V} \sum_{k \in N} e_k d_{i,i+1}^k x_{i,i+1}^k + \sum_{i \in V_1} \sum_{k \in N} e_i^{k,l} r_i^{k,l} \right) \\ & + \beta \left(\sum_{i \in V} \sum_{k \in N} g_k d_{i,i+1}^k x_{i,i+1}^k + \sum_{i \in V_1} \sum_{k \in N} g_i^{k,l} r_i^{k,l} \right) \end{aligned} \tag{4}$$

subject to:

$$\sum_{k \in N} x_{i,i+1}^k = 1, \quad \forall i \in V \tag{5}$$

$$\sum_{k \in N} \sum_{l \in N} y_{i,i+1}^k = 1, \quad \forall i \in V_1 \tag{6}$$

$$p(t) = \begin{cases} b \cdot M \cdot (t - T_2) & t > T_2 \\ 0 & t \geq T_2 \end{cases} \tag{7}$$

$$h(t) = \begin{cases} a \cdot M \cdot (T_1 - t_2) & t < T_1 \\ 0 & t \geq T_1 \end{cases} \quad (8)$$

$$x_{i,i+1}^k, y_i^{k,l} \in \{0, 1\} \quad \forall i \in V, k \in N, l \in N \quad (9)$$

$$0 \leq a \leq 1, \quad b = 1 - a, \quad (10)$$

where Eq. (2) is objective function of minimizing transport cost. Equation (3) is objective function of minimizing transport time. Equation (4) is objective function of minimizing transport energy consumption and emissions. Equation (5) makes sure there is only one kind mode to complete container transport between the RCCT or RCHS. Equation (6) makes one transport mode to another on the RCCT or RCHS be only once. Equation (7) is inventory costs with time window. Equation (8) is penalty costs with time window. Equation (9) is integer variable. Equation (10) is the weight to decide the important one between energy consumption and emissions.

4 Empirical Study

4.1 General Settings

An empirical case of container intermodal transport from Shanghai port to Urumqi RCCT is taken. With the Belt and Road Initiative, Shanghai to Urumqi container intermodal becomes more important to strengthen China and Central Asian-Europe trade and business, as shown in Fig. 3. Logistics enterprise needs to

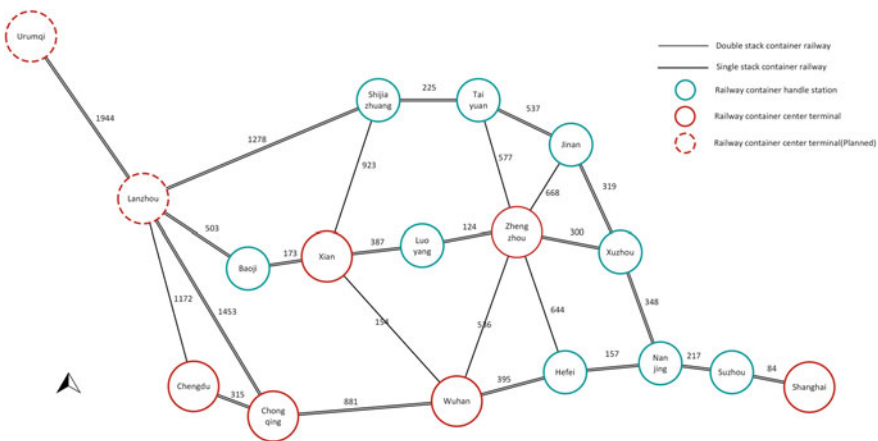


Fig. 3 Diagram of route between RCCH and RCHS from Shanghai to Urumqi

design the transport scheme carefully, to make themselves competitive and provide better service for shippers. With the environment protection pressure, the transport mode of each trip needs to be decided to minimize the total costs and time, besides, energy consumption and emissions need to be taken into consideration.

This paper presents four modes for container railway intermodal transport from Shanghai to Urumqi. Mode 1 and mode 3 are mainly container railway intermodal transport mode in 2016. Mode 2 and mode 4 will be the container railway intermodal transport mode in 2018, as shown in Table 1. This paper takes DF_{4B} diesel locomotive and HXD_{2B} electric locomotive which are both widely used in railway container transport of China. The freight data of container in railway transport [9]. Emission factor and fuel consumption calculation are obtained using reference [10], in order to minimize transportation distance from Shanghai to Urumqi, Vehicle Route Problem (VRP) algorithm is adopted by this paper [11].

4.2 Approach Description

Non-Dominated Sorting in Genetic Algorithms (NSGA) is a popular non-domination based genetic algorithm which is a very effective algorithm but lacks elitism and for choosing the optimal parameter value. This paper adopts Non-Dominated Sorting Genetic Algorithm II (NSGA-II) for solving the multi-objective optimization problem [12, 13], which has a better sorting algorithm, and incorporates elitism and no sharing parameter needs to be chosen a priori. The main differences between the NSGA and NSGA-II are the usages of (1) a fast non-dominated sorting method, (2) elitism mechanism; and (3) a parameter-less niching approach.

4.3 Empirical Test Result

The multi-objective optimization model is formulated by MATLAB 2015a, and runs on a personal computer with Intel Core i7 CPU and 16 GB RAM. The crossover probability is 0.8, mutation probability is 0.2, number of generations is

Table 1 Container intermodal transport mode in empirical study

Mode	Origin	Destination	Combination	Containers (per train)	Groups (per train)
1	Shanghai	Urumqi	HXD _{2B} + Single layer	98	49
2	Shanghai	Urumqi	HXD _{2B} + Double stack	168	42
3	Shanghai	Urumqi	DF _{4B} + Single layer	84	42
4	Shanghai	Urumqi	DF _{4B} + Double stack	152	38

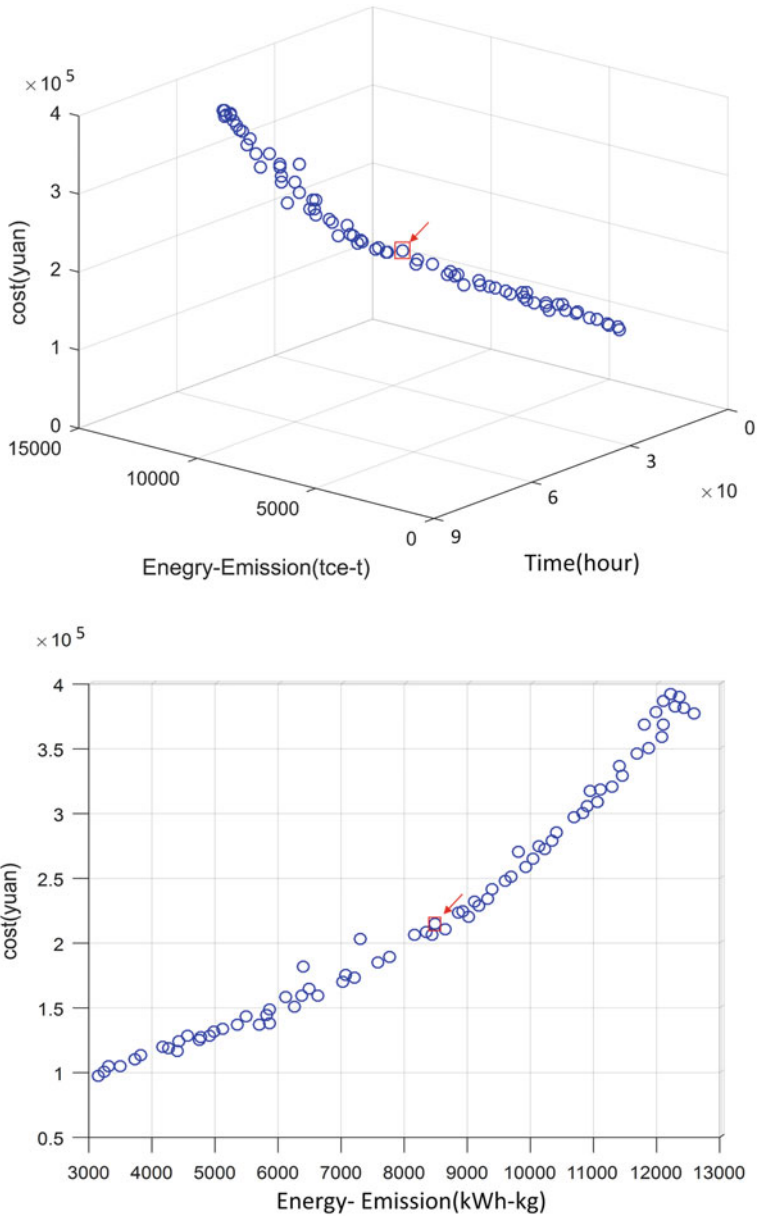


Fig. 4 Pareto-optimal fronts of mode I of empirical test

1000, and population size is 100. This set of multi-objective model resulted (Pareto-optimal fronts) in the solution of mode 1 as shown in Fig. 4. Choosing one group of the data in the Pareto set (red outline rectangle) to analyze the optimization result.

Table 2 Empirical result of different mode of container intermodal transport

Mode	Time (h)	Length (km)	Energy cost TEU/yuan	Energy consumption per TEU/kg		Emissions per TEU/kg			
				Power (kWh)	Fuel (kg)	CO ₂	NO _x	SO _x	PM
1	58.13	4080	607.95	1085.63	0	282.33	1.32	0.02	0.08
2	53.29	3627	364.77	651.38	0	169.40	0.79	0.01	0.04
3	62.81	4080	929.30	0	129.07	491.50	9.83	0.31	0.40
4	57.03	3627	513.29	0	71.29	271.46	5.43	0.17	0.22

The mode 1 and mode 2 both can reduce air pollution emissions in container intermodal transport, meanwhile, maintain service levels, and with the increase of utilization ratio of double stack, which allows the terminal operator to be more competitive on the Belt and Road Initiative. According to adjust the different weights between 0 to 1 in Eq. (10) and evaluate the implementation effect (Table 2).

5 Conclusions

This paper presents a multi-objective optimization model to minimize time, cost, energy consumption, and emissions of container intermodal transport, meanwhile, in order to improve the accuracy of locomotive energy consumption, dynamic simulation model of locomotive has been built. The numerical results show that the model allows the logistics enterprise to be more flexible in their decision to plan the transport scheme of container intermodal transport. There are also some limitations in this study. First, the container intermodal model in this study does not consider sensitivity analysis with price of energy and freight. Second, with the rapid development of waterborne transport, which has advantages of low freight and emission, this paper does not take it into consideration. Overcoming these limitations could be the topics for future research.

Acknowledgements This study was supported by the Conserve Energy and Reduce Emissions Grant (CEREG) No. 2015-JNJP-010-056 and No. 2014-JNJP-008-037 from the Ministry of Transport of the People’s Republic of China.

References

1. Reis V, Meier JF, Pace G et al (2013) Rail and multi-modal transport. *Res Transp Econ* 41(1):17–30
2. Liu D, Yang H (2012) Dynamic pricing model of container sea-rail intermodal transport on single OD line. *J Transp Syst Eng Inf Technol* 12(4):122–127

3. Yang H (2015) Joint slot allocation and dynamic pricing of container sea-rail multimodal transportation. *J Traffic Transp Eng* 2(3):198–208
4. Rodemann H, Templar S (2014) The enablers and inhibitors of intermodal rail freight between Asia and Europe. *J Rail Transp Plann Manage* 4(3):70–86
5. Huihui MO, Wang J, Song Z (2015) Economically suitable areas of China's transnational container transport by land in the silk road economic belt. *Prog Geogr* 34(5) (in Chinese)
6. Xue YF (2016) The legal guarantee thought of the container logistic transportation smooth in the "one belt one road": taking Lianyungang port as an example. *J Shandong Youth Univ Polit Sci* (in Chinese)
7. Zhou L, Jia Wei LI, Feng QB et al (2011) Bogie of HXD2B electric locomotive. *Electric Drive Locomotives* (in Chinese)
8. Ministry of Rail (2013) HXD2B electric locomotive. China Railway Publishing House (in Chinese)
9. Gould G, Niemeier D (2009) Review of regional locomotive emission modeling and the constraints posed by activity data. *Transp Res Rec J Transp Res Board* 2117(2117):24–32
10. Cordeau JF, Desaulniers G, Desrosiers J et al (2001) VRP with time windows. The vehicle routing problem. *Soc Ind Appl Math* 157–193
11. Mendes RS, Wanner EF, Sarubbi JFM et al (2016) Optimization of the vehicle routing problem with demand responsive transport using the NSGA-II algorithm. In: *IEEE, international conference on intelligent transportation systems*. IEEE, 2016, pp 2657–2662
12. Lamamra K, Belarbi K, Belhani A et al (2014) NSGA2 based of multi-criteria decision analysis for multi-objective optimization of fuzzy logic controller for non-linear system. *J Next Gener Inf Technol* 5(1):57–64
13. Wei J, Qiao J, Meng Q (2015) An improved NSGA2 algorithm with the adaptive differential mutation operator. In: *Control conference*. IEEE, 2015, pp 2633–2638

The Impact of Carbon Abatement Policies on Port Intermodal Freight Transportation Routing and Cost

Shenghua Wang, Qi Zhang and Wenyuan Wang

Abstract The rapid development of port intermodal freight transportation indicates the high demand for door-to-door transportation services. Satisfying such demand would bring high profits to logistics industry as well as produce great amounts of greenhouse gas emissions. The latter phenomenon is becoming the main threat to our environment. This paper developed an integrated optimization model to solve the intermodal transport routing problem. Furthermore, in order to investigate the effect of the carbon abatement policies (i.e., carbon tax and carbon credit) on optimal route, transportation cost, and carbon emissions, we adapted the original model by, respectively, considering each of the policies. Based on the illustrative example, the result shows that the carbon tax policy does not make an obvious influence on the optimal route with total and make the transportation cost increase by 1.7%. When the carbon credit policy is conducted, the carbon emissions reduce by about 1/3 and the transportation cost increases by 9%. Meanwhile, shippers are inclined to choose railway or water transportation to abide by the carbon emission limit.

Keywords Intermodal transportation · Routing optimization · Carbon abatement policy

1 Introduction

Increase in international trades and investments has become the driving force of port logistics development. Higher demands of door-to-door transportation services lead to the upgrade of the whole transportation system. Port intermodal freight transportation regards the original port as the main hub, delivers and transfers cargoes with multiple transportation modes. Collection and distribution capacity of

S. Wang · Q. Zhang · W. Wang (✉)
Faculty of Infrastructure Engineering, Dalian University of Technology,
Dalian, China
e-mail: wangwenyuan@dlut.edu.cn

intermodal transportation has become the key of promoting port handling capacity. As a result, many related researches have been conducted.

Reddy [1] proposed that the total transportation cost is the sum of the travel cost between cities besides the intercity transfer cost. Ziliaskopoulos [2] presented a time-dependent intermodal optimum path algorithm for multimodal transportation networks. Chang [3] solved the optimum routes selection of multi-objective multimodal multi-commodity flow problem (MMMFP) based on time windows and concave costs. He [4] improved the Dijkstra labeling algorithm to deal with the shortest path cost problem in container motor rail multimodal transportation. Wang [5] developed a shortest path model with time constraint by chat transform. Zhang [6] proposed a generalized shortest path method to solve the intermodal optimal transport path problem. Lei [7] developed a genetic algorithm to solve the path optimization model of the multimodal transport for long and bulky cargo. Xin [8, 9] discussed the multimodal hazardous materials transportation routing problem considering time-varying. Sun [10] dealt with the uncertainty and environment pollution of multimodal transport with a chance-constrained programming model. Chen [11] improved the Dijkstra algorithm to solve the multimodal transport routing model with the soft time windows restriction.

New troubles arose with the development of intermodal transportation, for instance, the energy consumption and greenhouse gases emissions. Due to the relatively rapid development of port capacity, large-scale equipment and busy traffic have produced numerous greenhouse gases, which gradually become a burden to our environment. To deal with such issues, some countermeasures, such as directed regulation policy, carbon tax policy, and carbon credit policy are put forward. In practice, the latter two policies provide more flexibility as well as incentive to the shippers, are more widely used. Since the lack of related regulations in our country as well as the absence of comparative analyses between these two policies, it's difficult to predict the final effects caused by the policies. In the view of the above, this paper develops an intermodal routing optimization model considering the influence of fixed departure time of port intermodal freight transportation. Furthermore, we discuss the different influence patterns of two carbon abatement policies on intermodal transportation routing and cost.

2 Modeling

2.1 Problem Description

The generally optimal routing problem of intermodal transportation can be described as below: a batch of cargoes needs to be delivered from origin node O to the destination node D . There are several network nodes, the set of which is denoted as I . Among these nodes, there are K types of different transportations available which differ in transport capacity, time, and cost. When transferring in a particular node occurs, such operation would consume an amount of time and the corresponding transfer cost.

To consider actual situations of port logistics, different transportation plans have their own departure time. With a given schedule, arriving too early would cause opportunity cost due to the time waiting for departure. On the other hand, not arriving on time would influence a normal delivery, or even disorder the latter transportation which might cause a high penalty cost. Moreover, the optimum routes of intermodal transportation differ with the constraint of different carbon abatement policies. Carbon tax would raise the cost of every leg of the transportation and raise the total transportation cost as a result. Besides, carriers have to rearrange their transport routes within the emission limit in the case of carbon credit condition. In this paper, we would analysis the differences of the influence on intermodal transportation between these two policies by adding constraints to the intermodal routing optimization model.

2.2 Intermodal Freight Transportation Network

Every leg of the intermodal transportation network has different transportation mode. Nodes between those legs have their own transfer mode as well. The original network has been extended for describing the network properly and more conveniently: Splitting all nodes into input and output ends except for node O and node D. M input nodes would be set up if there are M types of transportation mode between two nodes. Same amounts of legs are also set up to match different transportation mode. Output ends would be operated in the same way. Legs between input and output nodes (dotted line) are also set up, which stands for particular transfer modes (e.g., highway-railway, railway-water, depending on the transportation mode of input end and output end). The procedure of network conversion is shown in Fig. 1.

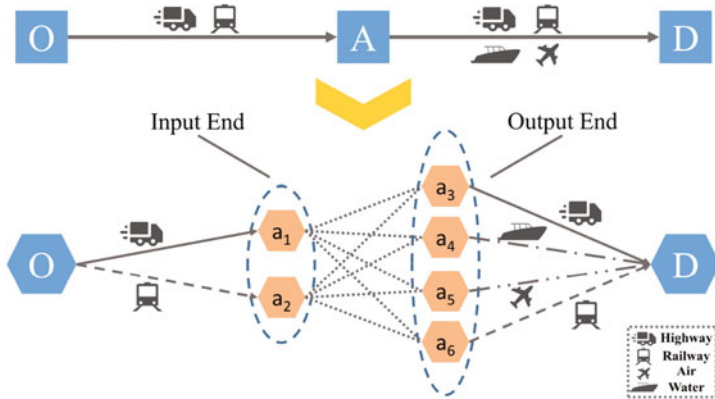


Fig. 1 Procedure of network conversion

2.3 Assumptions and Variables Description

2.3.1 Assumptions

For the simplicity of discussion, some hypotheses are made below:

1. The same batch of cargoes is indivisible during the whole transportation.
2. Transferring would only occur within a node.
3. Distances between nodes remain the same regardless of transportation modes.

2.3.2 Variables Description

Decision Variables: $x_{i,i+1}^m$ is a binary variable, if the cargo is being delivered between node i and $i + 1$ through the m th transportation mode, $x_{i,i+1}^m = 1$; or $x_{i,i+1}^m = 0$. y_i^{mn} is a binary variable, if the cargo is being transferred from the m th transportation mode to the n th transportation mode within the node I , $y_i^{mn} = 1$; or $y_i^{mn} = 0$.

Other Variables: Z denotes the total transportation cost. F_p denotes the fixed charge from handling the cargo p . $v_{i,i+1}^p$ denotes the quantity of cargo p being delivered between node i and $i + 1$ through the m th transportation mode. $l_{i,i+1}^m$ denotes the distance between node i and $i + 1$ through the m th transportation mode. $c(p, m)$ denotes per unit travel cost delivering cargo p through the m th transportation mode. d_p denotes the quantity of cargo p being transferred within a transit node. $c(m, n, p)$ denotes per unit transfer cost transferring from the m th transportation mode to the n th transportation mode when cargo p arrives at node i . t_{wi} denotes the waiting time before departure within node i . $c(A_i + t_i^{mn})$ denotes the opportunity cost causing by the waiting within a particular node.

2.4 Routing Optimization Model of Intermodal Freight Transportation Considering Carbon Abatement Policies

In practice, not all kinds of transfer mode exist. By banning some irrational transfer modes, an alternative transfer collection would be obtained which can reduce the scale of intermodal transportation network [12]. Also, we consider the diversity of travel cost between different cargoes. The objective function (OF) is shown below

$$\min Z = C_{\text{fixed}} + C_{\text{transport}} + C_{\text{transit}} + C_{\text{opportunity}} \quad (1)$$

$$\begin{aligned}
 C_{\text{fixed}} &= \sum_p F_p \\
 C_{\text{transport}} &= \sum_i \sum_p \sum_m x_{i,i+1}^m v_{i,i+1}^{pm} l_{i,i+1}^m c(p, m) \\
 C_{\text{transit}} &= \sum_i \sum_p \sum_m \sum_n y_i^{mn} d_p c(m, n, p) \\
 C_{\text{opportunity}} &= \sum_i \sum_m \sum_n y_i^{mn} t_{wi} c_w (A_i + t_i^{mn})
 \end{aligned}
 \tag{2}$$

s.t.

$$\sum_m x_{i,i+1}^m = 1 \quad \forall m \in M, \forall i \in I
 \tag{3}$$

$$\sum_m \sum_n y_i^{mn} = 1 \quad \forall n \in M, \forall i \in I
 \tag{4}$$

$$\sum_p \sum_m x_{i,i+1}^m v_{i,i+1}^{pm} \leq \text{Cap}_{i,i+1} \quad \forall p \in P, \forall m \in M, \forall i \in I
 \tag{5}$$

$$x_{i-1,i}^m + x_{i,i+1}^n \geq 2y_i^{mn} \quad \forall m, n \in M, \forall i \in I
 \tag{6}$$

$$t_{wi} = \max \{ S_i^n - (A_i + t_i^{mn}), 0 \} \quad \forall m, n \in M, \forall i \in I
 \tag{7}$$

$$\begin{aligned}
 &\sum_i \sum_m x_{i,i+1}^m t_{i,i+1}^m + \sum_i \sum_m \sum_n y_i^{mn} [\max \{ S_i^n \\
 &- (A_i + t_i^{mn}), 0 \}] \leq T \quad \forall m, n \in M, \forall i \in I
 \end{aligned}
 \tag{8}$$

$$\sum_p y_i^{mn} d_p < \text{Cap}_i^{mn} \quad \forall p \in P, \forall i \in I
 \tag{9}$$

$$x_{i,i+1}^m \in \{0, 1\} \quad \forall m \in M, \forall i \in I
 \tag{10}$$

$$y_i^{mn} \in \{0, 1\} \quad \forall m, n \in M, \forall i \in I
 \tag{11}$$

$$v_{i,i+1}^{pm} \geq 0, d_p \geq 0, t_{wi} \geq 0 \quad \forall i \in I, \forall p \in P, \forall m \in M
 \tag{12}$$

Objective function (1) is applied to minimize the total transportation cost which consists of four parts. C_{fixed} defines the fixed charge of delivering cargoes. $C_{\text{transport}}$ defines the travel cost produced during delivering. C_{transit} defines the transfer cost produced within the nodes. $C_{\text{opportunity}}$ defines the opportunity cost produced by waiting for departure. Constraint (3) specifies every two nodes have only one particular transportation mode. Constraint (4) specifies only one transfer mode can be chosen once within a node. Constraint (5) ensures the quantity of cargo delivering should no more than the capacity of one leg. Constraint (6) specifies when transportation mode transfers from the m th mode to the n th mode, cargo is delivered

through the m th mode between node $i - 1$ to i and through the n th mode between node i to $i + 1$. Constraint (7) ensures the waiting time should be more than 0. Constraint (8) ensures the whole transportation time should no more than the limit. Constraint (9) is the constraint of transfer capacity. Constraints (10), (11), and (12) enforce integrality conditions on some variables.

Considering the carbon tax policy, Constraint (13) is added to the original objective function:

$$C_{\text{tax}} = CE \times \sum_p d_p \left(\sum_i \sum_m e^m l_{i,i+1}^m x_{i,i+1}^m + \sum_i \sum_m \sum_n e^{mn} y_i^{mn} \right) \tag{13}$$

When considering the carbon credit policy, we add Constraint (14) as the constraint to the original objective function.

$$\sum_p d_p \left(\sum_i \sum_m e^m l_{i,i+1}^m x_{i,i+1}^m + \sum_i \sum_m \sum_n e^{mn} y_i^{mn} \right) \leq E \tag{14}$$

3 An Illustrative Example

In this section, a numerical example is provided to analyze the impact on the optimal routing of intermodal transportation from different carbon abatement policies. First, we supposed there is a batch of cargoes delivered from port O to destination D over an intermodal transportation network shown in Fig. 2. Node A, B, C, E, and F are transit nodes while the arcs between these nodes are travel legs including railway, highway, water, and Air transportation.

When converting the original network, non-exist transfer mode should be banned in advanced with respect to the practical condition. In this paper, we ban two types of transfer mode which are {Railway—Air} and {Water—Air}. Figure 3 shows the extended network considering the alternative set of transfer node. Moreover, network distances matrix is shown in Table 1.

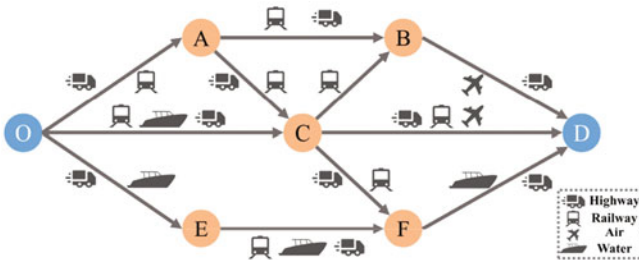


Fig. 2 An illustrative network of port intermodal freight transportation

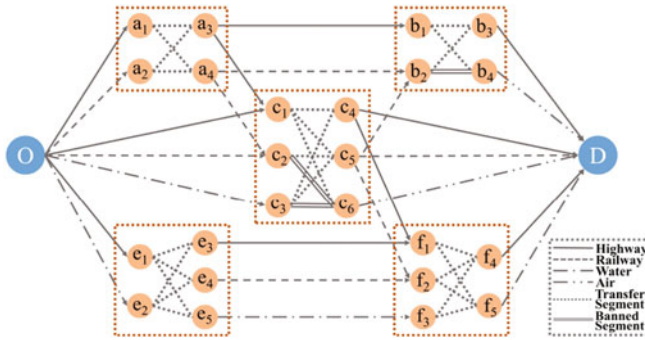


Fig. 3 Extended network considering alternative set of transfer node

Table 1 Network distances matrix (km)

OD	O	A	B	C	E	F	D
O	–	238	–	476	209	–	–
A	238	–	395	265	–	–	–
B	–	395	–	142	–	–	329
C	476	265	142	–	–	206	543
E	209	–	–	–	–	347	–
F	–	–	–	206	347	–	286
D	–	–	329	543	–	286	–

We assumed three different cargoes P_1 , P_2 , and P_3 that are to be delivered from O to D with a quantity of 35, 20, and 15 tons, respectively. The fixed charge of them is 600, 800 and 1200 CNY, respectively. Travel cost for each leg and transfer cost for each node are shown in Tables 2 and 3. According to some survey data, we assumed the average speed of every transportation mode, which are: Highway—85 km/h, Railway—55 km/h, Water—35 km/h, Air—300 km/h.

Table 2 Unit travel cost for different transportation modes [CNY/(km·t)]

Cargo	Highway	Railway	Water	Air
P_1	0.7	0.4	0.2	2.0
P_2	0.8	0.5	0.2	2.0
P_3	0.9	0.5	0.2	2.2

Table 3 Unit transfer cost for difference transfer modes (CNY/t)

Cargo	Highway–Railway	Highway–Water	Highway–Air	Railway–Water
P_1	0.2	0.1	0.5	0.3
P_2	0.4	0.2	0.6	0.3
P_3	0.3	0.2	0.7	0.3

Due to the departure schedule of railway, water, and air transportation, cargoes have to wait for departure within transit node when transferring occurs. Departure schedule is shown in Table 4. Waiting time within nodes would produce opportunity cost for our regulation. And the price is set as 10 CNY/(h·t).

In this example, delivery would start at 6:00 p.m. With the initial conditions above, we program to compute the optimal route of the original condition. The result shows that the minimum total transportation cost is 56,358 CNY with a time of 16.43 h. The optimal route is $O \rightarrow c_1 \rightarrow c_5 \rightarrow D$, or interpreted as Highway—(Highway → Railway)—Railway. The optimum is shown in Fig. 4a.

Furthermore, we extended the OF as constraint (13) as to consider the influence caused by carrying out carbon tax policy. Since the detailed charge of carbon tax remains unknown in China, we set the charge standard as the upper bound of recommendation from Ministry of Environment Protection (10–20 CNY/t) [13]. Then, such charge standard would be equivalent to 0.02 CNY/kg. The carbon emission factors are calculated according to a long-established equation as below:

$$m(CO^2) = V(\text{petrol}) \times 2.7 \tag{15}$$

Detailed carbon emission factors are listed in Table 5.

With the initial conditions unchanged, New optimal route for extended model is the same as the original one ($O \rightarrow c_1 \rightarrow c_5 \rightarrow D$) with a slightly higher cost of 56,453 CNY. And the total transportation time is 16.43 h remains the same. In this

Table 4 Departure schedule information

Nodes	Railway	Water	Air
O	8:00, 14:30	8:00, 14:00	–
A	14:00, 19:00	–	–
B	–	–	19:30
C	15:30, 19:45	–	13:30
E	15:00, 17:00	15:30, 19:00	–
F	–	05:30, 18:15	–

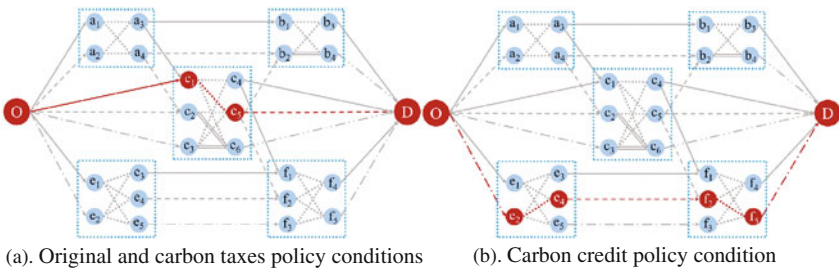


Fig. 4 Optimal route for different conditions

Table 5 Carbon emission factors

	Highway	Railway	Water	Air
Carbon emission factor 1 (kg/(t·km))	0.12	0.02	0.015	0.35
	Highway– Railway	Highway– Water	Highway– Air	Railway– Water
Carbon emission factor 2 (kg/t)	0.13	0.11	0.16	0.12

condition, the carbon emissions are 4768 kg and the corresponding charge is only 95.4 CNY. Such cost is not able to change the optimal route.

In the case of carbon credit policy, we add constraint (14) to the original model. The carbon credit is set as 3500 kg. Latest optimal route is $O \rightarrow e_2 \rightarrow e_4 \rightarrow f_2 \rightarrow f_5 \rightarrow D$, or interpreted as Water—(Water → Railway)—Railway—(Railway → Water)—Water. The total transportation cost is 61,452 CNY with a time of 20.47 h, and the corresponding carbon emissions are 3207 kg. Comparing with the last two optimums above, the carbon emissions are significantly abated in this case, though the transportation cost and time slightly rise up. The result is shown in Fig. 4b.

Table 6 shows three optimums considering different conditions. From this table, it is obvious that both two carbon abatement policies influence the original optimum, respectively. For carbon tax policy, total transportation cost rises up due to the taxation. However, such tax rate is not able to influence the optimal route as well as the total transportation time. On the other hand, with the limitation of emissions, carbon credit policy success in influencing the optimal route. Moreover, such limitation would raise the total transportation cost and time. Besides, there is one more difference between these two policies. In the case of carbon tax policy, there are not restrictions for transport shippers with relatively high-carbon emissions except paying carbon tax. Which means that carbon emissions would beyond control since the expense is not able to influence the transportation routes. As for carbon credit policy, shippers would have to rearrange their transportation strategy according to the emission limitation. Such outcome would not only lower the

Table 6 Optimums for different conditions

	Original	Carbon taxes	Carbon credit
Cost (CNY)	56,358	56,453	61,452
Time (h)	16.43	16.43	20.47
Carbon emission (kg)	–	4768	3207
Optimal route	$O \rightarrow c_1 \rightarrow c_5 \rightarrow D$	$O \rightarrow c_1 \rightarrow c_5 \rightarrow D$	$O \rightarrow e_2 \rightarrow e_4 \rightarrow f_2 \rightarrow f_5 \rightarrow D$

carbon emissions, but even upgrade the entire intermodal transportation system. Furthermore, promoting the importance of low-carbon transport pattern as well. In other words, putting the carbon credit policy into practice brings positive effects to both the environment and the intermodal transportation industry.

4 Conclusions

The impact of carbon credit policy on port intermodal freight transportation is more significant comparing with the carbon tax policy. Based on the standard of 0.02 CNY/kg for carbon emissions charging, the result shows that the proportion of carbon tax in total transportation cost is around 1.7‰ in the case of carbon tax policy. For carbon credit policy, in the scenario with a limitation of total 3500 kg carbon emissions, the results show that the emissions reduce by 33%, with total transportation cost raises less than 1/10. Such policy would do great help to control greenhouse gases as well as save fossil fuels. Restricting the carbon emissions, the port intermodal freight transportation routing would be inclined to choose low-power-consuming transportation modes such as railway and water transportation. In summary, the carbon credit policy may promote the development of low-power-consuming transportation mode as well as enhance the transportation capacities.

Acknowledgements This work is supported by National Natural Science Foundation of China (No. 51309049 and No. 51279026).

References

1. Leiding Y, You Y, Zhangying G, Pizhi D (2014) Path optimization model and algorithm of multimodal transportation for long and bulky cargo. *J Traffic Transp Eng* 01:75–83 (in Chinese)
2. Zhangyun H, Linbo L, Liang D, Gaohong Y (2006) Research on a generalized shortest path method of optimization intermodal transportation problems. *J China Railway Soc* 04:22–26 (in Chinese)
3. Reddy VR, Kasilingam RG (1995) Intermodal transportation considering transfer costs. In: *Proceedings of the 1995 global trends conference of the academy of business administration*, Aruba
4. Chendan D, Hong W, Jia Y (2015) Dynamic paths optimization of multimodal transport with stochastic factors. *J Chongqing Jiaotong Univ (Nat Sci)* 02:112–117 (in Chinese)
5. Sun B, Chenqiu S (2013) The routing optimization for multi-modal transport with carbon emission consideration under uncertainty. Paper presented at the proceeding of the 32nd Chinese control conference, July 26–28, 2013, Xi'an, China (in Chinese)
6. Xinchun L, Fengqian R, Zhangjian W (2016) Problem of distribution center location-routing optimization for multi-modal hazardous materials transportation. *China Safety Sci J* 09:73–78 (in Chinese)

7. Chen L, Linbo L, Wang L, Wenxu H, Li J (2015) Optimization model of mode selection for intermodal transportation based on carbon-reduction policy. *J Beijing Jiaotong Univ* 03:70–75 (in Chinese)
8. Wangqing B, Hanzeng X, Jiming J, Liyan M (2011) Path optimization of container multimodal transportation based on node operation randomness. *J Transp Syst Eng Inf Technol* 06:137–144 (in Chinese)
9. Xinchun L, Fengqian R, Zhangjian W (2016) Routing optimization for multi-modal hazardous materials transportation under a time-varying condition. *China Safety Sci J* 06:104–110 (in Chinese)
10. Liu J, Heshi W, Song R, Lihao D (2011) Study on optimization of dynamic paths of intermodal transportation network based on alternative set of transport modes. *J China Railway Soc* 10:1–6 (in Chinese)
11. Chang T (2008) Best routes selection in international intermodal networks. *Comput Oper Res* 35(9):2877–2891. <https://doi.org/10.1016/j.cor.2006.12.025>
12. Heguo X (2006) Computer aided algorithm of the cost weighted shortest path problem in container motor rail multimodal transportation. *J China Railway Soc* 01:1–5 (in Chinese)
13. Ziliaskopoulos A, Wardell W (2000) An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays. *Eur J Oper Res*

Improved Genetic Algorithm Based Passenger Flow Control in Subway

Man Jiang

Abstract This paper is in the background that improvement of organization for urban rail transit operation along with its rapid development is an urgent need. Aimed at solving the organization problem of mass passenger flow, passenger flow control in subway is analyzed. To increase the control quality and efficiency, the method of improved genetic algorithm is studied. Finally, the algorithm is applied to a case study to verify its validity.

Keywords Subway · Passenger flow control · Improved genetic algorithm
Train available capacity adjustment · Inbound passenger flow limiting

1 Introductions

Current researches on passenger flow control in the field of traffic and transportation focused on passenger flow organization in stations and train operation management.

Some literature analyzed the characteristics of urban rail transit passengers and structural characters of the network and made measures of passenger flow control according to these findings. Qiaomei [1] researched on the characteristics of mass passenger flow, influence factors of station operation, and relative organization method. Lu [2] used complex network-based transmission dynamics model to develop the crowding propagation model of burst passenger flow, analyzed the propagation characteristics of passenger flow, and proposed measures to respond to the passenger flow.

Some researches proposed station organization principles to solve the problems of inbound and outbound passenger flow. Jianlin [3] studied the contradiction between passenger volume and transportation capacity in Line 6 and 8 of Shanghai Subway, and changed the stations to apply measures of passenger flow control

M. Jiang (✉)

Transport Planning and Research Institute, China Railway Design Corporation,
No. 10 Minjiang Road, Hebei District, Tianjin, China
e-mail: applyitforu@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_71

701

according to principles of choosing control stations and time. Lianhua [4] analyzed the problems of increased passenger volume and complex transfers in Guangzhou Subway, calculated to obtain the control stations and volumes, and made relative passenger flow control strategies.

Some literature analyzed passenger flow regularity between two stations and made passenger flow control schemes. Yingsong [5] proposed a parameter calculation method based on priority for station coordination control, which provided a theoretical foundation for reasonable schemes of passenger flow control. Zheng [6] further put forward the idea of combined control in several subway stations.

Some researches adjusted train operation scheme to relieve the contradiction between passenger demand and line transportation capacity. Suh [7] applied stop-skipping service to Seoul subway in Korea and decided the stop-skipping stations by predicted OD. Sun [8] adopted nonlinear integer programming to realize the stop-skipping strategy in real-time operation. Cortés [9, 10] adopted the mixed strategy of holding and stop-skipping to control public traffic lines to reach the target of minimum waiting time and uniform headway.

This paper will combine two strategies of inbound passenger flow limiting and train available capacity adjustment to manage subway operation synthetically and improve the passenger flow conditions in peak hours.

2 Passenger Flow Control in Subway Stations

Passenger flow crowding results from real-time mismatching between train available capacity and passenger flow demand. So efforts should be made in two aspects: inbound passenger flow limiting and train available capacity adjustment.

- (1) Inbound passenger flow limiting. In the condition of large passenger flow, the control strategy to limit passengers outside the station entrances can be made to slow down inbound passenger speed and control redundant passengers outside the subway system. In the demand peak hours, large passengers want to go inside the stations and get a train to start their trip in short time and thus overcrowding often occurs and decreases passenger inbound efficiency and volume. What's more, there may be potential danger and accidents in the special condition. Purposive and sequential passenger flow control can guarantee passenger safety and promote passenger inbound efficiency.
- (2) Train available capacity adjustment. In the peak hours, trains usually run at the minimum headway in subway, which means that available transportation capacity for lines has reached the maximum, while passenger flow in crowded stations is still unsatisfied. Responding to this kind of situation, the strategy of train available capacity adjustment for different stations is practicable, which can be realized by the adjustment of train operation scheme. It can change the train stopping frequency at stations in the peak hours to reserve train available capacity for the stations with large passenger demand and thus relieve the pressure of the platform organization.

3 Algorithm Implementation

3.1 Improved Genetic Algorithm

In the application of genetic algorithm, researchers gradually found out that traditional methods had many shortcomings, which could not satisfy the need for solution in quality and speed. These shortcomings included: (1) there existed subjectivity and experience in coding of decision variables, fitness function, inheritance, and parameter setting; (2) evolution results depended on initial values of parameters and population to some extent; (3) inappropriate inheritance might lead to missing of population variability; (4) convergence speed decreased along with the increase of coding length of decision variables. To overcome these shortcomings, many improved genetic algorithms, incorporating improvement for coding way, heritance operators, control parameters, and operation strategies have been put forward.

The solving algorithm of this paper adopts improved genetic algorithm and the improvements include the adoption of order crossover in the process of individual crossover and the improvement of mutation sites in the process of individual mutation.

3.2 Algorithm Process of Passenger Flow Control

The improved genetic algorithm for passenger flow control in subway stations is implemented in the following steps:

- Step 1: Set initial train operation scheme. Demarcate passenger parameters (passenger arrival rates in stations and alighting rates from trains), station parameters (inbound capacity and platform holding capacity), and train parameters (train capacity, train running and holding time standards). Set the population scale, maximum cycling times, crossover way, and mutation way for the algorithm.
- Step 2: Realize the mutual constraints of passengers, stations, and trains. Mutual constraints of passengers and stations include passengers entering stations in the condition of inbound passenger flow limiting and passengers waiting for trains at platforms. Mutual constraints of trains and stations are expressed in the process of passengers boarding and alighting from trains. Constraints of trains incorporate train capacity, holding time, and running time.
- Step 3: Output current train operation scheme and inbound passenger flow limiting results and calculate the fitness function, which pursues the maximum boarding passenger number and minimum variance for boarding ratios. If the current solution is better than the existing solution, use the current solution to update the existing solution. Otherwise, keep the existing solution unchanged. Update i with $(i + 1)$, and enter the next step.

- Step 4: Judge if the cycling times have reached its maximum limit I. If it is, the final train operation scheme, passenger flow limiting solution, and objective function values will be output. Otherwise, enter the processes of heritance, crossover, and mutation and produce the next generation. Heritance refers to keeping some of the routings in the current train operation schemes, crossover means to mix some routings and get a new generation and mutation is to choose and mix two routings stochastically in unexplored schemes to get a new generation. Then enter the next step.
- Step 5: Begin a new scheme algorithm calculation and enter Step 2. When the cycling of the above steps ends, the algorithm ends and the optimal scheme and control results will be given as output. The flow-chart is shown in Fig. 1.

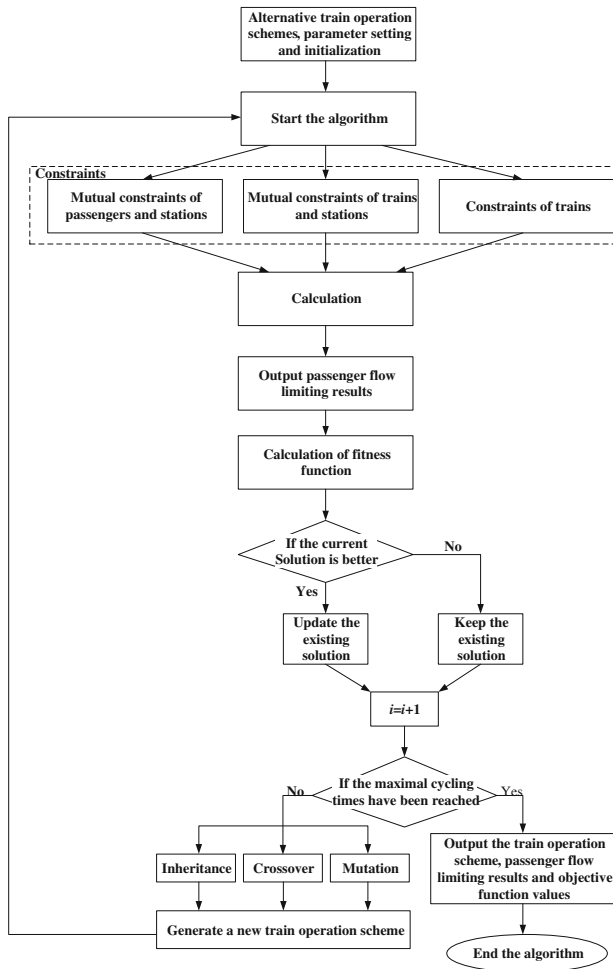


Fig. 1 Algorithm flow chart for passenger flow control in subway stations

The above algorithm depends on the platform of Visual Studio with the help of the programming language C#. Its interface of parameter setting and running results is shown in Fig. 2. Based on different solving efficiency and quality, population scale and cycling times can be set manually. For the process of crossover in genetic algorithm, crossover way and sites can be chosen freely. As to mutation, its probability can be changed to ensure the variability of solving effects. In all, improved genetic algorithm can facilitate summarizing the rules of solutions and finding the optimal solution rapidly.

4 Case Study

4.1 Parameter Setting

The case study chose a section in Line 3 of Guangzhou Subway, where stations locate as Lijiao (LJ), Datang (DT), Kecun (KC), Guangzhouta (GZ), Zhujiangxincheng (ZJ), and Tiyuxilu (TY) in the down direction. Passenger flow demand in each station in the morning peak hour is listed in Table 1.

Holding time of trains at each station, running time of trains on each section, arrival rates, and alighting rates of passengers are shown in Table 2.

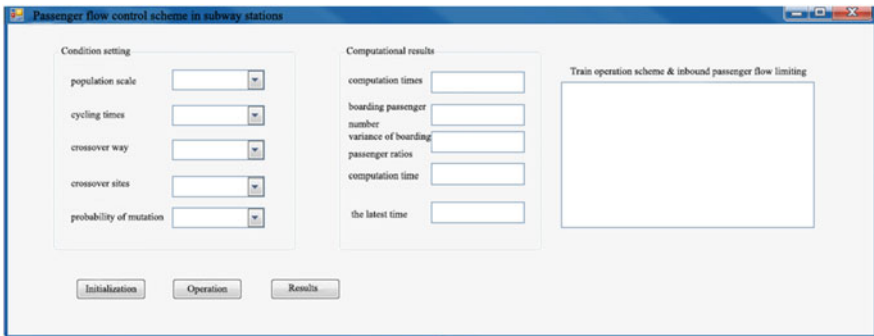


Fig. 2 System interface of condition setting and results

Table 1 Passenger flow demand in each station

Station	Passenger flow volume	Sequence
LJ	19,200	6
DT	23,400	5
KC	65,850	3
GZ	30,930	4
ZJ	70,740	1
TY	68,670	2

Table 2 Basic data of trains and passengers

Station		LJ	DT	KC	GZ	ZJ	TY
Train	Holding time (min)	0.5	0.5	1	0.5	1	1
	Running time to the next station (min)	2.5	2	1.5	2	2	
Passenger	Arrival rates (persons/min)	320	370	637	205	645	649
	Alighting rates (%)		6.3	68.8	3.7	72.2	58.3

Train headway is 2 min in the peak hour. Train capacity standard is 1350 persons. When the passengers are overcrowded, the maximal train capacity is 1882 persons.

4.2 Solving Results and Analysis

These parameters and initial settings are input to the solving algorithm system. The optimal train operation scheme is shown in Table 3.

Table 3 The optimal train operation scheme

Train NO.	Train operation scheme
1	LJ-KC-ZJ-TY
2	LJ-KC-ZJ-TY
3	LJ-DT-KC-GZ-ZJ-TY
4	LJ-KC-ZJ-TY
5	KC-ZJ-TY
6	LJ-DT-KC-GZ-ZJ-TY
7	LJ-DT-KC-GZ-ZJ-TY
8	LJ-DT-KC-GZ-ZJ-TY
9	KC-GZ-ZJ-TY
10	LJ-KC-ZJ-TY
11	KC-GZ-ZJ-TY
12	LJ-KC-ZJ-TY
13	KC-ZJ-TY
14	KC-GZ-ZJ-TY
15	KC-GZ-ZJ-TY
16	KC-GZ-ZJ-TY
17	KC-ZJ-TY
18	KC-GZ-ZJ-TY
19	LJ-KC-ZJ-TY
20	LJ-DT-KC-GZ-ZJ-TY
21	LJ-DT-KC-GZ-ZJ-TY
22	LJ-DT-KC-GZ-ZJ-TY
23	LJ-DT-KC-GZ-ZJ-TY

Table 4 Inbound passenger flow limiting schemes in each station

Station	Passengers limited in the peak hour (persons)	Average passenger flow speed after control (persons/min)
LJ	4386	73
DT	3775	63
KC	5605	93
GZ	3893	64
ZJ	5986	100
TY	5230	87

In this scheme, all boarding passengers are up to 51,525 persons on the line in the peak hour and the variance of boarding ratios at each station is 0.03. The control scheme of passenger flow is shown in Table 4. In the peak hour of passenger flow, limiting volume and speed at each station can be determined and thus control strategies are made coordinately to realize the objectives of the maximal boarding passengers and balanced inbound service.

5 Concluding Remarks

In this paper, two kinds of strategies, which are train available capacity adjustment and inbound passenger flow limiting, are analyzed to optimize the organization of large passenger flow in subway stations. To make detailed strategies in certain passenger flow condition in an efficient and high-quality way, improved genetic algorithm is utilized to solve the problem. An application case of line 3 in Guangzhou Subway is studied to verify the validity of the algorithm. In the future, the algorithm will be applied to more practical cases and thus relative modification can be made to promote the algorithm solving process.

References

1. Qiaomei T (2013) Research on the organization methods in urban rail transit under mass passenger flow. Southwest Jiaotong University (in Chinese)
2. Lu W (2013) Analysis of unexpected passenger demand and resulting congestion in urban rail transit. Southwest Jiaotong University (in Chinese)
3. Jianlin L (2011) Practice research on passenger transportation adjustment of passenger flow control in Shanghai rail transit. Mod Urban Rail Transit 4:81–83 (in Chinese)
4. Lianhua L, Jiang L (2005) Research on the method of passenger flow control in urban rail transit network. Railway Transp Econ 33(5):51–55 (in Chinese)
5. Yingsong H (2013) Analysis and research on the control measures of passenger flow in rail transit stations. Beijing Jiaotong University (in Chinese)

6. Zheng Z, Xi J, Yingsong H (2013) Research on safe and synergistic control of passenger flow in urban rail transit stations in the peak hours. *J Safety Sci Technol* 9(10):5–9 (in Chinese)
7. Suh W, Chon KS, Rhee SM (2002) Effect of skip-stop policy on a Korean subway system. *Transp Res Rec* 1793:33–39
8. Sun A, Hickman M (2005) The real-time stop-skipping problem. *J Intell Transp Syst* 9(2):91–109
9. Cortés CE, Saéz D, Alejandro T (2011) Integrating short turning and deadheading in the optimization of transit services. *Transp Res Part A* 45(6):419–434
10. Cortés CE, Saéz D, Milla F et al (2010) Hybrid predictive control for real-time optimization of public transport system's operations based on evolutionary multi-objective optimization. *Transp Res Part C* 18(5):757–769

Study on Protection from Stray Current in the Metro System

Hongbo Cheng, Zikang Xiao, Xun Wang and Nannan Sun

Abstract Aiming at solving an issue that the stray current increases the risk of corrosion of the buried metal with the increase of operation age in metro, some measures are proposed to protect the buried metal from stray current respectively on the stages of design, construction, operation, and protection in the metro system in this paper. Compared with the reference papers, CDEGS, professional software of grounding analysis, is applied for modeling as well as calculation, and factors are applied to decrease the leakage of stray current with regard to the electrical specification, like different supply modes, voltage classes, and so on. And when the metro is in operation, some measures, like the dust removal, waterproof, and passivation protection should be applied to reduce the risk of corrosion of buried metallic structures. Therefore, with the measures adopted, the risk of corrosion will be reduced greatly.

Keywords Risk of corrosion · CDEGS · Leakage of stray current
Buried metallic structures

1 Introduction

In recent years, construction of the subway has been developing rapidly in most provincial capital cities and some developed cities in China. A report shows that the number of metro cities is up to 31 and the mileage of these cities in total is up to 4000 km in China until January 2017. As the longitudinal resistance of the rail and the leakage of conductance exists, a small portion of traction current, which leaks from the return rail, is called stray current. And most of the stray current penetrates into the concrete reinforcement or buried metallic conductors. And the corrosion on

H. Cheng (✉) · Z. Xiao · X. Wang · N. Sun
School of Electricity and Automation, East China Jiaotong University (ECJTU),
Nanchang, China
e-mail: waitingbo@126.com

© Springer Nature Singapore Pte Ltd. 2018
L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_72

the surface of pipelines will cause the leakage or explosion of pipelines, which endangers the security of lives and properties of nearby residents. And the report shows that the inner structure reinforcement is severely corrosive along the first line of Beijing Metro since it has operated for 48 years [1]. Similarly, the gas pipeline caused by stray current corrosion was found perforated, which results in an accident of gas leakage in Hong Kong MTR [1].

A great number of experts and scholars have devoted themselves to research on protection against stray current in the advanced countries. Scholar M.M. Alamuti etc. built a distributed model to simulate the actual metro system, and the principle of stray current was found that the leakage of traction current is due to the voltage drop [2]; Expert Charalambous studied the evaluation of stray current leakage, total leak charge, and polarization potential of buried metals corrosion according to European standards EN 50122-2 and 50162 based on a topological model [3].

The construction of the first line of metro was taken relatively late in China, but stray current of metro had been paid more and more attention recently. A lot of experts and scholars devoted to study on regulation of stray current distribution and monitoring system of stray current. Guoli Wang [4] established three-dimensional model of metro system through Finite Element Method (FEM), and obtained the result that the total leakage of stray current in the rectangle tunnel is smaller than in shield configuration tunnel. Scholars Shujuan Li, etc., put forward that the existing monitoring system of stray current is mainly composed of slave computers, made up of a series of sensors and microcontroller with data acquisition module, and host computers. And last, the stray current is obtained by process and analysis of data [5]. Scholar Levin, etc., proposed a new method of monitoring stray current and confirmed the feasibility of monitoring stray current with optical fiber sensing technology through experiment [6]. With the further study of monitoring systems, these systems had been produced in abroad electrical companies, like Siemens, and domestic company, like Hang [7].

In this paper, some measures are applied to protect the buried metal from stray current respectively on the stages of design, construction, operation, and protection in the metro system. Compared with the reference papers, CDEGS, professional software of grounding analysis, is applied for modeling as well as calculation, and factors in regard to the electrical specification, like different supply modes, voltage classes, and so on, are applied to decrease the leakage of stray current. And when the metro is in operation, quite a few measures, like passivation protection, should be applied to reduce the risk of corrosion of buried metallic structures.

2 Reaches on the Stage of Design in Metro

The reference paper shows that in the stage of design of metro, the distance of traction substation, the application of the drainage network, the number of locomotives, and so on will affect the amplitude of stray current. However, this paper studies the different leakages of stray current based on the model of different

voltage levels or different resistivity and thickness of the concrete reinforcement or different power supply modes.

- (1) Currently, the DC traction power supply modes including third rail in 750 V and catenary in 1500 V are the main power supply modes in urban metro. The domestic standard GB 50157-2013 indicates that the suspension height of a catenary is not less than 4070 mm. Different from the catenary, the third rail gets close to the ground and only sets the installation beyond the third rail. And the schematic diagram of power supply system in catenary mode is shown in Fig. 1.

On the early stage of operation, due to the resistivity of insulation layer, the total leakage of stray current in third rail is similar to the one in catenary. However, with the operation years increasing, the resistivity of insulation is decreasing under the running rail and the third rail. And the leakage of stray current will be greater than the one in catenary, even two times as much as the leakage of stray current in the catenary mode as shown in Fig. 2.

- (2) DC 750 and 1500 V are the common voltage classes in the power supply system of metro. In order to meet the increasing demand of public, it is necessary to increase the number of locomotives at a time. The operation power is 2.25 kW and the operating current are 3000 and 1500 A, respectively, in 750 and 1500 V lines are assumed in this paper. From Fig. 3, the conclusion we can get that the leakage current of drainage network in 750 V line will be two times as much as the ones in 1500 V (Fig. 3).
- (3) The concrete reinforcement is close to the insulation layer and the rail, which may affect the distribution of stray current. Resistivity and thickness of concrete reinforcement are the key parameters in electricity. And the amount of stray current, which leaks from the drainage network, is shown in Figs. 4 and 5, respectively.

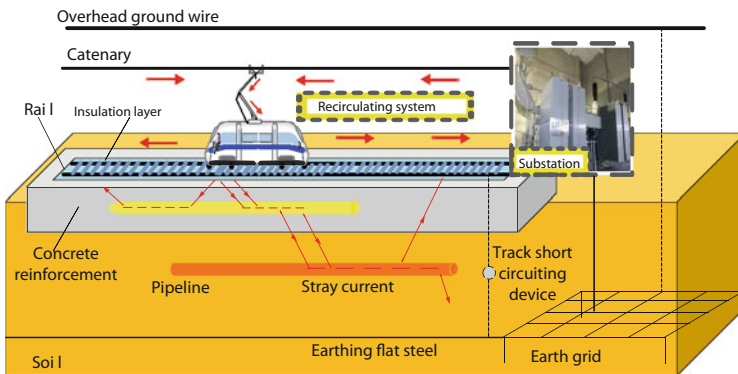


Fig. 1 The traction power supply system in the catenary mode

Fig. 2 The density of current leakage from drainage network

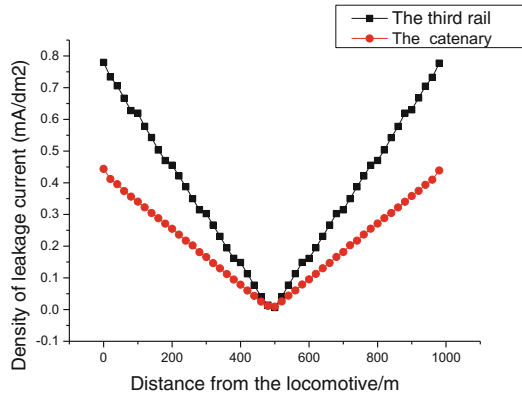
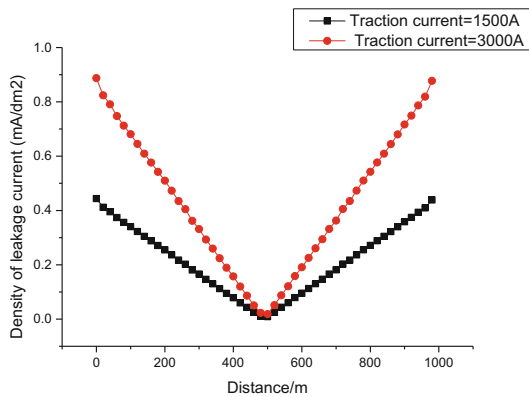


Fig. 3 The density of current leakage from the drainage network



From Fig. 4, we can find that the current leakage is decreasing with the rise of concrete’s resistivity. From Fig. 5, we can get the result that the thickness of concrete reinforcement makes little effect to the amount of leakage current. Above all, with the permit of cost and technology, it should be given great priority to applying 1500 V DC power supply for there is less traction current in 1500 V DC power supply under the condition of the same traction power. And the resistivity, but the thickness of concrete reinforcement plays an important role in the amount of leakage of stray current and some measures or the application of new materials in concrete reinforcement.

3 Reaches on the Stage of Construction in Metro

- (1) In order to reduce the stray current leakage, the seamless rail should be welded in low resistance. Seamless rails are in common use in metro, and the program

Fig. 4 The amount of current leaked from the drainage network affected by resistivity

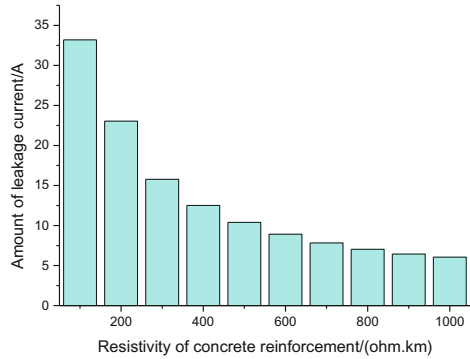
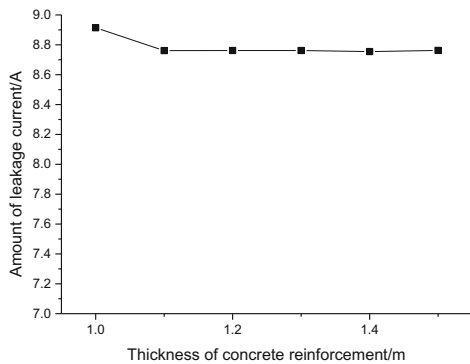


Fig. 5 The amount of current leaked from the drainage network by thickness



of manufacture and connection is as follows. First, 500-m-long rails are manufactured in steel factory at home and production 25 m, and fundamental welding tasks are completed. Second, 500-m-long rail will be transported to the sites and welding on the spot, and eventually cast into 1 or 2 km long rails. The longitudinal resistance of the rail depends on the quality of solder joint in Rail junctions. If any glitch or faulty welding occurs in the solder joints, on the one hand, resistivity of rail connections will sharply increase and most of traction current will leak to the soil from here, which will increase the amount of stray current leakage. On the other hand, it is likely to cause the rail slide, which increases the risk of unsafe operation of trains. Therefore, the characteristics of solder joints should be smooth, free of burrs, and virtual welding, so as to ensure the solid welding, which is a key link of protection from stray current on the stage of construction.

- (2) In order to reduce the amount of leakage, the performance of tunnel waterproof and drainage should be improved, which is able to reduce the risk of insulation layer affected by groundwater. Humidity is one of key factors to the resistivity of insulation layers. The place that is used for the construction of the subway tunnel should give priority to selecting sites with a certain slope, which

improves the performance of drainage and reduce the risk of standing pools of water in the tunnel. In the construction of the line and the erection of viaduct, some measures should be adopted, such as waterproof layer should be adopted on the surface of bridges and concrete structure. And drainage pipes should not be close to the surface of concrete structures and their diameter not be less than 150 mm, which enhances the effect of waterproofing and drainage.

4 Reaches on the Stage of Operation in Metro

Through addition of monitoring points, the amount of stray current leakage along the subway will be monitored online. If the amount of leakage exceeded relative criteria, some measures should be adopted at once. The technical specification CJJ49-1992 of protection from stray current corrosion in metro suggests that monitoring points m should be added in place adjacent to the signal acquisition on the station platform, adjacent to locomotive acceleration or breakpoints of rails, the ends of the bridges or subway lines. There are following five reasons for the addition of monitoring points. First, a series of cables and pipelines access to station was easily affected by stray current generated by a great number of air conditioners, escalators, large apparatus of drainage, sewage pump, and so on in the station. Second, in the acceleration process of locomotive, the value of transient surge current produced by locomotive is 1.7–2.5 times as much as normal traction current, which results in increasing the amount of stray current leakage. Third, the resistivity of solder joint is relative large, which results in increasing the amount of stray current leakage. Fourth, stray current may leak to nearby bridge of Pier and corrode steels embedded in the pier for the structure steel embed in roadbed between a bridge and the bridge nearby is not overall. Last, stray current cannot flow back to the station under no drainage net at the end of the line. And it may cause corrosion on the surface of underground metal structure adjacent to the end of the line.

5 Reaches on the Stage of Maintenance in Metro

- (1) With the increase of operation, the resistivity of the insulation layers will decrease and the amount of leakage stray current will increase to a great extent as shown in Fig. 5. And humidity and particle density are important factors in determining the life of insulation layer with high resistivity [8]. If the rail and roadbed are cleaned regularly, the life of insulation layers with high resistivity will be extended. Therefore, in addition to good drainage systems, to guarantee the cleanliness of insulation layer and concrete through clean is essential. Some effective measures should be completed as follows. One is that the cleaning

work should be completed monthly to keep the lines clean when the metro line comes into service. Second, improving the performance of dust removal by the way of adopting the advanced dust removal technology, like the improved suction inlet of cleaning car in metro [9]. Compared with the traditional apparatus, this new cleaner has the characteristic of high speed, low pressure of exhaust pipe, low consumption of energy, which makes dust particles float from the surface of insulation layers to the dust port, and then the negative pressure makes it easier for dust particles to float from the dust nozzle into the exhaust pipe, so as to enhance the performance of dust removal along the lines.

- (2) Reference to power supply equipment of AC electrified railway, such as the booster transformer is applied to force stray current return to the cathode at the traction station, so as to enhance the backflow coefficient of the metro. The aim of the booster transformer applied in AC electrification railway is to turn asymmetric two-phase power supply system into symmetric three-phase power supply system, which can reduce the interference from the power supply system to communications equipment adjacent to lines. And in DC traction power supply system, the concatenation of the primary side of the transformer and catenary, secondary winding and rail, which will force the value of current along the rail equals to the one of catenary through the electromagnetic coupling between the windings of transformers, so as to reduce the leakage of stray currents.
- (3) If the monitoring stray current exceeds the value expressed in the design codes, some appropriate metal protective measures should be applied. There are two ways of reducing the corrosion of buried metal pipelines. One is passivation protection and the other is activation protection. The principle of passivation protection is to reduce the amount of stray current penetrated into buried metals. Some specific measures are applied, such as pipeline coat with paint or insulating film. The object of activation protection is the metal structure corroded by stray current, and the principle of it is to force stray current along the metal pipe or force metal pipeline to act as the cathode. Some specific protection measures includes connecting buried metallic pipelines and the running rail and adding strong metals such as zinc, magnesium, aluminum, and other materials as anodes, and addition of external DC power supply.

6 Conclusion

This paper suggests that some measures should be adopted in four stages of metro. And it should be given great priority to the 1500 V DC catenary power supply, for it can decrease the amount of leakage of stray current as much as possible from the source in the stage of design of metro. And the resistivity of concrete reinforcement plays an important role in the amount of current leakage and some measures or the application of new materials in concrete reinforcement. And in order to extend the

life of insulation layers with high resistivity, dust removal and waterproof should be enhanced in the other three stages. Compared with the traditional measures on protection, these effective measures can decrease the risk of corrosion from the source.

Acknowledgements This work is supported by Jiangxi Provincial Natural Science Foundation (No. 20143ACB21019) and (No. GJJ13355). Hongbo Cheng is the corresponding author.

References

1. Zhou X, Gao B (1999) Experimental study on metro stray current corrosion of rebars in reinforced concrete. *West-China Exploration Engineering* (in Chinese)
2. Alamuti MM, Nouri H, Jamali S (2011) Effects of earthing systems on stray current for corrosion and safety behaviour in practical metro systems. *Iet Electr Syst Transp* 1(2):69–79
3. Charalambous CA, Aylott P (2014) Dynamic stray current evaluations on cut-and-cover sections of dc metro systems. *IEEE Trans Veh Technol* 63(8):3530–3538
4. Wang G, Pei X (2014) Three-dimensional finite element simulation of different subway tunnels under stray current fields. *J Railway Sci Eng* 6:85–91 (in Chinese)
5. Li S, Wang D, He R et al (2009) The monitoring scheme about buried field pipeline stray current based on SCADA system. *Pipeline Technol Equip* 1:19–21 (in Chinese)
6. Li W, Cai W, Wang Y et al (2008) Monitoring method of metro stray current based on an optical fiber sensing technique. *J China Univ Min* 37(6):848–853 (in Chinese)
7. Hang D (2016) Development of metro stray current monitoring system. *Southwest Jiaotong University* (in Chinese)
8. Lu C, Li F (2013) Study on the protection from corrosion by stray current in metro systems. *Urban Rapid Rail Transit* 26(1):64–67 (in Chinese)
9. Yang C (2014) Research on particle removal performance of vacuum dust suction mouth. *Central South University* (in Chinese)

Energy Consumption Analysis of High-Speed Maglev Train

Xiaochun Zhang, Siyuan Mu and Jinsong Kang

Abstract In this chapter, the finite element model of the linear synchronous motor is simulated and analyzed by the analysis of the high-speed maglev train. The dynamic model of the train is established from the force analysis's point of view. In addition, for the minimum energy consumption of the traction strategy, the minimum energy consumption model is solved by taking the actual situation of the project into account. The traditional traction strategy only considers one target variable, but this chapter takes the minimum time model of the train into account while considering the energy consumption, and combines the two traction strategies and models to get a new traction strategy. And an algorithm is designed to get the optimal train running process by combining the actual operating conditions of the train.

Keywords Maglev train · Finite element · Dynamic model · Energy consumption

1 Introduction

When the traditional wheel train is running at a high speed, the wheel wears seriously, and is not conducive to further improve the speed, but maglev train overcomes the condition, especially in the long-distance operation [1, 2]. The resistance greatly increases when maglev train is in high-speed operation, so energy consumption must be considered to save costs, and train energy loss is mainly distributed in two areas: traction system's energy consumption and auxiliary system's energy consumption [3].

The current energy consumption measurement model of power train mainly includes a data-based regression analysis model, an electrical power-based model, and a kinetic-based measurement model [4].

X. Zhang · S. Mu · J. Kang (✉)

College of Electronics and Information Engineering, Tongji University,
Shanghai, China

e-mail: kjs@tongji.edu.cn

This chapter mainly studies the energy consumption of train's traction system. The finite element model of traction motor is established, the train is analyzed from the dynamic point of view, the train's dynamic model is established, and the energy consumption model of the traction system is obtained. Then, the traction strategy is used to control the running condition of the train, and the consumption curve is obtained. In addition, taking the running time into account, the time model is established, and finally the two methods are combined to get the comprehensive strategy which has the guiding significance to the actual situation, and the solution and simulation are carried out.

2 Finite Element Model of Linear Synchronous Motor

JMAG is used to simulate the linear motor. The main processes are as below: First, the physical model of the electromagnetic field analysis object is established, and the finite element is defined, and the material attribute is defined. Second, the appropriate electromagnetic boundary condition and electromagnetic load are established for the model. Finally, the appropriate solver is used to calculate the simulation result [5] (Figs. 1 and 2).

3 Dynamic Model

The force model on the train is regarded as a single particle model, the train in the course of running is influenced by traction, resistance, and braking force. These three forces will not appear at the same time, and their resultant force is different according to different train's operating conditions [6, 7] (Fig. 3).

3.1 Traction

Traction is provided by the traction system to overcome the resistance encountered in the operation of the train. The floating electromagnet and a stator winding

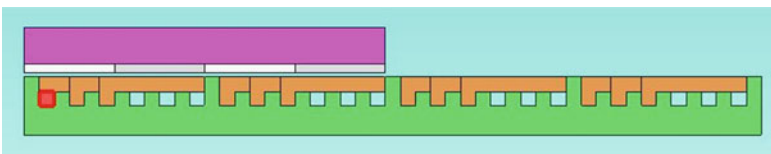


Fig. 1 Finite element model of linear synchronous motor

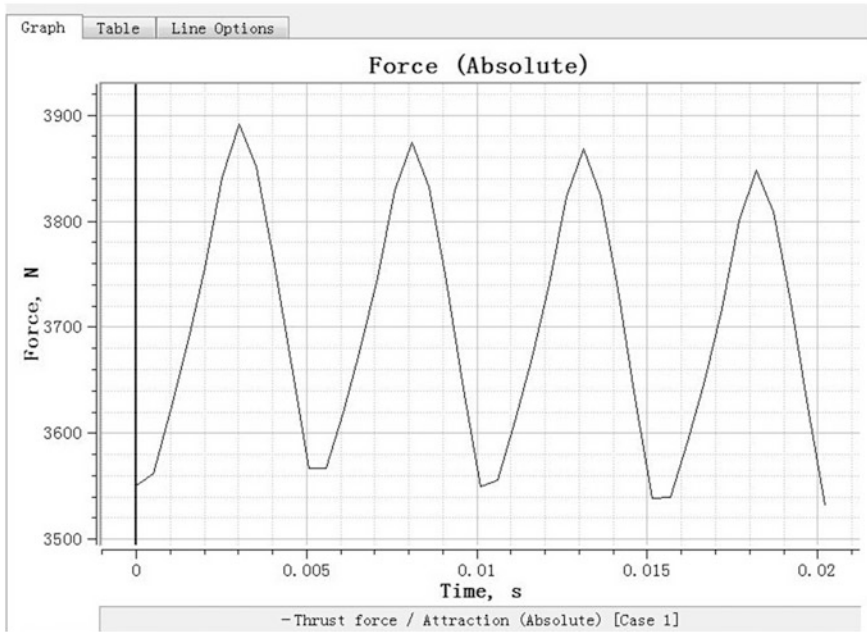
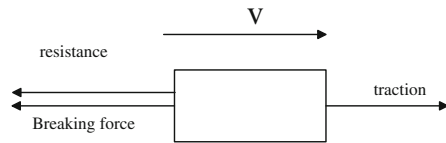


Fig. 2 The thrust on the train

Fig. 3 Train’s force diagram



generate a moving magnetic field, which causes an induced current in the suspension coil. An electromagnetic traction force to drive the train is generated by the induced current’s interacting with the moving magnetic field.

Traction can be expressed as follows:

$$F = \frac{3\pi}{2\tau} (\varphi_d i_q - \varphi_q i_d), \tag{1}$$

where F is electromagnetic thrust of the synchronous motor, τ is linear synchronous motor’s pole pitch, i_d, i_q are the current components of the stator winding in the $d - q - 0$ coordinate, φ_d, φ_q are the flux linkages of the stator windings in the $d - q - 0$ coordinates.

3.2 Resistance

The resistance of the train is divided into basic resistance and additional resistance. The basic resistance includes air resistance, linear motor's running resistance, electromagnetic eddy current resistance on both sides of the guide rail. Additional resistance is generally additional resistance to ramps.

Air resistance W_A usually uses the following empirical formula:

No lateral wind,

$$W_A = w_x \cdot 10^{-3} \cdot v^2, \quad (2)$$

where v is train's running speed, w_x is air resistance coefficient when the train is running.

With lateral wind,

$$W_A = w_x \cdot 10^{-3} \cdot v_w^2 \frac{1 + \left(\frac{0.54}{\pi}\right) \arctan\left(\frac{v_w}{v}\right)}{\sin^2\left[\arctan\left(\frac{v_w}{v}\right)\right]}, \quad (3)$$

where v_w is wind speed which is perpendicular to the running direction of the train.

The relationship among the running resistance caused by the linear motor, the train speed and the number of marshaling vehicles is

$$W_B = \begin{cases} 0 & v = 0 \\ 3.3n_s & 0 < v < 41.7 \text{ m/s} \\ \left(\frac{146}{v} - 0.2\right)n_s & v \geq 41.7 \text{ m/s}. \end{cases} \quad (4)$$

The electromagnetic eddy current resistance on the guide rails on both sides of the line is calculated as follows:

$$W_M = n_s \cdot 0.5 \left[(v/111)^{0.5} + 1.3 \cdot (v/111)^{0.7} \right]. \quad (5)$$

Additional resistance to ramps W_I is expressed as

$$W_I = M \cdot g \cdot i / 1000. \quad (6)$$

Where i is thousandth of slope.

The total resistance can be expressed as

$$W = W_A + W_B + W_M + W_I \quad (7)$$

3.3 Braking Force

In normal conditions, the brake of the maglev train that includes regenerative braking or resistive braking is realized by the long stator synchronous motor. It will start the vehicle's eddy current braking system, so that the train is safely docked to

the designated auxiliary parking area at the appropriate deceleration when the braking capacity of the traction system fails or other faults occur.

Conventional braking force B_P is as follows:

$$B_P = 3\pi/2\tau(\psi_d \times i_q - \psi_q - i_d) \quad (8)$$

Emergency braking force B_V is as follows:

$$B_V = \begin{cases} q \cdot 1.2M \cdot (1 - e^{(-v/B)}) & (v \geq 2.78 \text{ m/s}) \\ \mu \cdot Mg \cdot \sqrt{(I(s)/1000)^2} & (0 < v < 2.78 \text{ m/s}) \\ 0 & (v = 0) \end{cases} \quad (9)$$

Where q is the eddy current brake's level, B is a speed constant that is 30 m/s, and μ is the coefficient of friction.

3.4 Train's Dynamics Model

Let thousandth of slope i be a continuous function of distance, expressed as $I(s)$. Train's movement model is obtained according to Newton's law as follows:

$$\begin{cases} \frac{ds}{dt} = v \\ \frac{dv}{dt} = \frac{F - f(v)}{M} = a. \end{cases} \quad (10)$$

4 Strategy Model

Different traction strategy leads to different manipulation methods and calculation results. Extensive traction strategies include the fastest speed strategy, the most economic strategy, and hybrid optimization strategy [8].

4.1 Minimum Energy Consumption Model

The minimum energy consumption model is described as follows:

$$\min E = \min \int_0^E dE = \min \int_0^s F ds = \min \int_0^{t_n} |F|v(t) dt. \quad (11)$$

Constraints are

$$\begin{cases} v_{\min} \leq v_i \leq v_{\max} \\ a_i \\ \Delta t = t_{i+1} - t_i \\ v_{i+1} = v_i + a_{i+1} \Delta t \\ s_{i+1} = s_i + \frac{v_i + v_{i+1}}{2} \Delta t \\ v_n = v_1 = 0, \end{cases} \quad (12)$$

where v_i is the speed of the i th step, v_{i+1} is the speed of the next step, and s_i is the distance of the i th step, v_{\min} , v_{\max} are the speed limit, Δt is the step size, n is the step number, and a_i is the acceleration of the i th step.

4.2 Minimum Time Model

The minimum time model is described as follows:

$$\min = \int_0^{t_w} dt = t_w. \quad (13)$$

Constraints are the same as Eq. (12).

4.3 Optimal Model

Considering the minimum time model, the minimum energy consumption model, train's operating parameters, and line model, the most suitable operation strategy for the actual situation is obtained.

5 Simulation

For the above operation strategy, assuming that the number of marshaling vehicles is 5, the maximum traction is 500 kN, the maximum acceleration is 1 m/s², the total distance is divided into four sections, the train's operating conditions under these models are obtained by programming.

The following algorithm block diagram is used to solve the models: (Figs. 4, 5, 6, and 7)

Operation conditions for train:

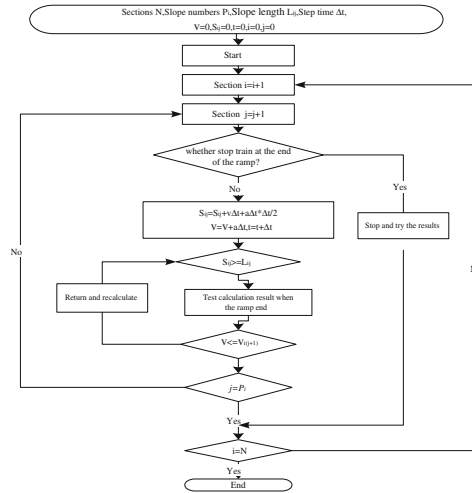


Fig. 4 Algorithm block diagram

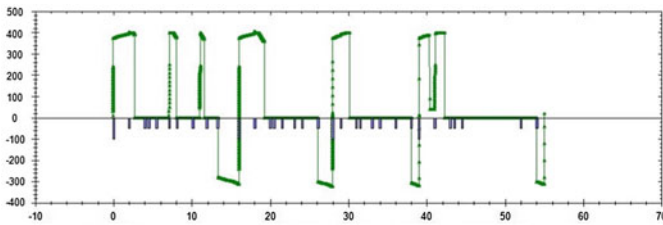


Fig. 5 F-S curve by minimum energy consumption model

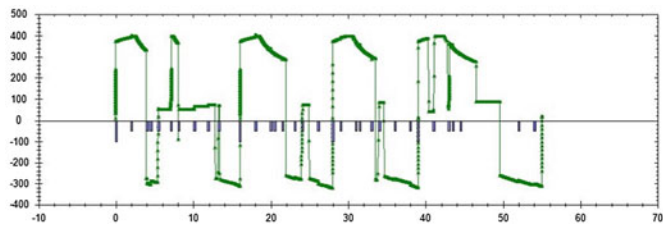


Fig. 6 F-S curve by minimum time model

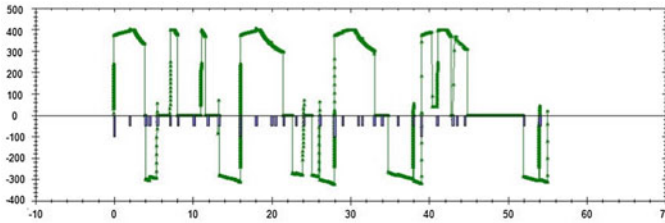


Fig. 7 F–S curve by optimal model

6 Conclusion

The purpose of the finite element analysis of the linear motor and the dynamic model of the high-speed maglev train is to obtain the accurate minimum energy consumption model of the train. However, the running time in practice must be considered at the same time. Therefore, the two models are combined to obtain the integrated optimization model. The specific operation of the train is obtained through the simulation, and the results have the advantages of both models.

Acknowledgements This work is supported by the “The thirteenth Five-Year Plan” National Key R&D Program of China (NO. 2016YFB1200602-02) and the National Key R&D Program of China (NO. 2017YFB1200900).

References

1. Motoharu O, Shunsaku K, Hisao O (2002) Japan’s superconducting maglev train. *IEEE Instrum Measur Mag* 9–15
2. Jinho L, Jeongmin J, Youngje H, Changyoung L (2013) Development of LSM control system for super speed maglev. *Control, automation and systems (ICCAS 2013)*, 2013 13th international conference on 20–23 Oct 2013 in Kimdaejung Convention Center, Gwangj, Korea, pp 466–469
3. Yuqing G, Jing Z, Youtong F, Roderick S (2016) Analysis of the energy use of various high speed trains and comparison with other modes of transport. *Ecological Vehicles and Renewable Energies (EVER)*, 2016 Eleventh International Conference on IEEE, pp 1–7
4. Xiao L, Shoudao H (2017) Magnetic field and thrust analysis of the U-channel air-core permanent magnet linear synchronous motor. *IEEE Trans Magn*, pp 1–4
5. Tong Y, Libing Z, Langru L (2008) Finite element analysis of linear induction motor for transportation systems. *IEEE Vehicle Power and Propulsion Conference (VPPC)*, 3–5 Sept 2008, Harbin, China, pp 1–4
6. Fei C, Hongze X (2006) Study on safety protection speed curve of German high speed maglev train. *Chin Sci Technol Inf* 17:248–251 (in Chinese)
7. Jing L, Wei H, Yuxiang L et al (2010) Analysis and calculation of traction characteristics of medium and low speed maglev train. *Electr Locomotive Urban Rail Veh* (6):21–22, 26 (in Chinese)
8. G Yang (2008) Energy-saving operation control strategy of high-speed maglev train. *Chinese Control and Decision Conference (CCDC 2008)*, 2008, pp 2250–2253 (in Chinese)

Risk Theory-Based Safety Evaluation of Passengers in Rail Transit Station

Yunxiao Zheng, Yong Qin, Jianyuan Guo, Limin Jia
and Jianghua Gao

Abstract Station passenger flow safety status is analyzed based on risk theory and a transit station, XI'ERQI Station, is analyzed in this paper. Station passenger flow safety status is an important part of station safety status. The results indicate that the safety of passenger flow of a rail transit station is mainly determined by both the occurrence probability and the influence degree of passenger flow. Comprehensive passenger density and passenger evacuation time are chosen as the key indices for safety evaluation of the rail transit station. Combining fuzzy theory with fuzzy comprehensive evaluation method, a fuzzy comprehensive evaluation model based on passenger flow density and evacuation time is established, which is used to identify and analyze the overall safety level of a rail transit station. The results produced from the proposed model can provide useful information for preparation of the early warning of a rail transit station, and the passenger safety assurance.

Keywords Risk theory · Comprehensive passenger density · Passenger evacuation time · Fuzzy theory · Safety-level identification

1 Introduction

Urban rail transit has been becoming the most important travel mode for urban residents, and takes on a large number of passenger transport, so it is essential to guarantee its operational safety. The transfer stations are the key nodes of the rail traffic network, and play a critical role in the connection of the rail traffic lines, of which the safety-level identification can identify the station status on a macro level in a timely manner. The demand of urban rail transit network operation for risk

Y. Zheng · Y. Qin (✉) · L. Jia · J. Gao
State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University,
100044 Beijing, China
e-mail: yqin@bjtu.edu.cn

J. Guo
School of Traffic and Transportation, Beijing Jiaotong University, 100044 Beijing, China

management has transformed from “post hoc analysis type, passive type” to “Beforehand prevention type, active type” stage, that is, from experience management to the modern system safety risk management stage [1]. The identification of safety status is an important part of safety risk management.

Many researches on the safety status of passenger flow in the hub are based on the classification of the service level of pedestrian facilities. Fruin [2] has established a method for calculating the service level of pedestrian walking facilities. Sarker [3] proposed that the service level of pedestrian transportation facilities should be divided by considering safety, comfort, convenience walking continuity, and so on. Combined with safety consideration and taking the speed and density of pedestrians as indices, Huang [4] put forward the evaluation method of the safety status of railway comprehensive passenger transport hub. Jie [5] proposes queue length and waiting time as the evaluation indices of passenger flow congestion level, and determined the grade division standard.

The analysis of risk has been going on for a long time, which is an important prerequisite for risk control. Many Western scholars have done a lot of research on risk theory. Douglass [6] pointed out that in economics, statistics, and insurance, generally the “cost-income” logic is implemented, using a unified formula “ $\text{risk}(R) = \text{injury degree}(H) \times \text{the possibility of occurrence}(P)$ ” to calculate the risk. This logic is also feasible in safety study and management. Zhang [7] combines risk with safety from the point of view of risk theory. Sui [8] pointed out that the loss of an accident depends on the frequency and severity of the accident, and discussed the safety principle systematically.

Based on the theory of risk, this chapter analyzes the safety of passenger flow from two aspects: the probability of passenger flow risk and the influence degree of passenger flow risk, and proposes the corresponding safety evaluation indices of station passenger flow. Combined with fuzzy theory and fuzzy comprehensive evaluation method, a model for identifying overall safety level of rail transit station is built, and based on the simulation of passenger flow using AnyLogic, an example is given to verify that the model could evaluate and identify the station safety level.

2 Risk Theory in Safety Rating

The risk is intended to be a possible danger, of which the theory is used to study both the possibility and the degree leading to harm or damage to humans and things. The concept and theory of risk have been widely introduced and analyzed in economics, insurance, statistics, and Actuarial science. As a potential state, risk includes two meanings, that is, the possibility of accident or disaster and its harmful consequences. The magnitude of risk depends on the severity of consequences and its probability of occurrence [7].

With reference to the risk classification method of the risk theory, the quantitative calculation of the level of risk is defined as follows [9]:

$$R = P \times I, \quad (1)$$

where R is the risk value, P is the risk occurrence probability, and I is the risk influence degree. Equation (1) is a semi-quantitative calculation, but not a full quantitative calculation. Safety is a relative concept, and danger is the membership degree of safety, when the level of risk is less than some extent, the people think that is safe. Safety (S) and danger (D) are the complementary to each other, which is [8]:

$$S = 1 - D. \quad (2)$$

Risk theory provides a basis for effective risk management, reducing risk and harm, which is important to ensure the safety of station. Therefore, this chapter applies the risk theory to the research on the safety-level identification of the station.

Risk (R) equals to danger (D) in this study. The quantitative calculation of passenger traffic safety value of rail transit station is defined as follows:

$$S = 1 - P \times I, \quad (3)$$

where R is the safety value, P is the passenger risk occurrence probability, and I is the passenger risk influence degree.

According to the above theory, the passenger flow safety of rail transit station is mainly determined by the passenger risk occurrence probability and passenger risk influence degree. Therefore, the safety guarantee of passenger flow in rail transit station is mainly considered in these two dimensions: decreasing the passenger risk occurrence probability which means the probability of occurrence of passenger flow risk events and reducing the passenger risk influence degree which means consequences of a dangerous event when an emergency occurs.

In this chapter, the identification and evaluation of passenger flow safety level for rail transit station are mainly carried out from above dimensions, and the corresponding evaluation indices are selected respectively.

3 Safety Evaluation Index

According to the risk theory as stated in the above section, we need to select the index that can reflect the passenger risk occurrence probability and the passenger risk influence degree. The most important indicators in the two aspects are selected to assess the safety of station passenger flow.

3.1 Comprehensive Passenger Density

Study on GROUP DYNAMICS demonstrated that the greater the population density is, the lower the speed of the group would be. Crowd gathering would result in risk events such as stampede [10]. The greater the density of passenger flow is, the higher the possibility of safety accidents such as stampede would be. Therefore, the density of passenger flow can directly reflect passenger risk occurrence probability well. The comprehensive passenger density is defined as an index to reflect the density of passenger flow in a station, which is calculated as follows:

$$P = \sum_{i=1}^n \omega_i \rho_i, \quad (4)$$

where P is the comprehensive passenger density, ρ_i is the density of various facilities, ω_i is the weight of various facilities, and n is the facility number.

3.2 Passenger Evacuation Time

Passenger evacuation time refers to the period from the occurrence time of dangerous events to the time when all passengers are evacuated from the station. When the passenger flow risk event occurs, the shorter the evacuation time is, the less the life and property loss would be. Therefore, passenger evacuation time can directly reflect the passenger risk influence degree.

Referring to GB505517-2003 code for subway design, the passenger evacuation time is calculated as follows:

$$T = \frac{Q_1 + Q_2}{0.9[A_1 b(N - 1) + A_1 B]} + 1, \quad (5)$$

where T is the safe evacuation time, Q_1 is the total number of passengers in the arriving train, Q_2 is the total number of passengers and staff in the station, A_1 is the evacuation capacity of escalators, A_2 is the evacuation capacity of pedestrian stairs, N is the number of escalators connected to the platform, b is the width of the escalator connected to the platform, B is the width of the stairs linked to the platform, and 0.9 is the correction rate affecting the efficiency of equipment use.

4 Safety Evaluation Model

Fuzzy comprehensive evaluation method is a comprehensive decision method in fuzzy scope, which can quantify some factors with unclear boundaries and convert the qualitative evaluation into quantitative evaluation [11]. Due to its characteristics of the clear result and strong systematicness, it is very suitable to solve problems with fuzziness and difficult to quantify. In this chapter, the fuzzy evaluation method is used to identify the passenger flow safety level of the station.

The establishment steps of fuzzy comprehensive evaluation model are as follows:

- (1) Determination of the evaluation factor set U .
According to the above discussion, the comprehensive passenger density and passenger evacuation time are taken as evaluation indices. They are used as evaluation factors in evaluation factor set, the evaluation factor is set as $U = \{P, T\}$.
- (2) Determination of the evaluation set V .
In order to describe the safety of passenger flow more clearly, according to the safety degree, the passenger flow safety level of rail transit station is divided into five levels—dangerous, less dangerous, general, less safe, and safe. The evaluation set is $V = \{V_1, V_2, V_3, V_4, V_5\}$, where V_1 means “dangerous”, V_2 means “less dangerous”, V_3 means “general”, V_4 means “less safe”, and V_5 means “safe”. The passenger flow safety levels are described in Table 1:
- (3) Determination of the membership function and building the judgment matrix

The membership function formula of the first-level index is

$$f_{v_1}(x) = \begin{cases} 1 & x < y_1 \\ \frac{y_2 - x}{y_2 - y_1} & y_1 \leq x \leq y_2 \\ 0 & x > y_2. \end{cases} \tag{6}$$

Table 1 Passenger flow safety levels and corresponding descriptions

Safety level	Safety-level description
Dangerous	Extreme poor and unacceptable safety situation. Passenger flow is seriously congested, and a great number of conflicts occur between pedestrians
Less dangerous	Poor safety situation. Passenger flow appears sluggish, congestion occurs from time to time, many conflicts occur between pedestrians
General	General safety situation. The flow of passengers is slow and some conflicts occur between pedestrians
Less safe	Good safety situation. The flow of passengers is basically smooth and a few conflicts occur between pedestrians
Safe	Very good safety situation. The flow of passengers flowing smoothly and orderly; almost no conflict occurs between pedestrians

The membership function formula of second to fourth index is

$$f_{v_i}(x) = \begin{cases} 0 & x < y_{i-1}, x > y_{i+1} \\ \frac{x-y_{i-1}}{y_i-y_{i-1}} & y_{i-1} \leq x \leq y_i \\ \frac{y_{i+1}-x}{y_{i+1}-y_i} & y_i < x \leq y_{i+1} \end{cases} \quad i = 2, 3, 4 \tag{7}$$

The membership function formula of the fifth index is

$$f_{v_5}(x) = \begin{cases} 0 & x < y_4 \\ \frac{x-y_4}{y_5-y_4} & y_4 \leq x \leq y_5 \\ 1 & x > y_5 \end{cases} \tag{8}$$

The thresholds y_j ($j = 1, 2, 3, 4, 5$) of comprehensive passenger density and the passenger evacuation time need to be determined respectively. The values of which are mainly determined by experience, relevant norms, and previous study in this chapter.

The comprehensive passenger density and the passenger evacuation time are substituted into the above formula. The judgment matrix is built as follows:

$$R = \begin{bmatrix} (R|P) \\ (R|T) \end{bmatrix} = \begin{bmatrix} \rho_{v_1}(x) & \rho_{v_2}(x) & \rho_{v_3}(x) & \rho_{v_4}(x) & \rho_{v_5}(x) \\ t_{v_1}(x) & t_{v_2}(x) & t_{v_3}(x) & t_{v_4}(x) & t_{v_5}(x) \end{bmatrix} \tag{9}$$

(4) Determination of the weights

Because various factors in evaluation factor set occupy different proportions in the comprehensive evaluation, it is necessary to analyze the influence of the comprehensive passenger density and passenger evacuation time on the safety status of rail transit station. Then, the fuzzy weight vector in the evaluation factor set is determined as $u = [u_\rho, u_t]$, where u_ρ and u_t are weight values of the comprehensive passenger density and the passenger evacuation time respectively.

(5) Identification of the safety level

The fuzzy judgment vector is obtained as

$$v = uR = [u_\rho, u_t] \begin{bmatrix} \rho_{v_1}(x) & \rho_{v_2}(x) & \rho_{v_3}(x) & \rho_{v_4}(x) & \rho_{v_5}(x) \\ t_{v_1}(x) & t_{v_2}(x) & t_{v_3}(x) & t_{v_4}(x) & t_{v_5}(x) \end{bmatrix} = [v_1, v_2, v_3, v_4, v_5] \tag{10}$$

After the fuzzy evaluation vector is obtained, according to the principle of maximum membership degree, the maximum value is observed, the safety level of the passenger flow is judged, and the safety-level recognition result is obtained.

5 Example Verification

5.1 Calculation of the Comprehensive Passenger Density

XI'ERQI Station is the southern terminal of Subway Changping Line in Beijing and its transfer station to Subway Line 13. The basic data, including the basic physical data of the station facilities, passenger flow data, service time of equipment, running ratio, and so on, are obtained through field investigation, to support the establishment of simulation model using AnyLogic of XI'ERQI Station in Beijing.

The acquisition scene of passenger flow data is a day's morning peak hour in XI'ERQI Station. 15 min passenger flow simulation are obtained using AnyLogic, and on this basis, the density values of each facility from different simulation times can be gained by data acquisition using AnyLogic Fig. 1.

Through the form of questionnaires, the safety level of each facility is scored by people who would select urban rail transit as the main travel mode for commuting. Weights can be obtained by using the method of Sect. 4 in this chapter. Then the comprehensive passenger density can be obtained by weighted average method as shown in Tables 2 and 3.

According to the relevant safety and service-level standards for the division of Danaher [12], which is designed to evaluate the level of service of walking facilities in transportation terminals and depicted in Fig. 2, the thresholds of comprehensive passenger density are set to 0.31, 0.43, 0.72, 1.08, and 2.15.

5.2 Calculation of the Passenger Evacuation Time

The starting of simulation time is set at the 0 moment, the train departing intervals for train directions going to Dongzhimen and Xizhimen of 13 lines is set as 3 min, while the train departing intervals of Changping line is set as 6 min.

The number of passengers loaded on the train was set to 1700, then the total number of stations is obtained by simulation data output. Physical device data were investigated. There are 4 stairs on the platform floor, of which the width is 2 m,

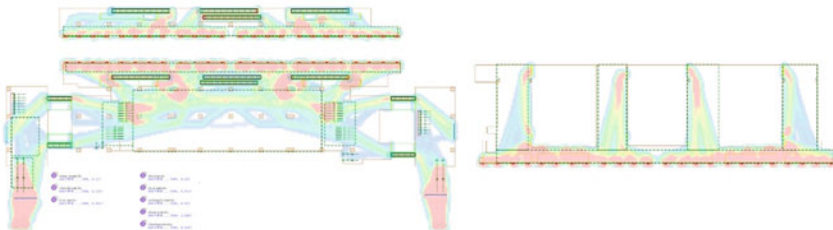


Fig. 1 Simulation diagram of XI'ERQI Station using AnyLogic software

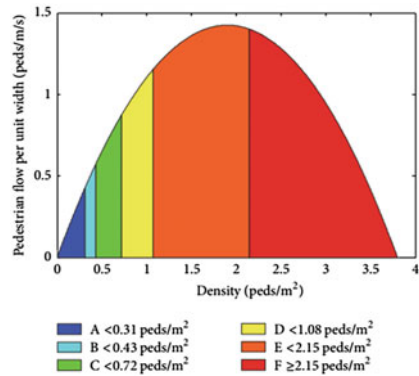
Table 2 Expert scoring weighting table

Facilities	Subway hall	Passage	Platform	Stairs/escalator	Gate
Weight	0.13	0.25	0.27	0.19	0.16

Table 3 Density of facilities at different simulation time

Simulation time (min)	Facilities densities (peds/m ²)					Comprehensive passenger density
	Subway hall	Passage	Platform	Stairs/escalators	Gate	
3	0.241	0.222	0.618	0.647	0.204	0.40926
6	0.120	0.151	1.061	0.860	0.596	0.59858
9	0.263	0.302	1.752	0.993	0.911	0.91716
12	0.202	0.240	1.857	1.003	0.800	0.90622
15	0.286	0.267	2.665	1.003	0.639	1.11629

Fig. 2 Pedestrian flow fundamental diagram and the level of service



and 4 stairs in the hall floor, of which the width is 2.8 m, thus $B = 19.2$ m. The width of escalators in the station is $b = 1.1$ m, and the number of down escalators is $N = 19.2$ m. Evacuation capacity is set as $A_1 = 135$ $A_2 = 62$.

In the code for subway design, it is stipulated that the station facilities shall guarantee when the safety accidents occur during the peak hours, both the passenger flow and the staff on the train and platform should be evacuated within 6 min. When more than 6 min, the passenger flow is in a “dangerous” state. Take 6 min as the value when the passenger flow is in a “less dangerous” safety condition, then the thresholds of the passenger evacuation time are set to 0, 2, 4, 6, and 8.

According to the Eq. (5), the passenger evacuation time is calculated as shown in Table 4.

The weights of comprehensive passenger density and passenger evacuation time are scored by experts, which indicates comprehensive passenger density is more

Table 4 Passenger evacuation time at different simulation times

Simulation time (min)	3	6	9	12	15
Number of people in the station	2242	2557	4038	4327	5496
Number of people in the train	3400	5100	3400	5100	3400
Evacuation number	5642	7657	7438	9427	8896
Passenger evacuation time	4.243256	5.401561	5.275671	6.419031	6.11379

Table 5 Safety-level identification

Simulation time (min)	The fuzzy judgment vector	Safety-level identification result
3	[0.13, 0.62, 0.22, 0.03, 0]	Less safe
6	[0, 0.31, 0.51, 0.18, 0]	General safe
9	[0, 0, 0.43, 0.57, 0]	Less dangerous
12	[0, 0, 0.36, 0.59, 0.05]	Less dangerous
15	[0, 0, 0, 0.96, 0.04]	Less dangerous

important than passenger evacuation time. According to fuzzy analytic hierarchy process, precedence relation matrix is built as $F = \frac{P}{T} \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix}$, which is converted to fuzzy consistent judgment matrix later as $M = \begin{bmatrix} 0.5 & 0.75 \\ 0.25 & 0.5 \end{bmatrix}$. After iterative computation by row normalization method, the sort vector is $W = (0.75, 0.25)$, which is used as a weight vector u .

Then the fuzzy judgment vector is calculated, and safety-level identification results can be obtained complying with the principle of maximum membership as shown in Table 5.

When the simulation time is 3, 6, 9, 12, and 15 min, the station passenger flow safety level is less safe, general safe, less dangerous, less dangerous, and less dangerous.

It can be seen that the safety levels continue to fall, and the danger degree is rising, which coincides that the simulation scene is morning peak hour, where with the increase of passenger flow, the passenger density and passenger evacuation time are increasing, and the station becomes more and more dangerous.

It shows that it is feasible to apply risk theory to evaluate the passenger flow safety of station.

6 Conclusion

On the basis of risk theory, this chapter applies the risk assessment method to the identification of rail transit station safety level, and provide the basis for safety management of rail transit station. From the two aspects of passenger risk occurrence probability and passenger risk influence degree, the influencing factors of passenger flow safety are analyzed, and the evaluation indices of passenger flow safety are selected as comprehensive passenger density and passenger evacuation time. The identification method of the overall passenger flow safety level is given by fuzzy comprehensive evaluation method. The simulation is carried out and the example is verified that the model is proved to be available. The purpose of this chapter is to provide some ideas and reference for the overall passenger flow safety-level identification of rail transit station, and further research will be done on the selection of passenger flow safety evaluation indices and the determination of the weights.

Acknowledgments The authors would like to express their thanks to the editor and anonymous reviewers for their help in revising the manuscript. This study is sponsored by the National key R & D project (2016YFB1200402). The support of State Key Lab of Rail Traffic Control & Safety of Beijing Jiaotong University is also gratefully acknowledged.

References

1. Xu T (2012) Study on risk assessment theory and methods on urban rail transit network operation. Beijing Jiaotong University, China (in Chinese)
2. Fruin JJ (1971) Pedestrian planing and design. Metropolitan Association of Urban Designers and Environmental Planners Inc, NY
3. Sarker S (1993) Determination of service levels for pedestrians with European examples. *Transp Res Rec* 1405:35 (Transportation Research Board)
4. Huang H (2011) Research of passenger safety status evaluation of comprehensive railway passenger transport hub. Beijing Jiaotong University, China (in Chinese)
5. Jie H (2014) Study on evaluation model of passenger flow congestion level in urban rail stations. Beijing Jiaotong University, China (in Chinese)
6. Douglass M (2002) Risk acceptability according to the social science. Routledge & Kegan Paul, UK
7. Zhang Q (2004) Risk analysis and management in the Beijing Olympics Park. Nankai University, China (in Chinese)
8. Sui P, Chen B, Sui X (2005) Safety principle. Chemical Industry Press, China, pp 11–13 (in Chinese)
9. Tang J (2015) Risk assessment and management of social stability. Peking University Press, China, pp 166–172 (in Chinese)
10. Ou K (2016) Causes of large-scale mass activities stampede and control strategies. *J Hunan Police Academy* 28(2):72–77 (in Chinese)
11. Du D, Pang Q, Wu Y (2008) Modern comprehensive evaluation method and case selection. Tsinghua University Press, China, pp 34–61 (in Chinese)
12. Danaher AR (2003) Transit capacity and quality of service manual. Transportation Research Circular, USA

Level of Service for Passage in Urban Rail Transit Station

Sai Chen, Jie Xu, Limin Jia, Yong Qin, Kunsheng Zhan
and Jian Zhang

Abstract The level of service determines the operational efficiency of the station. In order to accurately judge and predict the level of service of the passage, first of all, the evaluation index is set up. The satisfaction grading standard is established, The cloud model is used to describe the level of service based on fuzzy theory. The level of service of passage at a certain time is predicted based on Markov theory. Taking the facilities of Xizhimen station in Beijing Subway as an example, the level of passage service is evaluated and predicted, and the feasibility of the theory is verified.

Keywords Urban rail transit station · Level of service · Passage facilities
Cloud model · Markov theory

1 Introduction

As the main facility of passenger transfer and distributing in urban rail transit station. The purpose of this chapter is to grasp the dynamic level of service of the passage facilities, and to ensure the safety of the passage facilities. And also to provide the level of service of the regional facility node to site staff and improve the efficiency of the operation organization.

In the relevant theoretical research, Zhou and Zhang proposed a cloud model-based passenger flow-state recognition method [1, 2]. Yin and Li proposed a combination of fuzzy evaluation and level of service evaluation index calculation method [3]. Sheila put forward the qualitative analysis methods to assess the level of pedestrian service [4]. These methods focus on the indicators of fuzzy. And there

S. Chen (✉) · J. Xu · L. Jia · Y. Qin
State Key Laboratory of Rail Traffic Control and Safety, Beijing
Jiaotong University, Beijing 100044, China
e-mail: 16125723@bjtu.edu.cn

K. Zhan · J. Zhang
Operations Division, Guangzhou Metro Group Co., Ltd, Guangzhou 510330, China

are no strict grading standards. Pelechano and Malkawi proposed more specific indicators based on pedestrian comfort and convenience to assess facility level of services [5]. Khisty and Dixon proposed a five-point metric to classify the pedestrian level of services [6]. Combined with passenger perception, Zhang evaluated the level of service of the station's internal security check, ticket purchase, passage, and other facilities [7].

Taking the large passenger flow of urban rail transit station as the background, this chapter uses the research data to select the service index. The subway station level of service judgment method from the quantitative point of view is put forward by using the cloud model theory. And then the state level is forecasted by using the Markov theory.

2 Analysis of Service Index and Establishment of Satisfaction

2.1 Index Selection

The pedestrian speed, density, and flow volume as the index reflecting the crowdedness of pedestrian walking in the passage is selected. And the crowdedness reflects the level of service. With the increase in flow volume, the density increases and the speed decreases [8].

2.2 Satisfaction Establishment

In order to establish the relationship between passenger satisfaction and evaluation index, the concept of membership degree is introduced. The membership function $A(x)$ can be set to the triangular distribution. The questionnaire survey is used to obtain grading standard of satisfaction. The questionnaire was printed with the photos of different densities of the passage, then passengers select satisfaction reviews for different photos. Satisfaction reviews j corresponding to passenger density D_j is calculated as follows:

$$D_j = \sum \left(\frac{n_{ij}}{N} \times d_i \right) / \sum \frac{n_{ij}}{N}, \quad (1)$$

where d_i is the density of the i th photo, n_{ij} is the number of passengers who selected the photo i that selected satisfaction reviews j , and N is the total number of passengers surveyed.

Finally, we can get the satisfaction grade standard of the passage as shown in Table 1.

Table 1 Grading standard of satisfaction for passageway

Comment	Density (p/m ²)							Graded
	0.21	0.36	0.64	0.71	0.79	1	1.29	
Very satisfied	0.38	0.2	0.08	0.03	0	0	0	0.32
Satisfactory	0.58	0.63	0.5	0.38	0.2	0	0	0.46
General	0.05	0.18	0.3	0.4	1.33	0.15	0.08	0.72
Less satisfied	0	0	0.13	0.23	1.25	0.5	0.3	0.93
Very dissatisfied	0	0	0	0	0.2	0.35	0.65	1.16

3 Evaluation Cloud Model of Passage Service Level

The cloud model is used to deal with the transformation between qualitative and quantitative concepts. Through the cloud model, we can divide different levels of service into different intervals, so as to quantify the level of service through the evaluation index. The digital characteristics of the cloud include the expectation Ex that represents the center of the cloud distribution corresponding to the level of service, the entropy En that represents the range of values for the level of service, and the super entropy He represents the membership of a certain level of service random size [9, 10]. According to the cloud’s digital characteristics, the cloud drop process is as follows:

- Step1: Input the digital characteristics, which represents G . Generate the random numbers En'_i with expectation of En and variance of He^2 and the random numbers x_i with expectation of Ex and variance of En_i^2 .
- Step2: Complete the first cloud model operation of the concept G , which corresponds to the horizontal axis of the newly cloud drops.
- Step3: According to the formula, calculate the membership degree of the x drop for the concept G , which represents the ordinate of the newly cloud drops.

$$u_i(x) = Exp \frac{-(x_i - Ex)^2}{2(En'_i)^2}. \tag{2}$$

- Step4: Repeat step1 to step5, until the completion of the Nth positive cloud is completed.

3.1 Construction of Standard Cloud Model for Level of Service

According to the classification standard of service index calculated by satisfaction classification, the level of A, B, C, D, E, and F is established, and the level of A to F decreases from high to low. Standardize the threshold of the index, evaluation space will be divided into six fuzzy space through the five indicators of the threshold. The levels of service A and F select half liters of normal cloud and half down normal cloud, corresponding expectations Ex_{x1} , Ex_{x6} , entropy En_{x1} , and En_{x6} are as follows:

$$\begin{cases} Ex_{x1} = x_1^* \\ En_{x1} = En_{x2} \\ Ex_{x6} = x_5^* \\ En_{x6} = En_{x5} \\ He = 0.010, \end{cases} \tag{2}$$

where x_1^* and x_5^* are the normalized values of the first and fifth thresholds, respectively B, C, D, and E are represented by full normal clouds, and the digital features are calculated as follows:

$$\begin{cases} Ex_{xj} = x_{j-1}^* + x_j^*/2, \\ En_{xj} = x_{j-1}^* + x_j^*/6, \\ He = 0.010, \end{cases} \tag{3}$$

where x_j^* is the normalized value corresponding to the threshold x_j . Synthesis of six levels of service under the sub-cloud of indicators, that is, the standard cloud.

3.2 Judgment of Identification Cloud Model for Level of Service

The identification cloud model is the evaluation level of the facility at a given time, which is compared with the standard cloud according to the passenger flow index of a specific time to evaluate the facility. The cloud model to be identified mainly composes of five steps:

- Step1: Input the corresponding digital characteristics of the evaluation index at each level. Establish the positive normal cloud computation CG_{r1} of the metric in MATLAB according to the positive normal cloud generator algorithm.
- Step2: Standardize the real value r such that the pedestrian occupies the spatial indicators actually collected, which is recorded as r^* . For the target value belonging to the range $r \leq \min\{r_j\}$, the standardized result is 0, for the

target value belonging to the range $r \geq \max\{r_j\}$, the standardized result is 1.

- Step3: Input r^* to the positive normal cloud operation $CG_{r1}, CG_{r2}, CG_{r3}, CG_{r4}, CG_{r5},$ and CG_{r6} .
- Step4: Calculate the average speed and the flow value per unit width according to the above steps, and the corresponding sub-cloud eigenvalues are obtained.
- Step5: Set the sub-cloud calculated indicators as $R_r, S_s, T_t,$ respectively. Synthesize to be identified cloud.

$$U = R_r * S_s * T_t = (Ex, En, He) \tag{4}$$

4 Level of Service Prediction Based on Markov

4.1 Theoretical Model

Define the current facility level of service under A, B, C, D, E ,and F six levels. Obtain the probability P_{ij} in which the previous period of level of service is at i level, the other level of service is at j level. Construct the first step transition probability matrix, and the probability is

$$P_1 = (P_{ij}), \tag{5}$$

$$P_k = P_1^k. \tag{6}$$

The probability vector $P(t) = (P_1(t), P_2(t), \dots, P_i(t))^T$ is the absolute probability vector of the t -time period. Where $P_i(t)$ represents the absolute probability of the level of service at the i th order in the t -time period. According to the above equation, the absolute probability vector of the $t + k$ period is

$$P(t) = P(t)P_k = P(t)P_1^k \tag{7}$$

4.2 State Prediction

The flow forecasting model based on Markov state transfer is the following steps:

- Step1: Build the cloud model on the level of service at all times combined with the passage level of service indicators, and obtain the level of service of each time.

- Step2: Count the level of service of each known time. Obtain the transfer frequency matrix and the transfer probability matrix of the Markov chain with different steps.
- Step3: Take the current value of the passenger flow status indicator for several periods as the initial value. Using the transition probability matrix of each order to predict the probability P_i^k that next unknown time passenger flow status index value of the i level.
- Step4: The prediction probability of the passenger flow state in the next unknown period is the sum of the predicted probabilities of the same state. The predicted passenger flow condition of the period is $\max\{P_i, i = 1, 2, \dots, m\}$. Add the passenger flow status indicator of the time period to the known traffic flow sequence and return to Step1 to continue to predict the passenger flow status for the next unknown time.

5 Case Study

Take Xizhimen station of Beijing Subway as an example. The daily passenger flow of Xizhimen station changes significantly. And the passenger flow is very large during the morning and evening rush hours.

5.1 Construction of Standard Cloud Model

In the standard cloud model construction, the six-level evaluation criterion is established for the level of service status evaluation index of the selected passage by the satisfaction grading standard (Fig. 1).

Then calculate and get the digital characteristics of the passage cloud as shown in Table 2.

Finally, the standard cloud model is synthesized according to the digital characteristics of the passage cloud as seen in Fig. 2.

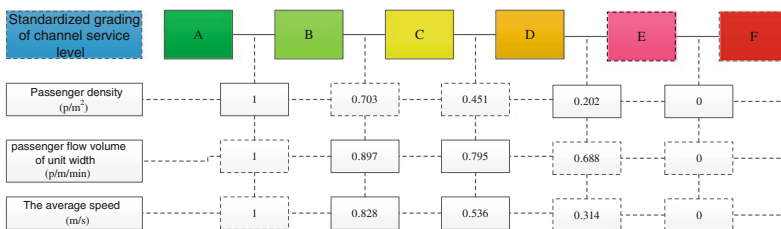


Fig. 1 Standardized passageway service index-level classification

Table 2 The digital characteristic of passageway cloud

Infrastructure	LOS	Ex	En	He
Passage	A	1.000	0.041	0.010
	B	0.886	0.041	0.010
	C	0.623	0.045	0.010
	D	0.475	0.041	0.010
	E	0.261	0.059	0.010
	F	0.000	0.059	0.010

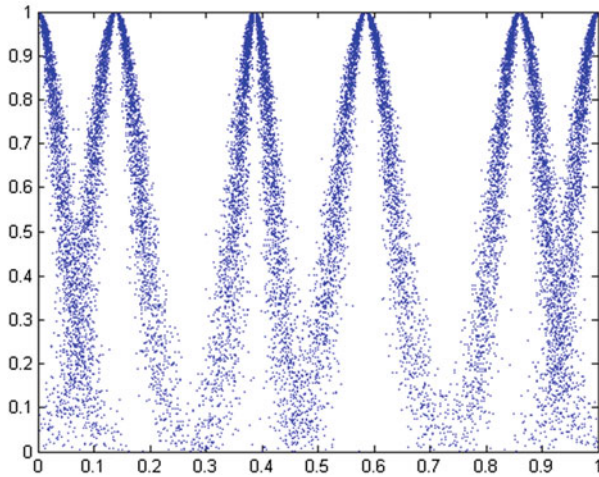


Fig. 2 Passageway standard cloud

5.2 Judgment of Identification Cloud Model

Select the data of passage facilities at the highest peak time 8:35 as an example of identification. When synthesizing the passage identification cloud graphics, standardize the source data obtained at first. And certainly, the degree membership standardized for each index is as shown in Table 3.

Calculating the numerical characteristics of the facility index identification sub-cloud are as shown in Table 4.

Table 3 The standardized weights

Infrastructure	Index	A	B	C	D	E	F
Passage	Passenger density (p/m ²)	0.000	0.000	0.000	0.000	0.656	0.344
	Average speed (m/s)	0.000	0.000	0.000	0.017	0.646	0.336
	Flow volume of unit width (p/m/s)	0.000	0.000	0.000	0.023	0.750	0.227

Table 4 The digital characteristic of identification cloud

Infrastructure	Index	Ex	En	He
Passage	Passage density (p/m^2)	0.152	0.063	0.010
	Average speed (m/s)	0.752	0.027	0.010
	Flow volume of unit width ($p/m/s$)	0.162	0.052	0.010

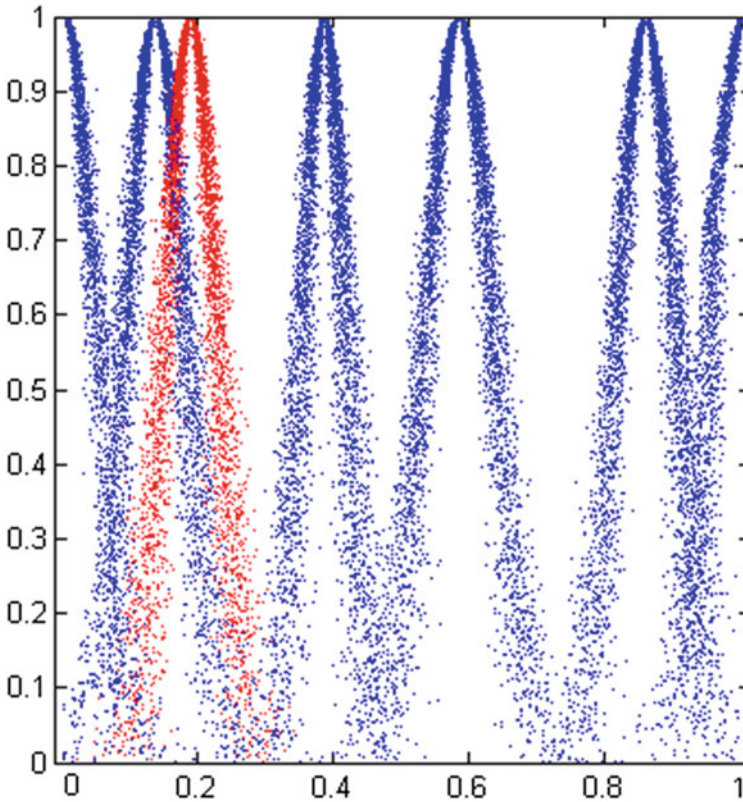


Fig. 3 Passageway identification cloud graphics

The synthetic identification cloud pattern of the passage is as seen in Fig. 3.

5.3 Level of Service Prediction Based on Markov Dynamic Transfer

Take the 25 data from 7:00–9:00 at the peak of the actual survey passage as an example. A standardized state transition matrix is established by statistical data of consecutive times. The statistics are shown in Table 5.

Table 5 Cloud model result of passage level of service during peak hours

Traffic flow parameters	Peak time data									
	7:00 a.m.	7:05 a.m.	7:10 a.m.	7:15 a.m.	7:20 a.m.	7:25 a.m.	7:30 a.m.	7:35 a.m.	7:40 a.m.	7:45 a.m.
Speed (m/s)	1.24	1.36	0.96	1.17	0.83	0.51	0.33			
Flow (p)	2073	1982	2843	2035	2960	3006	3152			
Density (p/m ²)	1.18	0.81	1.74	1.35	1.82	2.22	2.59			
Level of service	C	B	D	C	D	E	F			
~	7:35 a.m.	7:40 a.m.	7:45 a.m.	7:50 a.m.	7:55 a.m.	8:00 a.m.	8:05 a.m.			
Speed (m/s)	0.83	0.62	0.37	0.55	0.33	0.42	0.73			
Flow (p)	2834	2898	3016	2912	2927	2965	2867			
Density (p/m ²)	1.94	2.22	2.52	2.37	2.58	2.47	2.25			
Level of service	D	E	F	E	F	F	E			
~	8:10 a.m.	8:15 a.m.	8:20 a.m.	8:25 a.m.	8:30 a.m.	8:35 a.m.	8:40 a.m.			
Speed (m/s)	0.54	0.25	0.58	0.18	0.24	0.75	1.07			
Flow (p)	2795	2866	2743	2865	2904	2649	2387			
Density (p/m ²)	2.31	2.55	2.34	2.72	2.61	1.97	1.49			
Level of service	E	F	E	F	F	E	D			
~	8:45 a.m.	8:50 a.m.	8:55 a.m.	9:00 a.m.	9:05 a.m.					
Speed (m/s)	1.31	1.36	1.54	1.44	1.56					
Flow (p)	2239	1867	1706	1765	943					
Density (p/m ²)	1.04	0.62	0.46	0.51	0.43					
Level of service	C	B	A	B	A					

Table 6 Prediction calculation of five steps transfer

Initial period	Flow status	Step size to 9:05	A	B	C	D	E	F	Probability source
9:00 a.m.	B	1	0.75	0.18	0.07	0	0	0	p_{ij}^1
8:55 a.m.	A	2	0.886	0.114	0	0	0	0	p_{ij}^2
8:50 a.m.	B	3	0.253	0.583	0.164	0	0	0	p_{ij}^3
8:45 a.m.	C	4	0.124	0.243	0.451	0.182	0	0	p_{ij}^4
8:40 a.m.	D	5	0.281	0.126	0.255	0.263	0.075	0	p_{ij}^5
Probability sum p_i			2.294	1.246	0.94	0.445	0.075	0	5
The proportion of probability η_i			0.4588	0.2492	0.188	0.089	0.015	0	–

It can predict the passenger flow status at 9:00 a.m. based on the passenger status data table and the corresponding passenger flow state transition probability matrix. The results are shown in Table 6.

From the last line in the table, we can see that the maximum probability is $\max\{p_i, i = 1, 2, 3, 4\} = p_1 = 2.294$ after the probability of the same state has been superimposed. The probability ratio η_i reaches 45.88%, and the most likely service state is the A level.

6 Conclusion

Based on the survey data, this chapter determines the level of the basic index of facility service combining with satisfaction surveys. This chapter puts forward the theory of facility level of service evaluation based on Markov theory dynamic cloud model. The method can also be applied to the facilities used in the whole station. However, there are some shortcomings in the method, such as the process of dynamic cloud model processing still has a large amount of calculation and calls processing cumbersome problems to be further optimized.

Acknowledgments The authors gratefully acknowledge the support provided by China’s National “13th Five-Year” key research project “Safety assurance technology of urban rail system” (Grant No. 2016YFB1200402) and “National Engineering Laboratory for System Safety and Operation Assurance of Urban Rail Transit”.

References

1. Zhou J, Chen H, Yan B (2014) An empirical study on pedestrian traffic characteristics of subway transfer hub. *J Wuhan Univ Technol* 36(4):92–99 (in Chinese)
2. Zhang W (2014) Research on congestion recognition and early-warning control of subway transfer hub. Master thesis, Chang’an University, Shanxi, China (in Chinese)

3. Yin F, Li F (2000) Public transportation level of service fuzzy evaluation. *J Shanghai Jiaotong Univ* S1:100–104 (in Chinese)
4. Sheila S (1993) Determination of level of services for pedestrians with European examples
5. Transportation Research Record, TRB, National Research Council, Washington, D.C
6. Pelechano N, Malkawi A (2008) Evacuation simulation models: challenges in modeling high rise building evacuation with cellular automata approaches. *Autom Constr* 17(4):377–385
7. Dixon LB (1996) Bicycle and pedestrian level-of-service performance measures and standards for congestion management systems. *Transportation Research Record*, 1–9
8. Liu G (2007) Research of association rules based on cloud model. Master thesis, Shandong Normal University, Shandong, China (in Chinese)
9. Zhang H (2011) The level of service of urban rail transit station passenger traffic simulation evaluation research. Master thesis, Beijing Jiaotong University, Beijing, China (in Chinese)
10. Zhou J, Chen H, Yan B (2016) Identification method of congestion degree of subway transfer hub based on cloud model. *J Jilin Univ* 46(1):100–107 (in Chinese)

Acceptance Check Scheme of Integration Test and Commissioning for Urban Rail Transit

Mao Tian

Abstract Two methods of Integration Test and Commissioning (ITC) acceptance are presented in this chapter, and the calculation method, stage division, and interface relation of ITC acceptance are analyzed and illustrated. The ITC acceptance method is divided into expert scoring method and unit engineering acceptance method. The expert scoring method, based on the hierarchical analysis model, is used to determine whether the ITC is passed through data calculation; the unit engineering acceptance method based on process control model, tracks and controls the process of the ITC subproject, subproject and inspection batch acceptance, and gives the determination whether ITC is accepted or not. The ITC acceptance of the interface debugging includes vehicle, communication, signal, power supply, safety door, rail, and other system within the scope of rail line area space and ISCS, FAS, BAS, AFC, ACS, air-conditioning, water supply and drainage, escalator, and other system within the scope of the station.

Keywords Urban rail transit • Integration testing and commissioning
The expert scoring method • The unit project • Acceptance • Interface

1 The Background Analysis of Integration Test and Commissioning for Urban Rail Transit

The integration test and commissioning (ITC) is currently used in the field of urban rail transit and high-speed railway field, although called ITC, there is a great difference between city rail transit and high-speed railway in organization form, test method, and testing technology field. The ITC is a necessary part of dynamic acceptance [1] in high-speed railway, which has a clear definition and requirements; but the urban rail transit ITC definition generally refers to debugging the linkage

M. Tian (✉)

Postgraduate Department, China Academy of Railway Sciences,
100081 Beijing, China
e-mail: tmzsh@163.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_76

747

function among subsystems, and testing the performance of the main equipment system by professional testing equipment after each equipment subsystem completes the body debugging. Then the debugging and testing data are compared with the relevant national standards, industry standards, and design requirements, the detected items that [2] are not up to standard are modified, perfected, and optimized.

When construction units deal with integration test and commissioning, there are a number of factors to be considered such as management and technical capabilities, field installation quality, schedule status, operational security assessment, etc. The integration test and commissioning of urban rail transit have been developed and applied, because many cities in China have adopted the integration test and commissioning to a certain extent at present. The purpose [3] is to verify the function and performance of the equipment system, so as to ensure the smooth opening of the urban rail transit on schedule. However, how to check and accept the ITC project to ensure the design of the various functions and performance indicators to meet the requirements, to determine the availability of trial operations to provide the basis for theoretical and practical analysis is less. In this chapter, according to the practice of ITC in Beijing, Shenyang, and other cities, the contents and methods of ITC are analyzed preliminarily, in the hope of providing some help for integration test and commissioning of urban rail transit.

2 The Phase Division of Integration Test and Commissioning for Urban Rail Transit

After the investigation and analysis of the implementation of the ITC project in different cities, this chapter divides the core phase of the ITC into three important time nodes: the initial stage, the middle stage, and the late stage acceptance period. Single system commissioning [4], operation exercises [5], and reliability test can be added to the debugging range of the specific line according to the actual needs of the project stage division is shown in Fig. 1.

Initial stage of ITC: According to the technical specifications and design drawings of the system, prepare ITC program and system debugging rules, check and verify the single system debugging. All units jointly establish commissioning organization, and to determine the lead unit.

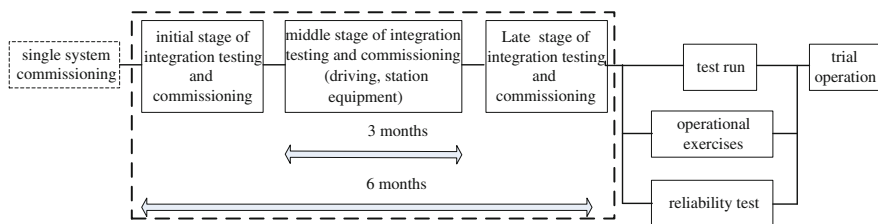


Fig. 1 Stage division about urban rail transit integration testing and commissioning

Mid stage of ITC: The lead unit will organize the commissioning units to carry out the system function debugging and performance test, record the debugging results, and rectify the unqualified items.

Late stage of ITC: The lead unit shall collate and analyze the records of ITC, and determine whether the ITC will be qualified by appropriate methods.

3 The Main Contents and Interface of Integration Test and Commissioning for Urban Rail Transit

The key task of the ITC is to check the functions and performance of each system linkage, and whether the acceptance indicators conform to the technical specifications, design documents, national regulations, and other relevant regulations.

In accordance with the scope of acceptance, the ITC includes the systems associated with the track section [6] and the systems associated with the station. The systems associated with the track section mainly include [7] vehicles, signals, power supply, communications, safety doors, and tracks; the systems associated with the station mainly include supervisory control system (ISCS), automatic fare

Table 1 The main contents and interface of the track section

Integration testing and commissioning associated with driving system	Vehicles	Signal	Communication	Power supply	Screen door	Track
ATO	○	●	○			
Train control center	○	●	○			
ATP stop function	○	●	○			
Emergency braking	●	○		○		
Train and safety door test	●	○	○		○	
Wireless subsystem	●	○	○			
Automatic broadcast system of the train	○	○	●			
Wheel rail relation test	●					○
Pantograph–catenary relationship test	●			○		
Field strength test of communication system	○	○	●			
Signal system interlock function and full running test	○	●	○	○		
Vibration and noise test	●			○	○	○
Electromagnetic compatibility test	○	○	●	○		

“black circle” Major system; “white circle” Coordination system

Table 2 The main contents and interface of the station system

Integration testing and commissioning associated with station system	ISCS	FAS	BAS	Ventilation and air-conditioning	Water supply and drainage	Power lighting	Communication	AFC	PIS	ACS	Escalator	Safety door
Comprehensive test of ISCS	●	○		○	○	○	○	○	○	○	○	○
Comprehensive test of FAS		●		○	○	○	○	○	○	○	○	○
Comprehensive test of BAS			●	○	○	○	○				○	
Comprehensive test of AFC	○	○					○	●				
Comprehensive test of ACS	○	○								●		
Comprehensive test of communication system							●	○				
Stress testing of large passenger flow	○	○						●				

“black circle” Major system; “white circle” Coordination system

collection (AFC), fire alarm system (FAS), building automation system (BAS), passenger information system (PIS), access control system (ACS), power lighting system, water supply and drainage system, escalator system, etc. The main contents and interface of acceptance check are shown in Tables 1 and 2.

4 Acceptance Check Methods of Integration Test and Commissioning for Urban Rail Transit

The acceptance check of ITC refers to the test and acceptance of the results of the combined test in urban rail transit, in order to verify whether the line function and performance meet the design standards and acceptance specifications requirements, and to provide a basic judgment for the entire line whether it can successfully enter the trial operation. According to the practice of ITC for urban rail transit, this chapter puts forward two kinds of acceptance test method, namely expert scoring method and unit project acceptance method, which can provide guidance for the acceptance of ITC, and determine the result of the combined test.

4.1 Expert Scoring Method Model

In the actual acceptance, according to their contribution to the successful acceptance, the achievement of the system functions and performance indicators are divided into two categories, one class index to judge the quality indicators, reflecting the cumulative amount; another index as key indicators, owning one vote veto. The two kind of index is a multiplication relation.

4.1.1 Calculation of Judging the Quality Indicator S

4.1.1.1 Calculation of Weight W of Function and Performance Indicators in ITC

Supposing the statistical indicators (n quantity) for statistics about function and performance are x_1, x_2, \dots, x_n , each indicator weight is $\omega_1, \omega_2, \dots, \omega_n$. If the relative weight of x_1 and x_2 is ω_i/ω_j ($i, j = 1, 2, \dots, n$), a comparison matrix can be obtained as follows:

$$A = \begin{bmatrix} \omega_1 & \frac{\omega_1}{\omega_2} & \dots & \frac{\omega_1}{\omega_n} \\ \frac{\omega_2}{\omega_1} & \omega_2 & \dots & \frac{\omega_2}{\omega_n} \\ \dots & \dots & \dots & \dots \\ \frac{\omega_n}{\omega_1} & \frac{\omega_n}{\omega_2} & \dots & \omega_n \end{bmatrix}. \tag{1}$$

Obviously, A is a consistent positive reciprocal matrix [8], setting vector $W = (\omega_1, \omega_2, \dots, \omega_n)$ T , get the formula as follows:

$$A = W \left(\frac{1}{\omega_1}, \frac{1}{\omega_2}, \dots, \frac{1}{\omega_n} \right). \tag{2}$$

Using Formula (2) and multiplying W on both sides, we get the following:

$$AW = W \left(\frac{1}{\omega_1}, \frac{1}{\omega_2}, \dots, \frac{1}{\omega_n} \right) W = nW. \tag{3}$$

This shows that W is the eigenvector of matrix A , and N is characteristic root. After normalization of W , it can be approximated as the weight vector of A , and we get $AW = \lambda_{\max} W$, λ_{\max} ($\lambda_{\max} = n$) is the largest characteristic root of A , and W is the characteristic vector corresponding to λ_{\max} .

4.1.1.2 Calculation of Expert Scoring Value Y of Function and Performance Indicators in ITC

According to the Likert scale [9], the experts give scores to different indicators, which stand for the completion of the function and performance in ITC. The score is k , as shown in Table 3.

Then, the weighted average is used to calculate the expert scoring value of Y . The formula is as follows:

$$y_i = \frac{\sum_{k=0}^9 k \times n_k}{9}. \tag{4}$$

In the formula, y_i is the average completion value of the i th indicator; n_k is the number of experts with the score of k .

Table 3 k scores and their meanings

k	Scores and their meanings of Likert scale
0	“Uncompleted”
1	“Completed in general -”
2	“Completed in general”
3	“Completed in general +”
4	“Completed good -”
5	“Completed good”
6	“Completed good +”
7	“Completed very well -”
8	“Completed very well”
9	“Completed very well +”

4.1.1.3 Calculation of Judging the Quality Indicator S

On the basis of the results of the first two steps, the expert scoring value Y and the weight W are multiplied to get the quality indicator S . The formula is as follows:

$$S = Y \times W = (y_1, y_2, \dots, y_n) \times (\omega_1, \omega_2, \dots, \omega_n)^T. \quad (5)$$

4.1.2 Calculation of the Key Indicator D

$$D = d_1 d_2 \dots d_i \dots d_n \quad (6)$$

In the formula,

$$d_i = \begin{cases} 1 & \text{Qualified about acceptance of the } i\text{th indicator} \\ 0 & \text{Unqualified about acceptance of the } i\text{th indicator.} \end{cases}$$

4.1.3 Calculation of Function and Performance Comprehensive Evaluation Value M About ITC

The performance comprehensive evaluation value M is equal to the product [10] of judging the quality indicator S and the key indicator D .

The formula is

$$M = SD = d_1 d_2 \dots d_n (y_1, y_2, \dots, y_n) \times (\omega_1, \omega_2, \dots, \omega_n)^T. \quad (7)$$

The formula means the bigger the comprehensive evaluation value M and the bigger the quality indicator value S , the better the ITC result.

4.1.4 The Application of Expert Scoring Method

The advantage of the expert scoring method is that it can evaluate and judge the results of ITC in a short period of time, especially when the installation is limited and the construction period is short, The construction work has not been fully completed at the time of operation, or the ITC link was added to the line as a construction step only in the latter part of construction.

4.2 Unit Project Acceptance Model

The ITC unit acceptance method based on process control is designed in this chapter. The method regards ITC as a unit project, and divides the acceptance of function and performance into three subparts, namely, subsection, sub item and

inspection batch, based on the station, subway section, and equipment specialty. The key monitoring process is the single system debugging, interface debugging, comprehensive debugging, system testing, and so on. Finally, the decision whether to pass ITC is given according to the monitoring situation.

4.2.1 Progressive Division Method of Unit Project of Equipment System

According to the classification method of Metro acceptance code [11], the acceptance scope of the ITC unit project can be divided into subsections, sub items, and inspection batches at intervals. The classification methods are shown in Table 4.

The acceptance scope of ITC unit project includes all equipment system of the urban rail transit line.

The subsection is divided into several sub items according to the function of a complete area or complete equipment system. The subsection working range of ITC can be urban rail transit stations, track area, and vehicle depot area.

The sub item is divided into several inspection batches according to the debugging procedure and the interface relation. The sub item of ITC can be used to test the functional performance of track area and station area.

The inspection batch is the basic unit of ITC acceptance, which is divided into the master control project and the ordinary project acceptance. The master control project is the function and performance indicator that must be finished before operation. These indicators are clearly defined in the contract, specifications, and design documents. The master control project shall not be passed if any an indicator is not passed. The ordinary project includes the indicators, which do not affect the opening of the line and not involve traffic safety and system reliability.

4.2.2 Application Analysis of Unit Project Acceptance Method

Different from the expert scoring method, the data of acceptance results are analyzed and determined, the unit project acceptance method can be used for tracking the whole process of debugging. According to the requirements of the design documents, the project personnel checks each function and performance of every a system, analyze the reliability of debugging, the realization rate of function and performance, etc. Through the process check, each of the real debugging of indicators is found, which provides the judgment information basis for the trial operation.

Table 4 *K* Division of unit project about urban rail transit using integration testing and commissioning

Unit project	Subsection	Sub item	Inspection batch
Urban rail transit line	Main line and station and interval	Station 1 ⋮ Station <i>n</i>	Interface verification, functional debugging, and performance testing among ISCS, BAS, FAS, electrical system, AFC, ACS, escalators, and other system
		Track area	Interface verification, functional debugging and performance testing among vehicle, signal, communication, power supply, track, and other system
	Vehicle depot	Repair depot and parking lot	Interface verification, functional debugging, and performance testing among ISCS, BAS, FAS and other system
		Track area	Interface verification, functional debugging, and performance testing among vehicle, signal, communication, power supply, track, and other system

5 Example Analysis

According to Tables 3 and 4, the key expert scoring *Y* values of urban rail transit equipment can be obtained and calculated as the linear difference. The results are shown in Table 5.

According to the formula (1), the comparison matrix is constructed, and the elements in it are given by the expert of weights:

$$A = \begin{bmatrix} 6/6 & 6/5 & 6/7 & 6/5 & 6/1 & 6/3 & 6/2 \\ 5/6 & 5/5 & 5/7 & 5/5 & 5/1 & 5/3 & 5/2 \\ 7/6 & 7/5 & 7/7 & 7/5 & 7/1 & 7/3 & 7/2 \\ 5/6 & 5/5 & 5/7 & 5/5 & 5/1 & 5/3 & 5/2 \\ 1/6 & 1/5 & 1/7 & 1/5 & 1/1 & 1/3 & 1/2 \\ 3/6 & 3/5 & 3/7 & 3/5 & 3/1 & 3/3 & 3/2 \\ 2/6 & 2/5 & 2/7 & 2/5 & 2/1 & 2/3 & 2/2 \end{bmatrix}.$$

The largest characteristic root $\lambda_{\max} = 7$, and its corresponding characteristic vector is

$$W = (0.21, 0.17, 0.24, 0.17, 0.03, 0.10, 0.07)^T.$$

According to the formula (7), take $d_i = 1$ and calculate the comprehensive evaluation value of the 3 lines, respectively, $M1 = 3.91$, $M2 = 6.72$, $M3 = 6.99$, which is in contrast to the ideal effect, as shown in Fig. 2.

Table 5 The key system function and performance realization of the 3 lines' equipment

The key system of subway equipment	Implementation of function and performance	Line 1	Line 2	Line 3
Signal system	y_1	3.13	5.18	7.22
Communication system	y_2	1	7.22	7.22
Vehicle system	y_3	5.96	7.22	7.22
Power supply system	y_4	5.61	7.22	7.22
Safety door system	y_5	5.78	7.22	7.09
ISCS system	y_6	2.4	7.22	6.00
FAS system	y_7	3.83	7.22	6.70

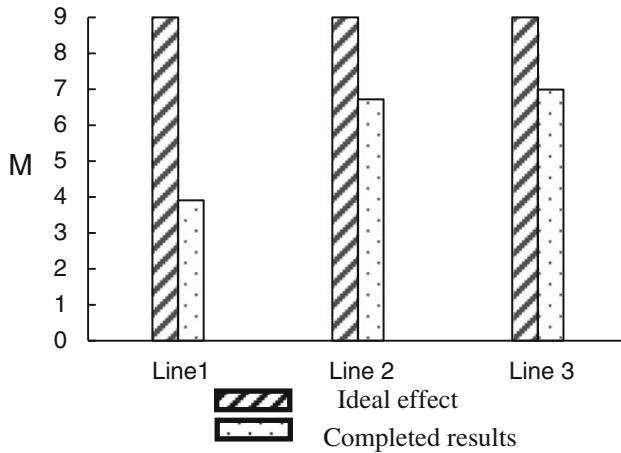


Fig. 2 Comparison of ideal results and actual results of 3 lines using integration testing and commissioning

Through the analysis of Fig. 2, we can see the gap between ITC completed results and ideal effect by design; In ITC, signal system, communication system, vehicle system, and power supply system are the main systems that affect the quality of debugging, and need to be paid attention to.

6 Conclusion

- (1) For urban rail transit ITC, different cities have different debugging methods and programs according to the actual situation, but there is no detailed research and sufficient experience for the acceptance of ITC. In this chapter, two methods of the ITC acceptance are put forward, and the calculation method, stage division,

- main contents, and interface relation of the ITC acceptance are analyzed and studied in detail, which can fully guide the order of the ITC acceptance work.
- (2) The ITC acceptance expert scoring method uses a hierarchical analysis model, and studies the weights of each system in the debugging and the calculation method. According to the data of 3 subway lines, an example is calculated. Through mapping, the gap between the ITC design value and the actual value can be clearly found.
 - (3) This chapter puts forward the acceptance method of the unit project of ITC for the first time. This method regards ITC as a unit project, which is divided into subsections, sub items, and inspection batches of three subparts. Through process control, the single system debugging, interface debugging, comprehensive debugging, and other aspects are tested and accepted, so as to improve the accuracy of acceptance of ITC.
 - (4) Through comparative analysis, we can see that the expert scoring method can be used to analyze, calculate, and judge the acceptance of ITC according to the scores given by experts after the construction; the unit project acceptance method can be used to check, record, analyze, and check the ITC indicators through three substages. The former method focuses on results, while the latter focuses on processes.
 - (5) Through the example analysis, it can be seen that in the case of better control of time and cost, the expert scoring method and the unit project acceptance method should be used together for acceptance analysis, which can achieve more accurate judgment effect.

Acknowledgments Thanks for Fund Project supported by the China Academy of Railway Sciences (1651CG5305) and (1551JJ9304).

References

1. Wang L (2013) Technical and engineering application of urban rail transit integration testing and commissioning. In: China Railway Publishing House (in Chinese)
2. Xiao Y (2014) Research on the evaluation method of the testing and commissioning system for urban rail transit system. *China Railway Sci* 35(4):124–127 (in Chinese)
3. Qin LI (2011) Research on the reliability analysis method of the testing and commissioning system for high-speed railway. *China Railway Sci* 32(2):110–115 (in Chinese)
4. Sun N (2011) The integration testing and commissioning for urban rail transit. In: China Railway Publishing House (in Chinese)
5. Xiao Y (2014) Research on the main technology and application of urban rail transit testing and commissioning. In: China Academy of Railway Sciences (in Chinese)
6. Ministry of Housing and Urban-Rural Development of the People's Republic of China (2006) GB 50578-2010 Code for quality acceptance of urban rail transit signaling system. In: China Planning Press (in Chinese)
7. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China (2013) GB/T30013-2013 basic condition for trial operation of urban rail transit. In: China Planning Press (in Chinese)

8. Han Z (2009) Mathematical modeling method and its application. In: Higher Education Press (in Chinese)
9. Hao W (2012) Core elements of public-private partnership in urban rail transit: case of Beijing Metro Line M. *China Civil Eng J* 45(10):175–180 (in Chinese)
10. Liu Z (2008) Study on sustainable mode of investment and financing of high speed railway in China. In: Tsinghua University (in Chinese)
11. Ministry of Housing and Urban-Rural Development of the People's Republic of China (2006) GB 50381-2006 code for constructional quality acceptance of urban rail transit automatic fare collection system engineering. In: China Planning Press (in Chinese)

Forecast of Short-Term Passenger Flow of Urban Railway Stations Based on Seasonal ARIMA Model

Zhirui Guang, Jun Yang and Jian Li

Abstract Forecast on short-term passenger flow of urban rail transit is the key to network operation and management, and the basis of passenger flow organization and train optimal allocation. In this chapter, a predictive model of passenger flow entering and departing is constructed, based on the model of seasonal Autoregressive Integrated Moving Average (ARIMA). First, outliers of passenger flow in time series were replaced by the linear interpolation method; Second, two methods, after considering weather conditions and air quality, are used for passenger flow forecast respectively. One is seasonal differencing, the other is by adding working day attributes as dummy variables. Third, the method of least squares was used to estimate the weight, thus a combined forecasting model for time series was constructed. After, the model has been calibrated and validated by the historical passenger flow data collected by AFC system of Beijing Metro: the error is less than 5%. This model not only considered dummy variables such as weather conditions, air quality, and working day attributes, but also quantified their impact for passenger flow. The results show that the prediction model has high accuracy.

Keywords Urban rail transit · Short-term passenger flow forecast
Autoregressive Integrated Moving Average (ARIMA) · The method of least squares · Dummy variables

1 Introduction

Urban rail transit's passenger flow forecast, which can be divided into long-term forecast, medium-term forecast, and short-term forecast is the basis of line planning and operation organization [1]. Among them, short-term passenger flow forecast is

Z. Guang (✉) · J. Yang · J. Li
Beijing Metro Network Control Center, No. 6 Xiao Ying North Road,
Yang District, Beijing, China
e-mail: zhiruiguang@163.com

of great importance for operation organization to achieve a quick response to traffic fluctuations and effective allocation of resources to maximize social and economic benefits. Thus, it is of great theoretical and practical significance to forecast short-term passenger flow of urban railway stations.

The main methods of passenger flow forecasting include four-step method of traffic planning [2], nonparametric model [3–5], and statistical model [6–9]. The traditional four-step method, which is mainly used in line planning, is not timely because of fluctuations in bad weather, air quality, and other factors. Nonparametric models, such as Support vector machine (SVM), fuzzy logic, and neural network model, need massive data for iterative calculation. Meanwhile, it is not easy to track the prediction error, as a result, the operation management cannot be guided effectively, due to its “black box” learning mode. In order to overcome problems above, this chapter chose ARIMA model, after considering weather, land use, and other influencing factors, based on the statistical analysis of historical passenger flow.

2 Prediction Models

Based on the analysis of the historical passenger flow under normal circumstances, a combined forecasting model for time series was constructed. First, outliers of passenger flow in time series were replaced by the linear interpolation method; Second, two methods, after considering weather conditions and air quality, are used for passenger flow forecast respectively. One is seasonal differencing, the other is by adding working day attributes as dummy variables. Third, the method of least squares was used to estimate the weight.

2.1 *Outliers Processing of Short-Term Passenger Flow*

The accuracy of traffic time series prediction for normal circumstances is affected by outliers including holidays and vacations, and one day before and after the holidays. Given the passenger flow of same working day attribute has the characteristics of short-term stability, outliers were replaced by the linear interpolation method. In mathematics, linear interpolation is a method of curve fitting using linear polynomials to construct new data points within the range of a discrete set of known data points.

The specific linear interpolation method is applied as follows:

If the two known points are given by the coordinates (x_0, y_0) and (x_1, y_1) , the linear interpolant is the straight line between these points. For a value x in the interval (x_0, x_1) , the value y along the straight line is given from the equation which can be derived geometrically from the figure below (Fig. 1).

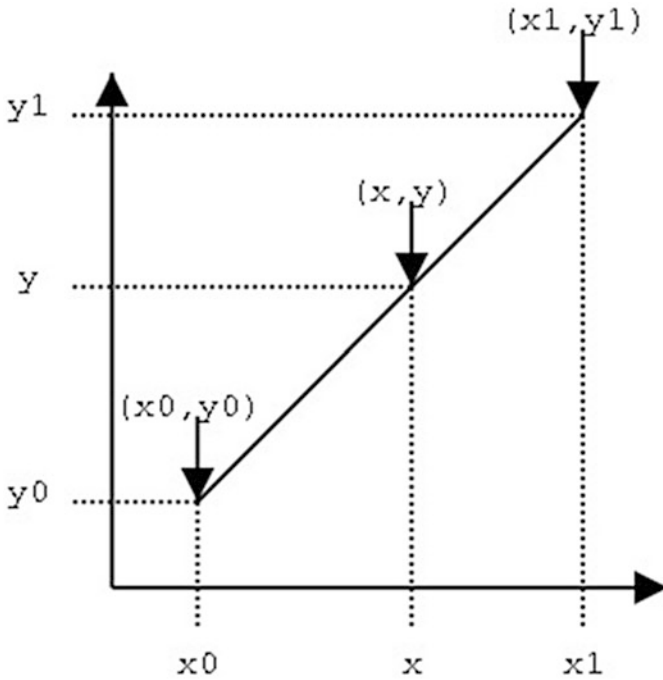


Fig. 1 Schematic diagram of the model

$$\frac{y - y_0}{y_1 - y_0} = \frac{x - x_0}{x_1 - x_0}, \tag{1}$$

$$y = \frac{x - x_0}{x_1 - x_0} (y_1 - y_0) + y_0. \tag{2}$$

2.2 The Seasonal ARIMA Models

Considering the complex relationship among the seasonal effect, long-term trend effect, and random fluctuations, the multiply ARIMA on the forecast of the entrance and exit passenger flow under normal circumstances was established. At the same time, weather conditions and air quality is added as dummy variables to prevent from their influences.

Working day attributes (such as Sunday, Monday) have a direct impact on the entrance and exit passenger flow. In method 1, seasonal differencing is used to remove the changes by working day attributes, and to stabilize passenger flow time series. In method 2, working day attributes were added as dummy variables in order to accurately grasp the influence of each working day attribute on passenger flow.

For example, when day 1 = 1 on Monday, the other days' day 1 = 0; when day 5 = 1 on Friday, the other days' day 5 = 0.

Seasonal ARIMA models $ARIMA(p, d, q) \times (P, D, Q)_S$ are usually denoted, where S refers to the number of periods in each season, and the lowercase $p, d,$ and q refer to the autoregressive, differencing, and moving average terms for the non-seasonal part of the ARIMA model, and the uppercase $P, D,$ and Q refer to the autoregressive, differencing, and moving average terms for the seasonal part.

The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The seasonal ARIMA model [10, 11] formula is as follows:

$$\begin{aligned}
 &ARIMA(p, d, q) \times (P, D, Q)_S, \\
 &\nabla^d \nabla_S^D x_t = \frac{\Theta(B)\Theta_S(B)}{\Phi(B)\Phi_S(B)} \varepsilon_t, \\
 &\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q \\
 &\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p \\
 &\Theta_S(B) = 1 - \theta_1 B^S - \dots - \theta_Q B^{QS} \\
 &\Phi_S(B) = 1 - \phi_1 B^S - \dots - \phi_P B^{PS},
 \end{aligned} \tag{3}$$

where B refers to delay operator.

2.3 The Combined Model

In order to improve the prediction accuracy, the linear combination forecasting model is used to combine the two ARIMA models, and the combined model formula is as follows:

$$\hat{F}(t) = \hat{\beta}_1 \hat{F}_1(t) + \hat{\beta}_2 \hat{F}_2(t), \tag{4}$$

where $\hat{\beta}_1$ and $\hat{\beta}_2$ are the weights of the two prediction methods.

The target of the forecast weight calibration is to minimize the error between the predicted value and the actual value as far as possible. That is the minimum sum of the squares residuals. The combination forecasting parameters can be expressed as the optimization model as follows:

$$\begin{aligned}
 \min Q &= \sum (F(t) - \hat{\beta}_1 \hat{F}_1(t) - \hat{\beta}_2 \hat{F}_2(t))^2 \\
 s.t. & \\
 \hat{\beta}_1 + \hat{\beta}_2 &= 1 \\
 \hat{\beta}_1, \hat{\beta}_2 &\in [0, 1]
 \end{aligned}
 \tag{5}$$

3 Case Study

In this chapter, the passenger flow data collected by the AFC of Beijing urban rail transit were used to calibrate the parameters of the model. Taking Tiantongyuan station as an example, the model parameters were calibrated by the time series from March to July in 2016. Then the calibrated model was used to dynamically predict the passenger flow in August and September of that year. Finally, the errors among two ARIMA models and the combined forecasting model were compared and analyzed.

In order to judge the accuracy of the model, the mean absolute error (MAE) and the mean absolute percentage error (MAPE) were used to analyze the prediction results.

First, outliers (holidays and vacations, and one day before and after the holidays) have been replaced by the linear interpolation method. Take passenger flow at $t-7$ and $t+7$ as the known point, and then calculate passenger flow at t in outliers.

Second, a stationary test of corrected time series of station quantities is carried out. As the Fig. 2 shows, passenger flow of Tiantongyuan station is a nonstationary time series. It takes 7 days as a cycle to fluctuate. By a first and seasonal difference, the residuals fluctuated randomly near 0, thus it is a stationary time series.

Then a white noise test was carried out for this stationary time series. The results show that the autocorrelation and partial autocorrelation coefficients (Fig. 3) are greater than 2 times the standard deviation. And this time series has short-term correlation after difference, thus it is a stationary non-white noise time series. And two methods, after considering working day attributes, weather conditions, and air quality are used for passenger flow forecast as follows.

Method 1: Modeling of autocorrelation and partial autocorrelation coefficients

1. After a first and seasonal difference, the behavior of ACF and PACF both tails off, thus the $ARMA(1, 1)$ model is used to extract the short-term autocorrelation information of the post-differential sequence;
2. ACF of lag 7 was significantly not 0, but ACF of lag 14 was 0. PACF of lag 7, 14, and 21 were significantly not 0. Thus the behavior of ACF cuts off, and PACF tails off. Then seasonal $ARMA(0, 1)_7$ model is used to extract the seasonal autocorrelation information of the post-differential sequence;

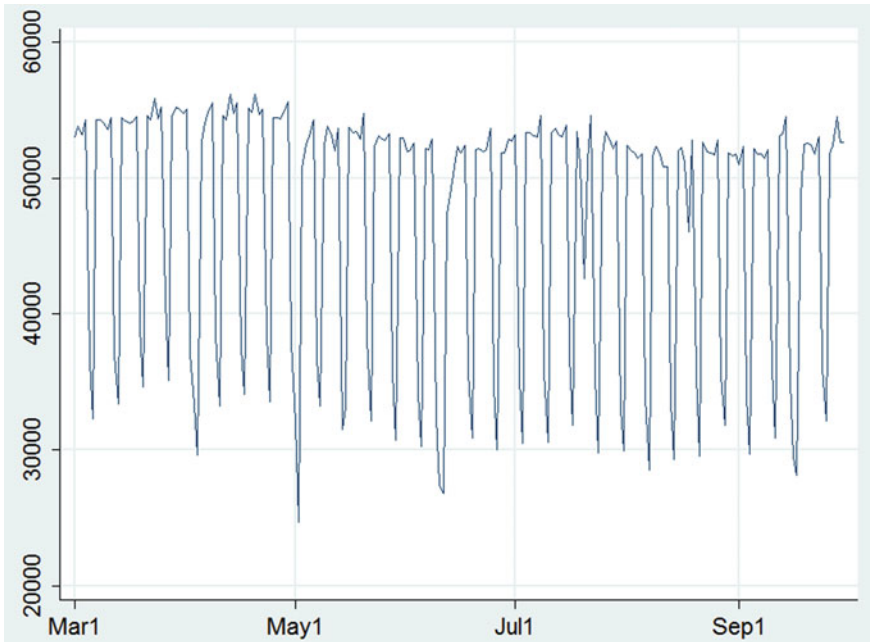


Fig. 2 Passenger flow of Tiantongyuan station

3. The model $ARIMA(1, 1, 1) \times (0, 1, 1)_7$, which combined the models above, was constructed by adding dummy variables (weather conditions, air quality, and working day attributes). The fitting results showed that the coefficient of air quality was not statistically significant, thus this dummy variable is removed, and other factors were significantly correlated except $AR(1)$, thus $AR(1)$ is removed. The residuals were a non-white noise sequence.
4. Model $ARIMA(0, 1, 1) \times (0, 1, 1)_7$ passed the white noise test. The prediction equation is as follows:

$$F_1(t) = F_1(t-7) + F_1(t-1) - F_1(t-8) + -4.114 - 623.261x_{\text{wea}} - 0.724\varepsilon_{t-1} - 0.812\varepsilon_{t-7} + 0.724 \times 0.812\varepsilon_{t-8}. \quad (6)$$

Method 2: Modeling by adding working day attributes as dummy variables

To quantify their impact for passenger flow, method 2 sets a model $ARIMA(1, 0, 1) \times (1, 0, 1)_7$ by adding dummy variables such as weather conditions, air quality, and working day attributes. The fitting results showed that except coefficients of seasonal $AR(1)$ and $MA(1)$ were not statistically significant, other factors were significantly correlated. They are still useful, because the residuals would not pass the white noise test if the two terms were removed.

The prediction equation is as follows:

$$\begin{aligned}
 F_2(t) = & 32058.24 - 624.71x_{wea} + 21186.55x_{day1} + 21355.49x_{day2} + 21426.77x_{day3} \\
 & + 21098.23x_{day4} + 21998.37x_{day5} + 4315.71x_{day6} + 0.969(F_2(t-1) - \hat{F}_2(t-1)) \\
 & + 0.187(F_2(t-7) - \hat{F}_2(t-7)) - 0.969 \times 0.187(F_2(t-8) - \hat{F}_2(t-8)) + \varepsilon_t \\
 & - 0.706\varepsilon_{t-1} - 0.053\varepsilon_{t-7} + 0.706 \times 0.053\varepsilon_{t-8}.
 \end{aligned}
 \tag{7}$$

From the formula, we can see impact extent of the working day attributions on the passenger flow. The largest impact to passenger flow lies in Friday, and the smallest impact to passenger flow in the weekend.

Finally, the combination forecasting model has been calibrated by the method of least squares. The combination prediction equation is as follows:

$$\hat{F}(t) = 0.2 \times \hat{F}_1(t) + 0.8 \times \hat{F}_2(t).
 \tag{8}$$

Table 1 Prediction performance of different models

Variable	Mean	Median	Std. Dev.
Error of the combined model	935	796	876
Error of method 1	1643	1180	1410
Error of method 2	1101	918	882
Error rate of the combined model (%)	2.20	1.67	2.25
Error rate of method 1 (%)	3.77	2.41	3.59
Error rate of method 2 (%)	2.51	1.81	2.31

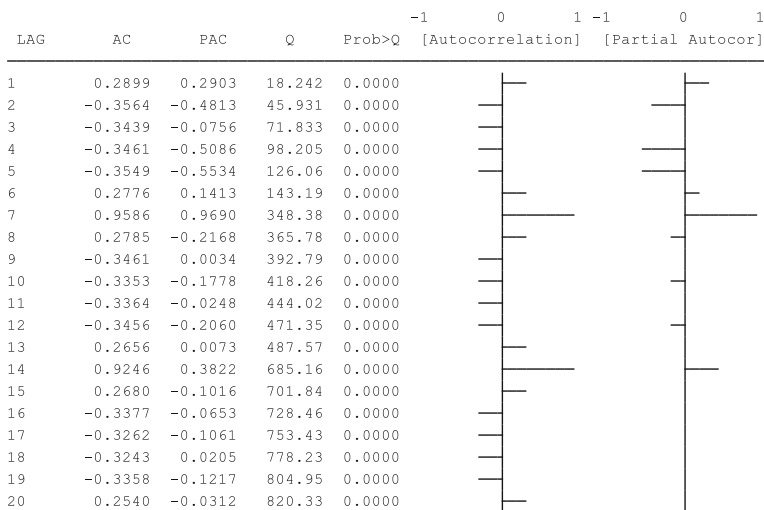


Fig. 3 Figure of the ACF and PACF

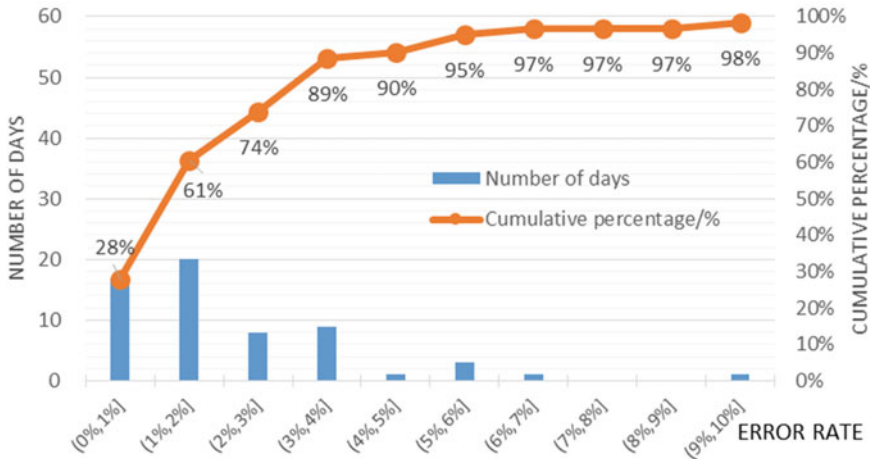


Fig. 4 Error analyses for the combined forecasting model

The table of prediction error (Table 1) showed that the mean absolute error among the three models is less than 1700 people, and the mean absolute percentage error is less than 4%. Among them, the error of the combined forecasting model is the smallest: the mean absolute error is 935 people, and the mean absolute percentage error is 2.20%.

The error statistics (Fig. 4) showed that prediction errors of 89% days are less than 4%, indicating that the combined forecasting model has high accuracy and good applicability.

4 Conclusion

The study on predicting the short-term passenger flow is of great practical importance. In this chapter, two short-term passenger flow forecasting models were set up based on the statistical analysis of historical passenger flow after considering the working day attributes, the weather, air quality, and other influencing factors. Then the combination forecasting model was calibrated by the method of least squares. Finally, the model is evaluated with extensive tests carried out from Beijing rail transit network. The results show that the proposed method is easier, reliable, and helpful to urban rail transit operation and management decision-making.

Due to the complex factors of urban rail transit passenger flow, it is necessary to carry out post evaluation of the passenger flow forecast results, and analyze the specific reasons for the error. The model established in this chapter still needs to be improved and optimized on the basis of distinguishing controllable and uncontrollable factors.

Acknowledgments This research is supported by the State Key Laboratory of Rail Traffic Control and Safety (Contract No. I11K00050, Beijing Metro Network Control Center).

References

1. Vlahogianni EI, Karlaftis MG, Golias JC (2014) Short-term traffic forecasting: where we are and where we're going. *Transp Res Part C Emerging Technol* 43:3–19
2. Chunfu S (2004) *Traffic planning*. China Railway Publishing House, Beijing (in Chinese)
3. Zhou J, Zhang D (2014) Direct ridership forecast model of urban rail transit stations based on spatial weighted LS-SVM. *J China Railway Soc* 36(1):1–7
4. Li X, Lv X (2011) Forecast of railway short-term passenger flow based on RBF neural network *Railway Transp Econ* 33(6):86–89. (in Chinese)
5. Khan A, Bayesian M (2012) predictive travel time methodology for advanced traveler information system. *J Adv Transp* 46(1):67–79
6. Mu T, Jiang J, Wang Y (2013) Heterogeneous delay embedding for travel time and energy cost prediction via regression analysis. *Intell Transp Syst IEEE Trans Intell Transp Syst* 14(1): 214–224
7. Lina WANG (2011) *Traffic forecasting and scheduling of the urban passenger rail network based on historical data*. Beijing Jiaotong University, Beijing (in Chinese)
8. Dewei L, Yixing Y, Xianfeng Z (2017) Combined short-term prediction model of station entry flow in urban rail transit. *Urban Rapid Rail Transit* 30(1):54–58 (in Chinese)
9. Junchang C, Enjian Y (2014) Prediction of urban railway station's passenger entrance and exit flow based on multiplies ARIMA model. *J Beijing Jiaotong University* 38(2):135–140
10. Geoge E, Box P (1997) *Time series analysis: forecasting and control*. In: Lan G, (ed) FAN Jinheng, trans. China Statistic Press, Beijing, 101–135. (in Chinese)
11. Wang Y (2005) *Applied time series analysis*. China Renmin University Press, Beijing. (in Chinese)

Research on the Route Choice Behavior of Subway Passengers Based on AFC Data

Liping Xie, Haiying Li and Xinyue Xu

Abstract This paper studies the route choice behavior of passengers from auto fare collection and timetable data using a method combined with Bayesian and Metropolis–Hasting sampling. First, influential factors of route choice such as in-vehicle travel time, transfer time, and in-vehicle crowding are selected. Then, formulations of these factors are established for a single passenger, which are merged into a logit model to model route choice behavior of subway passengers. Next, an algorithm that integrates Bayesian inference and Metropolis–Hasting sampling is designed to calibrate the parameters of the logit model. Finally, a case study of Beijing subway is applied to verify the validity of the developed model and algorithm.

Keywords Subway · Big data · Route choice · Bayesian · Crowding

1 Introduction

Along with the operation in network wise in Chinese subways, route choice behavior has become essential in demand forecasting, and even passenger flow management during peak hours [1]. Strategies for passenger flow control have been applied to solve the daily crowding problems in Chinese subways by guiding passengers away from the crowded routes [2]. Therefore, understanding the nature of route choice behavior is in urgent need to relieve the tension of congestion effectively during peak hours in subways.

AFC data becomes possible for subways to learn the route choice behavior of passengers. AFC data consists of the entry station, exit station, and the corresponding

L. Xie (✉) · H. Li · X. Xu
State Key Laboratory of Rail Traffic Control & Safety,
Beijing Jiaotong University, Beijing, China
e-mail: 15120897@bjtu.edu.cn

L. Xie
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

timing when a passenger swipes the card. Besides, each smart card is assigned a unique serial number, that is, the movement of an individual passenger can be tracked by the key serial number.

To our knowledge, few studies have considered the influence of crowding on passenger route choice behavior, despite its significance is proven [3, 4]. The statistics feature of various parts of the travel time can be captured by mining records of a large number of passengers [5], which ultimately and effectively overcomes the difficulty to reproduce details of the whole trip for each passenger.

The research applies a Bayesian inference approach to model route choice behaviors using AFC data of subway. First, we construct a data-driven model of passenger route choice behavior by AFC data. Then, its parameters are calibrated using the Bayesian and MH sampling method. Furthermore, Markov Chain Monte Carlo (MCMC) method is employed to compute the Bayesian posterior jointly, providing a brute force approach for calculating complex models whose maximum likelihood is hard to achieve through the conventional optimization process.

The remainder of the paper is structured as follows: In Sect. 2, factors influencing the route choice behavior are analyzed. In Sect. 3, we construct a logit model of route choice behavior whose parameters are calibrated by an integrated Bayesian and MCMC sampling method based on AFC and train timetable data. In Sect. 4, a case study is done to validate the proposed methods. Finally, in Sect. 5, we draw the conclusion of this study.

2 Modeling of Passenger Route Choice Based on AFC Data

2.1 Effect Factors of Route Choice

In-vehicle travel time, transfer time, and in-vehicle crowding are selected as the principal factors by [6–9] based on an SP survey, which was conducted for a month during the peak hours at workdays in December 2015. Note that the in-vehicle crowding is proposed based on the traffic volume and train capacity, which can be expressed as

$$Y_a(x_a) = \begin{cases} \frac{x_a - z_a}{z_a} B & z_a < x_a < C_a \\ \frac{x_a - z_a}{z_a} B + \frac{x_a - C_a}{C_a} D & x_a > C_a, \end{cases} \quad (1)$$

where Y_a , x_a , z_a , and C_a are the congestion coefficient, traffic flow volume, seat number of the train, and maximum capacity of the train in a section a of subway networks, respectively; B and D are the calibration parameters.

2.2 A Logit Model for Passenger Route Choice

The route choice behavior is examined by a classical logit model in the subway network. The utility of a selected route r from the feasible set $\Omega_{n,o,d}$ for a passenger n of the OD pair (o, d) is calculated by

$$U_{n,o,d}^r = u_{o,d}^r + \varepsilon_{o,d}^r \quad \forall r \in \Omega_{n,o,d}, (o, d) \in W, \tag{2}$$

where $u_{o,d}^r$ is the generalized cost and $\varepsilon_{o,d}^r$ is a random item with $E[\varepsilon_{p,w}^r] = 0$.

The generalized cost $u_{o,d}^r$ can be further calculated by

$$u_{o,d}^r = \theta_1 R_r + \theta_2 T_r + \theta_3 Y_a, \tag{3}$$

where θ_1 , θ_2 , and θ_3 are the corresponding weight coefficients of in-vehicle time, transfer time, and in-vehicle crowding perception level respectively. R_r and T_r are in-vehicle travel time and transfer time for the route r .

By the rules of utility maximization, the route choice probability of passengers choosing route r between pair (o, d) is given as

$$P_{n,o,d}^r = \frac{\exp(u_{o,d}^r)}{\sum_{r \in \Omega_{n,o,d}} \exp(u_{o,d}^r)}. \tag{4}$$

Finally, the problem of modeling passenger route choice can be transformed into the estimation of parameters, the variables of passenger perception $\theta = (\theta_1, \theta_2, \theta_3)$. Next, we will introduce methods to achieve these estimations.

3 Calibration Method of Passenger Route Choice Model

3.1 Bayesian Inference Processes

Let Γ_w is the set of data observations obtained from passengers traveling on an OD pair w , including travel time and crowding information. Let $\Gamma = \cup_{w \in W} \Gamma_w$ be the whole set of all OD pairs. If we know the prior distribution of parameters, their posterior probability density $\pi(\theta|\Gamma)$ can be inferred by the Bayesian criterion as follows:

$$\pi(\theta|\Gamma) = \frac{p(\Gamma|\theta)\pi(\theta)}{p(\Gamma)}, \tag{5}$$

where $\pi(\theta)$ is the joint prior probability density of parameter θ , and $p(\Gamma|\theta)$ is the probability of observing data Γ conditional on all unknown parameters. If the full

probability $p(\Gamma)$ is a constant factor, this equation can be simplified into $\pi(\theta|\Gamma) \propto p(\Gamma|\theta)\pi(\theta)$.

If all unknown parameter vectors are independent, we have

$$\pi(\theta|\Gamma) \propto p(\Gamma|\theta)\pi(\theta), \tag{6}$$

where $\pi(\theta)$ is the prior probability density of an unknown parameter θ . Note that $p(\Gamma|\theta)$ equals the likelihood of all parameters given by T observations, which is

$$p(\Gamma|\theta) = \prod_{w \in W} p(\Gamma_w|\theta). \tag{7}$$

Further, the probability of observation data $z_n \in \Gamma_w$ from a passenger n also depends on its alternative routes with more than one possible route of an OD pair w . Therefore, the probability of observing data z_n of an OD pair w can be expressed as

$$p_w(z_n|\theta) = \prod_{r \in R_w} g(z_n|r, \theta) p_{n,w}^r(\theta), \tag{8}$$

where $g(z_n|r, \theta)$ is the conditional probability of observing data z_n at a given route r of an OD pair w , and $p_{n,w}^r(\theta)$ is calculated by Eq. (4). Note that $g(z_n|r, \theta)$ is a normal distribution function based on the assumption that running time of each section follows normal distribution independently.

Therefore, the likelihood of the OD pair w can be given as follows:

$$p(\Gamma_w|\theta) = \prod_{z_n \in \Gamma_w} \left(\prod_{r \in R_w} g(z_n|r, \theta) p_{n,w}^r(\theta) \right). \tag{9}$$

Finally, we can rewrite the posterior probability the likelihood of the OD pair w by substituting Eqs. (7) and (9) into Eq. (6) as follows:

$$\begin{aligned} \pi(\theta|\Gamma) &\propto p(\Gamma|\theta)\pi(\theta) \\ &\propto \prod_{w \in W} \left(\prod_{z_n \in \Gamma_w} \left(\prod_{r \in R_w} g(z_n|r, \theta) p_{n,w}^r(\theta) \right) \right) \pi(\theta). \end{aligned} \tag{10}$$

So, point estimations for the unknown parameters can be given by the posterior probability distribution with the observing data from AFC, if we know the exact prior distribution of the unknown parameters θ . The detailed way to calculate posterior probability is explained in the next chapter.

3.2 MCMC Algorithm

The main steps are as follows:

Step 1: Generate an initial value $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)})$ and set $t = 1$.

Step 2: Initialize the parameter position, set $i = 0$

Step 3: MH sampling

Step 3.1: $i = i + 1$

Step 3.2: The candidate θ_i^* can be randomly selected from the proposed distribution function $q(\theta_i/\theta_i^{(t)})$ with the current state of the parameter $\theta_i^{(t)}$, which means $\theta_i^* \sim q(\theta_i^*/\theta_i^{(t)})$.

Step 3.3: Calculate the acceptance probability of θ_i^* .

$$\rho = \min \left\{ 1, \frac{\pi'(\theta_i^*)}{\pi'(\theta_i^{(t)})} \right\} \tag{11}$$

$$\pi'(\theta_i^*) = p(\Gamma|\theta_i^*, \theta_{-i}^{(t)})\pi(\theta_i^*, \theta_{-i}^{(t)})/p(\Gamma) \propto p(\Gamma|\theta_i^{(t+1)}, \theta_{i+1}^{(t)}, \theta_{i+4}^{(t+1)})\pi(\theta_i^*) \tag{12}$$

$$\pi'(\theta_i^{(t)}) = p(\Gamma|\theta_i^{(t)}, \theta_{-i}^{(t)})\pi(\theta_i^{(t)}, \theta_{-i}^{(t)})/p(\Gamma) \propto p(\Gamma|\theta_i^{(t+1)}, \theta_{i+1}^{(t)}, \theta_{i+4}^{(t)})\pi(\theta_i^{(t)}) \tag{13}$$

Step 3.4: After generating a number u from Uniform (0, 1) distribution, we should determine whether the parameter is updated. If $u \leq \rho$, accept the candidate and set $\theta_i^{(t)} = \theta_i^*$. Otherwise, set $\theta_i^{(t)} = \theta_i^{(t-1)}$.

Step 3.5: Determine whether the current sampling process is going on or not. If the constraint $i < |A| + 3$ is satisfied, return to Step 3.1. Otherwise, stop sampling, and the generation of the parameter vector $\theta^{(t)}$ in the iteration of t is done.

Step 4: Check the constraint that the current iteration number t is less than the maximum iteration number M .

If $t < M$, set $t = t + 1$, and return to Step 2. If not, go on to the next step.

Step 5: Obtain the value of each parameter.

After collecting each parameter 0 of the Markov chain, we can achieve the value of this parameter by

$$\theta_i = \frac{1}{M - B^*} \sum_{t=B^*+1}^M f(\theta_i^{(t)}), \tag{14}$$

where B^* is fixed to represent the number of iterations of the burn-in phase that has been discarded.

4 Result Analysis

A case study, consisting of Line 1, Line 2, Line 4, and Line 6 of Beijing Subway, is chosen to verify the proposed algorithm during morning peak hours. Learning route behavior of passengers in this area is in urgent need, which is helpful to make policies to reduce crowding and improve service.

4.1 Data and Parameter Setting

We collected a primary data set such as the AFC data of several days during December 2015 from the Beijing Mass Transit Railway Operation Corporation Limited, as well as average running time at sections, headways of lines, and loading rates of sections. Transfer time and transfer walking time in each transfer station were both obtained through the survey of stations during December 20–25, 2015 during peak hours in Beijing subway. Regarding the in-vehicle crowding factor, the number of vehicle seats z_a is 276, the train capacity C_a is 1452 persons, and the correction coefficient B and D values are 1. Furthermore, headway of each line is 2 m 30 in the peak hour, and waiting time of passengers for boarding a train is considered as a half headway of the relevant line.

4.2 Analysis of Markov Chains of Parameters

We consider a large number $M = 12,000$ as the maximum iterations to identify the valid Markov chain of each parameter with the implementation of the proposed algorithm by MATLAB. Some results are given as shown in Fig. 1, which indicate that the burn-in period of these parameters with $B^* = 5000$ is suitable.

From Fig. 1, perceptual parameters of passenger’s in-vehicle travel time θ_1 , transfer time θ_2 , and in-vehicle crowding θ_3 are -1.73 , -2.32 , and -1.48 ,

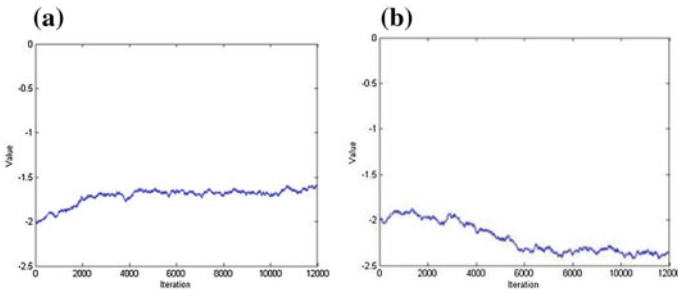


Fig. 1 Markov chains of relevant parameters: **a** θ_1 ; **b** θ_2

Table 1 Route choice probabilities of passengers in some OD pairs

OD	Route no	Route	Transfer time	Cost	Portability
SH-BJN	1	SH-JGM-DD-XD-XWM-BJN	1	53.43	0.237
	2	SH-JGM-BJZ-CWM-XWM-BJN	2	51.238	0.763
XD-BJZ	1	XD-DD-JGM-BJZ	1	25.295	0.966
	2	XD-XWM-CWM-BJZ	1	31.042	0.034
SH-HDHZ	1	SH-JGM-DD-XD-PAL-XZM-HDHZ	1	76.249	0.407
	2	SH-JGM-CYM-DZM-YHG-JST-XZM-HDHZ	2	75.636	0.574
	3	SH-JGM-CYM-DS-NLGX-PAL-XZM-HDHZ	3	81.53	0.019

respectively. By Eq. (4), route choice probabilities for passengers are calculated, as shown in Table 1.

We can also find the following results:

- (1) θ_1 and θ_2 are -1.73 , -2.32 , respectively, which means $\theta_2/\theta_1 > 1$. It is in accordance with the existing result [10], which proves that transfer time plays a key role in the choice of passenger path.
- (2) Passengers' subjective perception of transfer time is stronger than that of in-vehicle time, because one needs to consume physical energy during transfer time.
- (3) In-vehicle crowding may increase passengers' anxiety, stress, and feeling of exhaustion, because they have to stand or to share a limited space with other passengers during the crowding routes [3, 11, 12], but commute passengers have a certain understanding of the route, so in-vehicle crowding is not associated with the timeliness.

5 Conclusions

This paper presents a Bayesian inference approach to model route choice behavior using big data such as AFC, timetable, and related operational data. First, a discrete logit model of route choice behavior is proposed considering in-vehicle time, transfer time factors, and in-vehicle crowding. Then, the data mining approach integrating Bayesian and MH sampling method is employed to calibrate the corresponding parameters. Finally, a case study is made to verify the proposed model and algorithm.

The main contributions of this paper are as follows:

- (1) The influence of in-vehicle crowding to subway passengers on route choice behavior is for the first time explicitly considered in the process of model development using AFC data. Our method has the advantage of RP method than traditional SP methods with cheaper, faster, more reliable, and reproducible data.
- (2) The adopted AFC data-based mining approaches such as Bayesian inference and MH sampling method can solve the difficult problem of calibration parameters with high-dimensional probability distributions existing in the traditional analytical methods.
- (3) Our method provides a wider analysis framework for passenger demand of subways, which helps to make efficient strategies of crowding control. With the help of other data such as timetable and load factors, the exact route and transfer stations can be inferred from AFC data, as well as the exact trains, the OD demand matrix, and the section passenger flow.

Acknowledgments This work is financially supported by the National Natural Science Foundation of China (No. 71601018, 61403288), and the Beijing Natural Science Foundation (No. 9164033).

References

1. Eboli L, Mazzulla G (2008) A stated preference experiment for measuring service quality in public transport. *Transportation Planning and Technology* 31(5):509–523
2. Xu X, Liu J, Li H et al (2016) Capacity-oriented passenger flow control under uncertain demand: algorithm development and real-world case study. *Transp Res Part E Logistics Transp Rev* 87:130–148
3. Li Z, Hensher DA (2011) Crowding and public transport: a review of willingness to pay evidence and its relevance in project appraisal. *Transp Policy* 18(6):880–887
4. Prud'homme R, Koning M, Lenormand L et al (2012) Public transport congestion costs: the case of the Paris subway. *Transp Policy* 21:101–109
5. Sun L, Lu Y, Jin JG et al (2015) An integrated Bayesian approach for passenger flow assignment in metro networks. *Transp Res Part C Emerg Technol* 52:116–131
6. Sadhukhan S, Banerjee UK, Maitra B (2014) Commuters' perception towards transfer facility attributes in and around metro stations: experience in Kolkata. *J Urban Plan Develop* 141(4): 04014038
7. Chen T, Mao B, Gao L et al (2007) Research about passenger travel choice behavior of dedicated passenger railway line. *J China Railway Soc* 3:001
8. Kim KM, Hong SP, Ko SJ et al (2015) Does crowding affect the path choice of metro passengers? *Transp Res Part A Policy Pract* 77:292–304
9. Jara-Díaz S, Gschwender A (2003) Towards a general microeconomic model for the operation of public transport. *Transp Rev* 23(4):453–469

10. Raveau S, Guo Z, Muñoz JC, Wilson NH (2014) A behavioural comparison of route choice on metro networks: time, transfers, crowding, topology and socio-demographics. *Transp Res Part A Policy Pract* 66:185–195
11. Tirachini A, Hensher DA, Rose JM (2013) Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transp Res Part A Policy Pract* 53:36–52
12. Sun S, Li H, Xu X (2017) A key station identification method for urban rail transit: a case study of Beijing subway. *Promet-Traf Transp* 29(3):267–273

A Model of High-Density Passenger Boarding and Alighting in Urban Rail Transit Station

Ning Jia and Yanhui Wang

Abstract Passengers' boarding and alighting is an important link between station and the train in urban rail transit system. The paper presents features of boarding and alighting of passenger flow. A model was established with two continuity equations and a formula. In that model, continuity equations about boarding and alighting are deduced based on these two characteristics, revealing the relationship between time and density. The formula of resultant about passenger flow during their boarding and alighting activities is aimed at exposing the main direction of the whole procedure. What's more, the model was applied to Chongwenmen station in Beijing subway Line 1, the model is simplified for simulation by a mathematical application software-MATLAB, graphing directly. The result is formally described and experimented in experimental and real-world situation, which is proved to be correct and reliable. The model provides the theory foundation for safer process of boarding and alighting for actual operation and plays an important guiding meaning in daily management.

Keywords Urban rail transit · Boarding and alighting · Model Simulation · MATLAB

1 Introduction

In all the service facilities of urban rail transit, subway door is the only way to get on and off. For increasing the passenger flow and limitation of entrance's apertures, boarding and alighting are becoming a key point in daily transportation activities. Fluency of this part not only effects train dwell time but also means the safety of passengers in life and property. Therefore, a study on the characteristics of

N. Jia (✉) · Y. Wang

State Key Laboratory of Rail Traffic Control and Safety, School of Traffic and Transportation, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 15120945@bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_79

779

passenger flow and a describing model present an important guiding meaning in the actual operation.

The research on boarding and alighting of urban rail transit stems from Kraft [1]. He presented the time was the main factor for train's dwelling. Thereafter, Weston [2] built a mathematical model for train's drawing-up related to passenger. Lam et al. [4] brought a new regression model forward about the influence of passenger number. Wiggenraad [3] raised the connection between the width of door and time of on/off. For the feature of unevenly distributed passenger on the platform, Wu et al. [5] utilized four criteria to quantify the standard of crowd. What's more, other scholars are also contributed to the model of boarding and alighting [6–11].

This paper combines passenger boarding and alighting in urban rail transit station with physics, by observing the situation of passengers' self-organization, we describe passenger activities in mathematic models with time and density variables. According to the Newton's second law and the theorem of momentum, a formula for passengers' resultant when boarding and alighting has been set. Afterward, based on the data from Chongwenmen station in Beijing subway Line 1, some examples are conducted to prove the availability of my theory.

2 Problem Statement

2.1 *The Similarity Between Passenger Flow and Liquid Flow*

Liquid is one of the three major forms of the world without definite shape. When the shape of the container is fixed, the liquid tends to flow in the same open direction.

In the process of high-density boarding and alighting, the distance between passengers is small. In the front of some door, the boarding passengers move into the train collectively, and the passengers waiting for getting off move out of the train together, which conform the figures of liquid flow under some kind of force. Therefore, the high-density boarding and alighting passenger flow have the property of liquidity.

Because of the features of liquidity, some mathematical relationships of liquid can be applied to this problem. And the typical motions are as follows:

- a. Flow discharge (Q), $\text{peo}/(\text{m} \times \text{s})$.
- b. Flow rate (μ), m/s .
- c. Flow density (ρ), peo/m^2 .

In the process of high-density boarding and alighting, the relationship among discharge, rate, and density of passenger flow is similar to liquid.

$$Q = \rho \times \mu \tag{1}$$

2.2 Model Research

The high-density passenger can be treated as liquid. In the process of waiting, the waiting area of boarding and alighting can be treated as containers. It can be revealed that the shape of passengers can be treated as isosceles trapezoid. The time and stability of boarding and alighting are related to some parameters.

For discrimination, we introduce definitions as follows:

Def. 1: Self-organized trapezoid: The shape that passengers in front of train doors are normalized is defined as self-organized trapezoid.

Def. 2: Self-organized trapezoid angle: The angle of self-organized trapezoid and perpendicular to the train is called self-organized trapezoid.

The process of boarding and alighting is shown in Fig. 1.

2.3 Notations Summary

In order to make these parameters clear and model equations explicit, we summarize all the parameters in Table 1.

Fig. 1 Example of high-density boarding and alighting

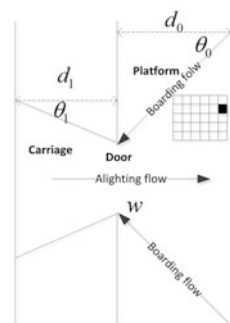


Table 1 Notations summary

Notations	Meanings	Notations	Meanings
Q	Passenger volume, $peo/m^2 \times s$	w	Width of door, m
μ	Rate of boarding, m/s	d_0	Altitude of boarding self-organized trapezoid, m
v	Rate of alighting, m/s	d_1	Altitude of alighting self-organized trapezoid, m
ρ	Density of boarding, peo/m^2	θ_0	Angle of boarding self-organized trapezoid
η	Density of alighting, peo/m^2	θ_1	Angle of alighting self-organized trapezoid
P_0	Aggregation of boarding, peo	$A_1(t)$	Area of boarding self-organized trapezoid, m^2
P_1	Aggregation of alighting, peo	$A_0(t)$	Area of alighting self-organized trapezoid, m^2
k_0	Momentum of boarding, $kg \times m/s$	k_1	Momentum of alighting, $kg \times m/s$
t_0	Time of stopping, s	\bar{m}	Average weight of passengers, kg

3 My Model

3.1 The Continuity Equation

3.1.1 Boarding Continuity Equation

The mini-unit is shown in Fig. 2. In the flow field at the time of t , we take any two-dimensional mini-unit $\tau(x, y)$ in boarding self-organized trapezoid, and the lengths are respectively dx and dy .

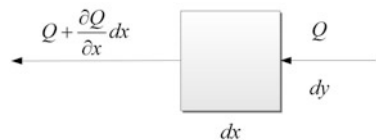
In unit time, passengers inflow volume is Q , and outflow volume is $[Q + \frac{\partial Q}{\partial x} dx] dy$.

So in unit time, the boarding flow volume is (2).

$$\left[Q + \frac{\partial Q}{\partial x} dx \right] dy - Q dy = \frac{\partial Q}{\partial x} dx dy = \frac{\partial(\rho\mu)}{\partial x} dx dy \tag{2}$$

Due to $P_0 = \rho \cdot A_0(t)$, in unit time the loss of passengers is (3).

Fig. 2 Diagram of two-dimensional mini-unit



$$A_0(0) \frac{\partial(\rho\mu)}{\partial x} dx dy + \frac{\partial\rho}{\partial t} \cdot A_0(t) dt + \frac{\partial A_0(t)}{\partial t} \cdot \rho dt = 0 \tag{3}$$

Trapezoidal area formula is (4).

$$A_0(0) = (w + d_0 \tan \theta_0) \cdot d_0 \tag{4}$$

If the speed of passengers is uniform speed, $x = d_0 - \mu t$

$$\frac{\partial A_0(t)}{\partial t} = -\mu w - 2(d_0 - \mu t) \cdot \mu \tan \theta_0 \tag{5}$$

The relationship between dx and dy is (6).

$$dy = 2 \tan \theta_0 dx \tag{6}$$

We bring (2)–(6) into (7), the boarding continuity equation is (7).

$$\begin{aligned} [w + (d_0 - \mu t) \tan \theta_0] (d_0 - \mu t) \cdot \frac{\partial \rho^2}{\partial t^2} + [2d_0 \mu^2 (w + d_0 \tan \theta_0) \cdot \tan \theta_0 \\ - 4\mu (d_0 - \mu t) \tan \theta_0 - 2\mu w] \frac{\partial \rho}{\partial t} + 2\mu^2 \tan \theta_0 \cdot \rho = 0 \end{aligned} \tag{7}$$

In (7), t is the independent variable, and ρ is dependent variable. $t \in (0, t_0)$.

3.1.2 Alighting Continuity Equation

The theory of alighting continuity equation is same as boarding, but in the reverse direction. So the alighting continuity equation is (8).

$$\begin{aligned} [w + (d_1 + vt) \tan \theta_1] (d_1 + vt) \cdot \frac{\partial \eta^2}{\partial t^2} + [2d_1 v^2 (w + d_1 \tan \theta_1) \cdot \tan \theta_1 \\ + 4v (d_1 + vt) \tan \theta_1 + 2vw] \frac{\partial \eta}{\partial t} + 2v^2 \tan \theta_1 \cdot \eta = 0 \end{aligned} \tag{8}$$

3.2 Resultant of Boarding and Alighting

The theorem of momentum is (9).

$$\frac{Dk_x}{Dt} = \sum F_x \tag{9}$$

We take boarding as the example, the resultant of boarding is (10).

$$\begin{aligned}
 dk_0 &= \mu dm = \mu \bar{m} \frac{\partial P_0}{\partial t} dt = \mu \bar{m} \left(\frac{\partial \rho}{\partial t} \cdot A_0(t) + \rho \frac{\partial A_0(t)}{\partial t} \right) \cdot dt \\
 &= \mu \bar{m} \left\{ \frac{\partial \rho}{\partial t} [w + (d_0 - \mu t) \tan \theta_0] (d_0 - \mu t) - \rho (\mu w + 2(d_0 - \mu t) \cdot \mu \tan \theta_0) \right\} dt
 \end{aligned}
 \tag{10}$$

Similarly, the resultant of alighting is (11).

$$dk_1 = v \bar{m} \left\{ \frac{\partial \eta}{\partial t} [w + (d_1 + vt) \tan \theta_1] (d_1 + vt) + \eta (vw + 2(d_1 + vt) \cdot v \tan \theta_1) \right\} dt
 \tag{11}$$

So the resultant of boarding and alighting is (12).

$$\begin{aligned}
 \sum F_y = F_0 - F_1 &= \mu \bar{m} \left\{ \frac{\partial \rho}{\partial t} [w + (d_0 - \mu t) \tan \theta_0] (d_0 - \mu t) - \rho (\mu w + 2(d_0 - \mu t) \cdot \mu \tan \theta_0) \right\} \\
 &\quad - v \bar{m} \left\{ \frac{\partial \eta}{\partial t} [w + (d_1 + vt) \tan \theta_1] (d_1 + vt) + \eta (vw + 2(d_1 + vt) \cdot v \tan \theta_1) \right\}
 \end{aligned}
 \tag{12}$$

4 Simulation

4.1 Experimental Environment

We take an example of Chongwenmen station in Beijing subway Line 1. We investigated multiple sets of data about the number of boarding and alighting at different times for data supporting. Three groups are picked out in Table 2.

At this moment in Table 3, the angle of boarding self-organized trapezoid is at the right angle. The examples are specialized all the angles of self-organized trapezoid with right angle for simplified calculation. Besides, the altitudes of

Table 2 Data of investigation results

Group	Boarding passengers (P_0)	Alighting passengers (P_1)	Altitude of boarding self-organized trapezoid (d_0)	Times of train stop (t_0)
1	20	6	3.5	45
2	10	10	1.9	43
3	4	15	0.6	42

Table 3 Contrast of density and rate

Group	Boarding		Alighting	
	Density ($\rho(0)$, peo/m ²)	Rate (μ , m/s)	Density ($\eta(0)$, peo/m ²)	Rate (v , m/s)
1	4.4	0.38	1.0	-0.90
2	4.0	0.40	3.0	-0.56
3	5.1	0.31	5.0	-0.32

alighting self-organized trapezoid cannot be measured. As the result, they would be assumed as $\eta_1(0) = 1\text{peo/m}^2, \eta_2(0) = 3\text{peo/m}^2, \eta_3(0) = 5\text{peo/m}^2$.

According to the GB 50157-2013 Code for design of metro, the width of train doors in Beijing subway Line 1 is 1.3 m, so $w = 1.3$ m.

In addition, according to Ando et al. [11] about the relationship between density and speed in passenger flow, the rate of boarding and alighting can be calculated based on Table 2. Results are given in Table 3.

4.2 Simulation Results

A new definition named critical density is introduced in images of continuity equations for understandable. When the actual density is less than critical one, the activities of boarding or alighting can be deemed to be finished.

Def. 3: Critical density (ρ_0, η_0): Critical density is the ratio of 1 to the area of self-organized trapezoid at some point. The formulas are (13a, b).

$$\rho_0 = \frac{1}{A_0(t)} = \frac{1}{w(d_0 - \mu t)} \tag{13a}$$

$$\eta_0 = \frac{1}{A_1(t)} = \frac{1}{w(d_1 + vt)} \tag{13b}$$

4.2.1 Simulation of Continuity Equations

(1) Boarding simulation

By MATLAB, the results of boarding are shown in Fig. 3.

In Fig. 3, $p_i, (i = 1, 2, 3)$ are the graphs of boarding continuity equations. With time going by, boarding passengers at the platform are fewer, and the density is lower at the same time. The area of self-organized trapezoid decreases and density

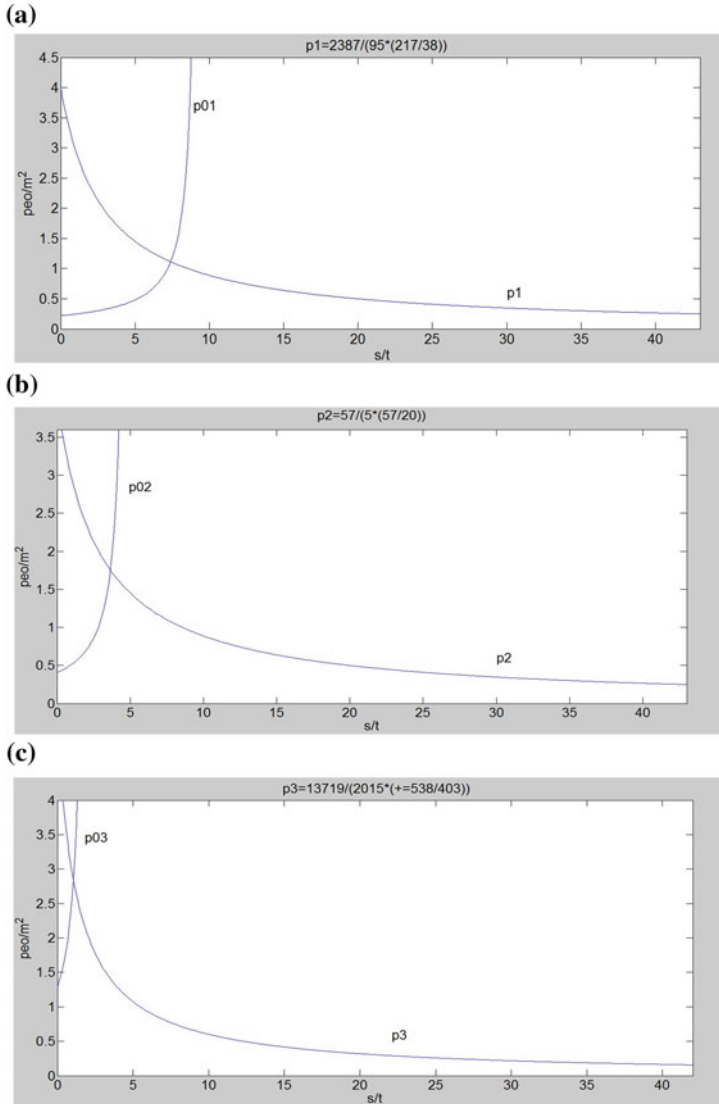


Fig. 3 **a** Boarding continuity equation result of Group 1, **b** Boarding continuity equation result of Group 2, **c** Boarding continuity equation result of Group 3

risers. The abscissa of intersection points represents the time when boarding finishes. These images present that the finish time of boarding are 9 s, 4 s, 2 s, respectively. The boarding time are very short compared with train stop time, which means safety and stabilization.

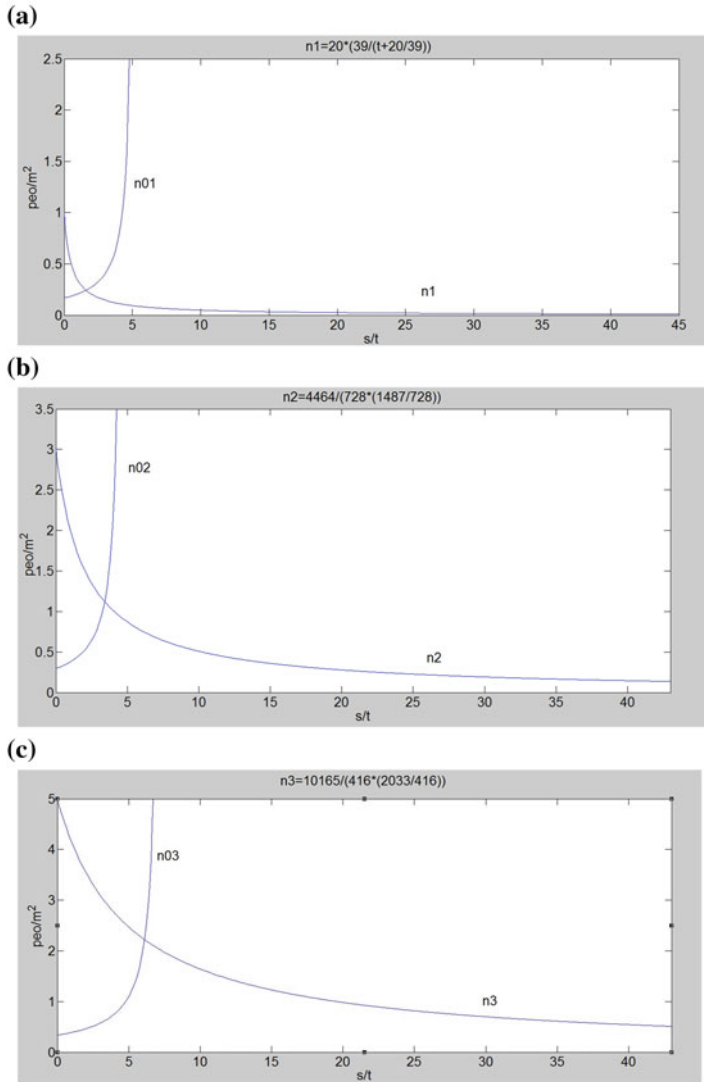


Fig. 4 a Alighting continuity equation result of Group 1, b Alighting continuity equation result of Group 2, c Alighting continuity equation result of Group 3

(2) Alighting simulation

The results of alighting are shown in Fig. 4.

In Fig. 4, n_i , ($i = 1, 2, 3$) are the graphs of boarding continuity equations. n_{0i} , ($i = 1, 2, 3$) is the graph of densities. These images present the finish time of

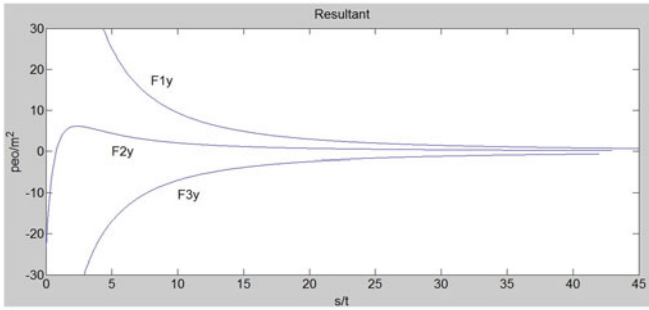


Fig. 5 Resultant formula result

boarding are 2 s, 3 s, 7 s, respectively. The boarding time is very short compared with train stop time, which means safety and stabilization.

4.2.2 Simulation of Resultant

The results of resultant are shown in Fig. 5.

In Fig. 5, with the high-density passengers, the direction of resultant is boarding direction. As boarding passengers at platform decreases, resultant becomes smaller until to zero. On the contrary, when the number of alighting passengers is high, resultant is with alighting direction, decreasing to zero. And when boarding and alighting are almost same, passengers get off and then on, the resultant represents appears to be minus plus, until its end. The graphs cater for practical situation, which are meaningful in actual operation.

5 Conclusion

The paper presents two features of passenger flow, based on which two continuity equations are deduced. We established a formula calculating the resultant during the process of boarding and alighting. The situation of Chongwenmen station in Beijing subway Line 1 has been simulated by MATLAB, which proves the validity of my model.

The paper provides the theory foundation for safer process of boarding and alighting, which can be the theoretical support to the actual operation and plays an important guiding meaning in daily management.

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No. 2016YFB1200402-002).

References

1. Kraft WH (1975) An analysis of the passenger vehicle interface of street transit systems with applications to design optimization. New Jersey Institute of Technology, New Jersey
2. Weston JG (1989) Train service model—technical guide. Underground Operational Research Note, London
3. Wiggendaad IPBL (2001) Alighting and boarding times of passengers at Dutch railway stations. Delft University of Technology, Delft
4. Lam WHK, Cheung CY, Lam CF (1999) A study of crowding effects at the Hong Kong light rail transit stations. *Transp Res Part A Policy Pract* 33(5):401–415
5. Wu J, Ma S (2012) Crowdedness classification method for island platform in metro station. *J Transp Eng* 139(6):612–624
6. Bandini S, Mondini M, Vizzari G (2014) Modelling negative interactions among pedestrians in high density situations. *Transp Res Part C Emerg Technol* 40:251–270
7. Seriani S, Fernandez R (2015) Pedestrian traffic management of boarding and alighting in metro stations. *Transp Res Part C Emerg Technol* 53:76–92
8. Cao Shouhua (2009) Analysis and modeling on passengers traffic characteristics for urban rail transit. University of Beijingjiaotong, Beijing (In Chinese)
9. Li Rifei (2009) Study on passenger flow model at metro platform end based on fluid mechanics. University of Beijingjiaotong, Beijing (In Chinese)
10. Li Juncheng (2015) Analysis and modeling on passengers traffic characteristics for urban rail transit. *East China Sci Technol* 3:33–34
11. Ando K, Ota H, Oki T (1988) Forecasting the flow of people. *Railway Res Rev* 45(8):8–14

A Multi-objective Programming Model for Subway Last Train Timetabling

Wanxiao Xiang and Yanhui Wang

Abstract Subway last train timetabling focuses more on train coordination to make passengers transfer smoothly. Aiming to find an optimal balanced timetabling, a multi-objectives programming model is proposed with speed profile control for finding the optimal train velocity on sections and train dwell time at stations. Our work makes two contributions. First, this paper analyzes train movements in a line and provides an energy flow transferring method. Second, based on the analysis of train velocity and energy flow, an optimization model, in which passenger satisfaction, train trip time and network energy consumption are model objectives, is brought into the last train timetabling. Finally, to obtain an approximately optimal scheduling strategy, an artificial intelligence algorithm is proposed in particular to effectively solve the proposed model. The results of numerical experiments demonstrate the efficiency and effectiveness of the proposed method.

Keywords Subway · Last train timetabling · Multi-objective programming
Train coordination

1 Introduction

Last train is the final choice for passengers to ride or transfer by subway so that the purpose of its timetabling problem tends to the coordination of trains in different lines, which differs from other trains. Nowadays, numerous researchers investigate train timetabling problem (TTP), in which energy consumption and trip time are usually regarded as the optimized objects. However, as a branch in this domain, last train timetabling problem has been paid attention by only a few scholars [1–8].

W. Xiang (✉) · Y. Wang

State Key Laboratory of Rail Traffic Control and Safety, School of Traffic and Transportation,
Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 15120778@bjtu.edu.cn

Y. Wang

e-mail: wangyanhui@bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_80

791

Ning et al. [3] proposed a timetable optimization model to minimize transfer waiting times with the distribution characteristics of transfer passengers. Kang and Zhu [1] provided two practical optimization models for last trains to minimize the standard deviation of transfer redundant times and to balance the last train transfers in subway networks. Kang and Meng [2] developed a global optimization method that can solve the last train departure time choice problem for large-scale urban subway networks. Kang et al. [4] proposed a rescheduling model for last trains with the consideration of train delays caused by incidents that occurred in train operations.

However, several researchers use network energy consumption to be the objective in this issue. Thus, this paper presents a multi-objective programming model for last train to make passengers transfer more smoothly and reduce network energy consumption and trip time of last trains in the network.

2 Problem Statements

Subway network consists of lines, in which there are stations and sections between every two stations. In a network, line set can be represented as $L = \{1, 2, 3, \dots, n\}$, $l \in L$. In each line l , station set can be represented as $S(l) = \{l(1), l(2), l(3), \dots, l(m)\}$, $s \in S(l)$. To transfer station, it can be described as $Tr = \{(l(s), l'(s')), \dots\}$, in which $(l(s), l'(s'))$ represents that station s in line l and station s' in line l' are the same transfer station. Meanwhile, section set can be represented as $Sec(l) = \{l(1, 2), l(2, 3), \dots, l(m-1, m), l(m, m-1), \dots, l(2, 1)\}$, $l(s, s+1), l(s, s-1) \in Sec(l)$, in which $l(s, s+1)$ represents the section between stations s and $s+1$. In addition, d_s^l represents the distance of section $l(s, s+1)$.

2.1 Assumptions

In this section, several assumptions made to formulate the model are explained.

Assumption 1 Train speed profile in each section is divided into four phases, i.e., accelerating phase, uniform phase, coasting phase, and braking phase. Train runs to the allowed maximum velocity v_{\max} with maximum tractive force $f_{a\max}$ in accelerating phase and maximum braking force $f_{b\max}$ in braking phase.

Assumption 2 For simplicity, train mass M , maximum tractive force $f_{a\max}$, maximum braking force $f_{b\max}$, resistance r , and passenger transfer time in a station are considered as constants.

Assumption 3 There is an energy storing device in each station, and regenerated energy of trains will be transferred into the device. Within the useful time t_{es} of the device, the energy can be used to be the tractive energy of the following trains; otherwise, it will be dissipated.

2.2 Train Movement and Energy Consumption

In this paper, we consider that trains run in lines with two directions and the train trip in a line l is shown in Fig. 1. The last up-direction train arrives in station s at time t_{sa}^{+l} , and departs from this station at time t_{sd}^{+l} after the dwell time t_{sdw}^{+l} . It runs on section $l(s, s + 1)$ during the running time t_{sr}^{+l} and then arrives in next station $s + 1$. Moreover, to the last train with down direction at station $s + 1$, it arrives and departs at time $t_{(s+1)a}^{-l}$ and $t_{(s+1)d}^{-l}$. During the running time $t_{(s+1)r}^{-l}$ on section $l(s + 1, s)$, it arrives in station s .

The arrival time and the departure time in Fig. 1 can be computed by two equations Eq. (1). To each line l , Eq. (1) tracks the arrival time and the departure time of the last up-direction train at station s . The departure time t_{sd}^{+l} in the second row of Eq. (1) can be obtained by the dwell time t_{sdw}^{+l} to the arrival time in the first row.

$$\begin{cases} t_{sa}^{+l} = t_{1a}^{+l} + \sum_{i=1}^{s-1} t_{ir}^{+l} + \sum_{j=1}^{s-1} t_{jdw}^{+l} \\ t_{sd}^{+l} = t_{1a}^{+l} + \sum_{i=1}^{s-1} t_{ir}^{+l} + \sum_{j=1}^s t_{jdw}^{+l} \end{cases} \tag{1}$$

Moreover, the running time of the last up-direction train on a section can be obtained by adding up the time of four phases, shown in Fig. 2. In the first phase, the train accelerates with maximum tractive force f_{amax} to the allowed maximum velocity v_{max} during the time t_{sra}^{+l} . It runs with maximum velocity in the second phase and then coasts to velocity v_{src}^{+l} during a period of time t_{src}^{+l} . Within final distance to next station, it brakes from the switching velocity v_{src}^{+l} to 0.

The speed profile in detail which corresponds to Fig. 2 can be formulated by Eq. (2), in which the first row denotes the accelerating phase, the second row denotes the uniform-velocity phase, the third row denotes the coasting phase, and the last row denotes the braking phase. Note that v_{src}^{+l} is a decision variable and the uniform-velocity time t_{src}^{+l} can be determined by v_{src}^{+l} , shown in Eq. (3). Thus, based

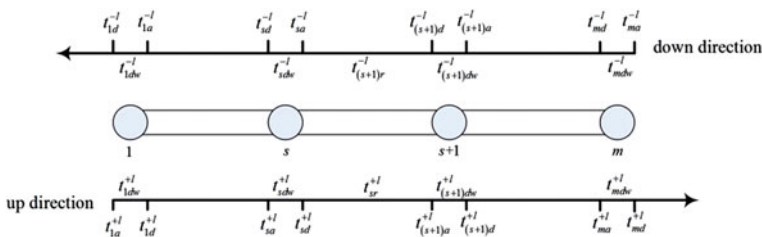
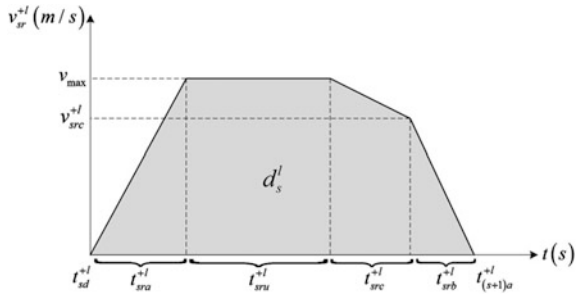


Fig. 1 Train trip in line l

Fig. 2 Train speed profile on section $l(s, s + 1)$



on the equation, the running time of the last up-direction train on section $l(s, s + 1)$ can be formulated by Eq. (4).

$$v_{sr}^{+l} = \begin{cases} (f_{amax} - r)(t - t_{sd}^{+l})/M, & t \in [t_{sd}^{+l}, t_{sd}^{+l} + t_{sra}^{+l}] \\ v_{max}, & t \in (t_{sd}^{+l} + t_{sra}^{+l}, t_{sd}^{+l} + t_{sra}^{+l} + t_{sru}^{+l}) \\ v_{max} - r(t - t_{sd}^{+l} - t_{sra}^{+l} - t_{sru}^{+l})/M, & t \in [t_{sd}^{+l} + t_{sra}^{+l} + t_{sru}^{+l}, t_{sd}^{+l} + t_{sra}^{+l} + t_{sru}^{+l} + t_{src}^{+l}] \\ v_{src}^{+l} - (f_{bmax} + r)(t - t_{sd}^{+l} - t_{sra}^{+l} - t_{sru}^{+l} - t_{src}^{+l})/M, & t \in (t_{sd}^{+l} + t_{sra}^{+l} + t_{sru}^{+l} + t_{src}^{+l}, t_{(s+1)a}^{+l}] \end{cases} \quad (2)$$

$$t_{sru}^{+l} = (v_{src}^{+l})^2 f_{bmax} M / 2v_{max} r (f_{bmax} + r) + d_s^l / v_{max} - v_{max} f_{amax} M / 2r (f_{amax} - r) \quad (3)$$

$$t_{src}^{+l} = (v_{src}^{+l})^2 f_{bmax} M / 2v_{max} r (f_{bmax} + r) - v_{src}^{+l} f_{bmax} M / r (f_{bmax} + r) + d_s^l / v_{max} + v_{max} f_{amax} M / 2r (f_{amax} - r) \quad (4)$$

In addition, in the former two phases, the last up-direction train on section $l(s, s + 1)$ requires tractive energy, given by the first row of Eq. (5). And the second row gives the regenerated energy in braking phase, in which t_{srb}^{+l} represents the braking time of the last up-direction train on section $l(s, s + 1)$.

$$\begin{cases} E_{st}^{+l} = \int_{t_{sd}^{+l}}^{t_{sd}^{+l} + t_{sra}^{+l}} f_{amax} v_{sr}^{+l} dt + \int_{t_{sd}^{+l} + t_{sra}^{+l}}^{t_{sd}^{+l} + t_{sra}^{+l} + t_{sru}^{+l}} r v_{sr}^{+l} dt \\ E_{sb}^{+l} = \int_{t_{(s+1)a}^{+l}}^{t_{(s+1)a}^{+l} - t_{srb}^{+l}} f_{bmax} v_{sr}^{+l} dt \end{cases} \quad (5)$$

3 The Presented Multi-Objective Programming Model

In this section, the presented multi-objective programming model is formulated in detail.

3.1 Model Constraints

In this section, three constraints are provided in detail.

(a) Train velocity constraint

In this model, the velocities v_{sr}^{+l}, v_{sr}^{-l} of running trains should not exceed the maximum velocity v_{max} . To the switching velocity from coasting phase to braking phase, it has been selected in a smaller interval on account of the uniform-velocity time t_{stru}^{+l} which is not allowed to be less than 0, shown in Eq. (6).

$$v_{src}^{+l}, v_{src}^{-l} \in \left[\sqrt{\max\left(0, \frac{(f_{bmax} + r)(v_{max}^2 M f_{amax} - 2d_s^l r(f_{amax} - r))}{M f_{bmax}(f_{amax} - r)}\right)}, v_{max} \right] \quad (6)$$

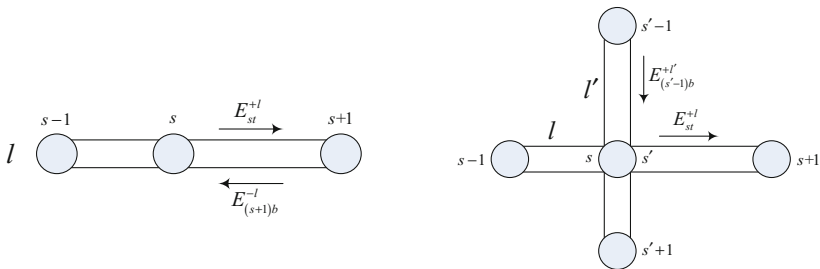
(b) Dwell time constraint

Within the dwell time, passengers get on and off trains. There are minimum and maximum dwell times of trains at a station. To each train and each station, these two dwell times are fixed.

$$t_{dwmin} \leq t_{sdw}^{+l}, t_{sdw}^{-l} \leq t_{dwmax} \quad (7)$$

(c) Useful regenerated energy constraint

Due to the assumption 3, regenerated energy transferred into the device will be dissipated except that another train departs from the same station within the useful time of the device. So we should judge if there is useful regenerated energy or not before the computation of network energy consumption. To station s without transfer channel, Eq. (8) provides the useful regenerated energy of the last down-direction train in line l . And Eq. (9) corresponds to the last up-direction train at transfer station s' in line l . Besides, the energy flows in two stations are shown in Fig. 3.



(a) energy flow in station without transfer channel (b) energy flow in transfer station

Fig. 3 Energy flows in different stations

$$E_{(s+1)ub}^{-l} = \begin{cases} \min(E_{st}^{+l}, E_{(s+1)b}^{-l}), & t_{sd}^{+l} \in (t_{sa}^{-l}, t_{sa}^{-l} + t_{es}) \\ 0, & \text{others} \end{cases} \tag{8}$$

$$E_{(s'-1)ub}^{+l'} = \begin{cases} \min(E_{st}^{+l}, E_{(s'-1)b}^{+l'}), & t_{sd}^{+l} \in (t_{s'a}^{+l'}, t_{s'a}^{+l'} + t_{es}) \\ 0, & \text{others} \end{cases} \tag{9}$$

3.2 Model Objectives

Three model objectives are introduced in this section.

(a) Passenger satisfaction

In this paper, we regard passenger satisfaction PS as one of the model objectives. It can be described by two satisfactions of passengers waiting on trains and wanting to transfer. For passengers waiting on trains, there is an acceptable waiting time t_w . Equation (10) shows the satisfaction of passengers on the last up-direction train at station s in line l . Equation (11) represents the satisfaction of passengers who transfer from station s in up-direction line l to station s' in up-direction line l' , in which $t_{ls}^{l's'}$ represents passenger transfer time in transfer station $(l(s), l'(s'))$. Thus, passenger satisfaction PS can be obtained by Eq. (12).

$$psw_{+ls} = \begin{cases} 1, & 0 < t_{sdw}^{+l} \leq t_w \\ 0, & t_{sdw}^{+l} > t_w \end{cases} \tag{10}$$

$$pst_{+ls}^{+l's'} = \begin{cases} 0, & t_{sa}^{+l} + t_{ls}^{l's'} \geq t_{s'd}^{+l'} \\ 1, & t_{sa}^{+l} + t_{ls}^{l's'} < t_{s'd}^{+l'} \end{cases} \tag{11}$$

$$PS = \sum_{l=1}^n \sum_{s=1}^m (psw_{+ls} + psw_{-ls}) + \sum_{(l(s), l'(s')) \in Tr} (pst_{+ls}^{+l's'} + pst_{-ls}^{+l's'} + pst_{+ls}^{-l's'} + pst_{-ls}^{-l's'}) \tag{12}$$

(b) Train trip time and network energy consumption

In train timetabling domain, train trip time and network energy consumption are usually regarded as the targets which need to be optimized. Thus, trip time T of last trains in a network and network energy consumption E are used to be the remaining two objectives in this model. The previous one is determined by the arrival time of last trains at the first station and the departure time of last trains from the last station, shown in Eq. (13). And network energy consumption, which is denoted by tractive energy and useful regenerated energy, is given by Eq. (14).

$$T = \max_{l \in L} (t_{md}^{+l}, t_{1d}^{-l}) - \min_{l \in L} (t_{1a}^{+l}, t_{ma}^{-l}) \tag{13}$$

$$E = E_t - E_{ub} = \sum_{l=1}^n \left(\sum_{s=1}^{m-1} E_{st}^{+l} + \sum_{s=2}^m E_{st}^{-l} \right) - \sum_{l=1}^n \left(\sum_{s=1}^{m-1} E_{sub}^{+l} + \sum_{s=2}^m E_{sub}^{-l} \right) \tag{14}$$

3.3 The Presented Model

All the constraints and objective functions are introduced in the above. Thus, the presented multi-objective programming model can be formulated as shown in Eq. (15), in which α , β and γ are the coefficients to show the importance of objectives and balance the magnitude.

$$\begin{aligned} \min \quad & \alpha E + \beta T - \gamma PS \\ \text{s.t.} \quad & v_{sr}^{+l}, v_{sr}^{-l} \in [0, v_{\max}] \\ & \dots \dots \\ & t_{dwm} \leq t_{sdw}^{+l}, t_{sdw}^{-l} \leq t_{dwm} \\ & E_{(s+1)ub}^{-l} = \begin{cases} \min(E_{st}^{+l}, E_{(s+1)b}^{-l}), & t_{sd}^{+l} \in (t_{sa}^{-l}, t_{sa}^{-l} + t_{es}) \\ 0, & \text{others} \end{cases} \\ & \dots \dots \end{aligned} \tag{15}$$

4 Case Study

In this section, a case study based on a test network, which contains three lines and eight stations, is presented to illustrate the practicability of the proposed model, shown in Fig. 4.

Fig. 4 A subway network for testing

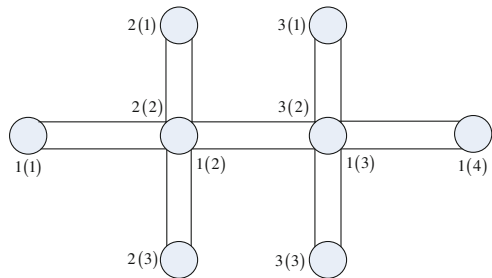


Table 1 Values and units of rest parameters

Parameter	M	f_{amax}	f_{bmax}	r	t_{es}	t_w	α	β	γ
Value	195	315	258	3	600	40	0.000001	1	20
Unit	t	kN	kN	kN	s	s	–	–	–

In this test, the distance of each section is 2.5 km. The last train runs on the section with velocity no larger than 20 m/s, of which the dwell time at each station is donated in an interval [20, 60]. Their departure times are seen as 0, 30, 75, 20, 10, and 40. In particular, 0 and 30 are corresponding to the last trains in line 1 with up and down directions. Besides, in transfer stations (1(2), 2(2)) and (1(3), 3(2)), passenger transfer times are 180 and 240. And rest parameters are initialized in Table 1.

In this paper, as a great algorithm searching for solutions efficiently, GA (see Kang [4]) is employed to solve the presented model. We use MATLAB to carry out the algorithm, in which the sizes of the generation and the populations in one generation are both 50, and the probability of mutation operation is 0.5.

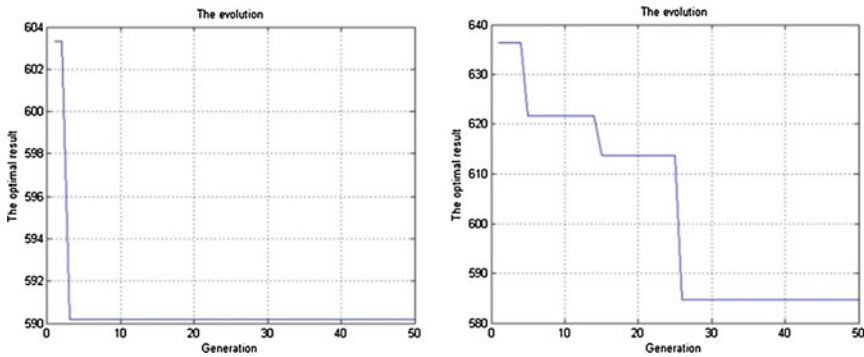
After one test, we obtain some information. In this test, the optimal result is 582.0212, and values of decision variables are shown in Table 2. In this table, every two rows correspond to a line with different directions. Meanwhile, the values in odd columns represent train dwell time at stations, and the switching velocities from uniform-velocity phase to coasting phase on sections are given by even columns. For example, in the first two rows, the preceding row is related to line 1 with up direction, and another one is correlated with line 1 with down direction. In this preceding row, The values in 1, 3, 5, 7 columns are the dwell times of the last up-direction train at stations 1, 2, 3, 4 in line 1, and those in 2, 4, 6 columns are the switching velocities from uniform-velocity phase to coasting phase on Sections 1(1, 2), 1(2, 3), 1(3, 4). In addition, due to only three stations in line 2 and 3, the values from the third row to the sixth row in final two columns are 0.

In Table 2, we found all the switching velocities are concentrated in an interval [18, 20], which is caused by the nonnegative uniform time of last trains on sections. To make passengers transfer more smoothly, we can control running time on sections and dwell time at stations. Due to the small selectable range of these switching velocities, train running times on sections are also limited in a small interval. Thus, there is only one way to select, dwelling more time. It may be the reason of the large dwell time at stations without transfer channel. In addition, it can be seen that the dwell times of all the last trains at their determinations are 20 s, which is the minimum dwell time. It is because they do not need to be coordinated with others.

To explore the accuracy of the algorithm, 20 tests are made and there are two kinds of computational performances shown in Fig. 5. In 16 tests, there is a convergence directly within five iterations, shown in Fig. 5a. In others, through some local optimums, the optimal results can also be found within 30 iterations, given by Fig. 5b. The computational times of all the tests do not exceed 52 s. And the

Table 2 Values of decision variables

Dwell time	Switching velocity	Dwell time	Switching velocity	Dwell time	Switching velocity	Dwell time
41.4	18.3	44.1	19	29.8	18.7	20
20	18.5	20.5	18.8	42.6	18.6	31.2
46.1	19.9	53.4	19.5	20	0	0
20	18.8	34.1	18.7	58.5	0	0
41.3	18.4	46.1	19.5	20	0	0
20	19.9	34.7	20	56.6	0	0



(a) the case getting the convergen (b) the case getting theconvergence through some ce directly local optimums

Fig. 5 The computational performance

obtained optimal results are listed in an interval [575.7997, 595.5322]. The length of this interval is no larger than 20 while the minimum result is 575.7997. It makes the results acceptable.

5 Conclusions

In this paper, we proposed a multi-objectives programming model to optimize the timetabling of last trains. As a basis, train speed profile control and transferred energy flow are discussed to determined train movements in a line. Under the consideration of train movements, this paper provides three model objectives, passenger satisfaction, train trip time, and network energy consumption, and some constraints, train velocity, dwell time, and transferred energy to formulate the optimization model and get the optimal timetable. Finally, we employ GA to practice the model with a test network on account of the lack of real raw data.

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No. 2016YFB1200402-002).

References

1. Kang L, Zhu X (2017) Strategic timetable scheduling for last trains in urban railway transit networks. *Appl Math Model* 45:209–225
2. Kang L, Meng Q (2017) Two-phase decomposition method for the last train departure time choice in subway networks. *Transportation Research Part B: Methodological* 000:1–15
3. Liqiao N, Peng Z, Wen-kai XU et al (2016) Timetable synchronization optimization for urban rail transit in the last trains’ hour. *J Transp Syst Eng Inf Technol* 16(6):108–114 (in Chinese)
4. Kang L, Wu J, Sun H et al (2015) A practical model for last train rescheduling with train delay in urban railway transit networks. *Omega* 50:29–42
5. Kang L, Wu J, Sun H et al (2015) A case study on the coordination of last trains for the Beijing subway network. *Transp Res Part B Method* 72:112–127
6. Feng Z, Jungang S, Hanchuan P (2013) Optimization method for last train coordination plan of urban rail transit based on network operation. *Procedia Soc Behav Sci* 96:2706–2712
7. Xu R, Li X (2012) Comprehensive optimization for connection scheme of last trains in urban mass transit network. *J Tongji Uni (Nat Sci)* 40(10):1510–1516 (in Chinese)
8. Xu R, Zhang M, Jiang Z et al (2008) Study on departure time domain of the first and last trains of urban mass transit network based on operation coordination. *J China Railway Soc* 30(2):7–11 (in Chinese)

The Research of the Synergic Passenger Flow Control of Urban Rail Transit Line Operation

Yiru Cui and Yanhui Wang

Abstract As one of the main means of public transportation, the urban rail transit plays more and more important role in passenger transport. Under the condition of limited network capacity, the realization of safe operation has become an important topic of urban rail transit. Based on the passenger flow density which intuitively reflects the traffic aggregation degree and the line collaborative operation theory, in this paper, a collaborative safety operation optimization method for urban rail transit lines based on the passenger flow is proposed. Which provides the theoretical basis for the relief of the operation pressure, the improvement of the operational efficiency, and the ensuring of the safety in daily operation of the urban rail transit.

Keywords Urban rail transit · Operation management · Synergic control
Passenger inflow control

1 Introduction

As a kind of large capacity, high-intensity traffic, urban rail transportation plays a very important role in city public transportation, makes a great contribution to relieve city traffic congestion and pollution, provides a powerful guarantee for the residents and the city economic development, and has become one of the important ways to improve the structure of the city public transportation public transportation [1] and city residents commuting rely on the most. With the development of the city, the contradiction between the rapid growth of city rail transit passenger

Y. Cui · Y. Wang (✉)

State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: wangyanhui@bjtu.edu.cn

Y. Cui · Y. Wang

School of Traffic and Transportation, Beijing Jiaotong University,
Beijing 100044, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_81

801

demands and improves the speed of the limited transport capacity increasing, especially in the morning and evening peak hours and weekends, line traffic congestion phenomenon is very serious. In the network operation stage, operation management tends to be fine, this change brings new opportunities for China’s rail transportation industry but also puts forward new and higher requirements for operation and safety assurance [2].

In the face of huge traffic pressure, city rail transit should first consider the optimization and the shorting of the train headway, the optimization of train formation plan to improve transport capacity, alleviate the conflict of large passenger demand and the limited transport capacity [3–5]. Passenger flow limit is a kind of short-term response measures to deal with large passenger flow situation. The appropriate flow limiting measures can effectively alleviate the transportation pressure of the network and the service pressure of the station, and reduce the risk of accidents and improve the service level of the stations [6]. However, the index is still put forward for passenger safety guarantee of a single station. It cannot guarantee the overall efficiency and safety of the road network, and cannot form a cooperative relationship with the control measures of other stations on the line.

In view of the existing literatures listed above, there is no comprehensive consideration of the cooperative relationship in the network operation environment. Therefore, the cooperative control model and control method of passenger flow management are proposed, it is of great significance to improve the efficiency of train transportation organization, maintain operation safety, and guide the operation and management units to formulate a reasonable organization plan.

2 Operation Model of Urban Rail Transit Line

Considering the coordination between the train, station, and passengers, taking the most typical route model as an example, the cooperative operation model of urban rail transit line is established in Fig. 1. The line consists of m stations, two vehicles located at the first and last stations. The operation of vehicle online includes forward and reverse downlink.

2.1 Traffic Model

The train line on the run is mainly used to complete the displacement in the train between stations, so we only consider the three elements of cooperative operation of

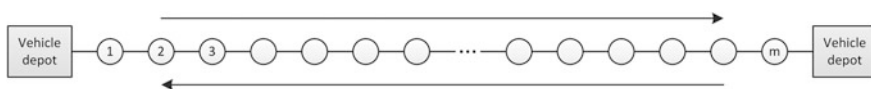


Fig. 1 Cooperative operation model of urban rail transit

urban rail lines, for trains, platforms, people, and the coordination relation can be expressed by train organization equation and passenger flow load and transfer equation.

2.1.1 Train Organization Equation

The working state of the train on the line can be divided into two kinds: one is the running state between the stations; another is the stopping state of the stations. Line stations working state can be divided into two types: First, the train leaves and the following train has yet to reach the vacant state, passengers continue to arrive the platform station in this state. Second, the platform is the occupied because of the train stops, there is not only the arrival or transfer passengers in the platform, but also the passenger flow exchange between the train and the platform in this state. Using matrix $X = [x_{ij}]_{m \times n}$ to represent the states and correlations of the above two, i indicates the train number ($1 < i < m$), and the j indicates the platform number ($1 < j < n$), x_{ij} indicates the time at which the train i arrives at the platform j . According to the law of traffic organization of urban rail transit, the following train organization equations can be established:

$$X = [x_{ij}]_{m \times n}, D = [d_{ij}]_{m \times (n-1)}, R = [r_{ij}]_{m \times (n-1)}, H = [h_{ij}]_{(m-1) \times n} \tag{1}$$

$$L = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}_{(m-1) \times m}, U = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 1 & -1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}_{n \times (n-1)} \tag{2}$$

$$XU = H \tag{3}$$

$$LX = D + R, \tag{4}$$

where d_{ij} refers to the stopping time of the train i at the platform j ; r_{ij} refers to the running time between the train i at the station j and the station $j + 1$; $x_{i+1,j} - x_{ij} - d_{ij}$ refers to the vacant time of platform j before the train $i + 1$ arrives and after the train i leaves; h_{ij} refers to the time headway at the platform j before the train $i + 1$ arrives and after the train i leaves.

2.1.2 Load and Transfer Equation of Passenger Flow

The working time of the platform is divided into several stopping cycles by the time headway, and the basic change process of passenger flow in each stop cycle is regular. During a long stop period, the load and transfer of passenger flow in the platform and carriage can be shown as shown in Fig. 2.

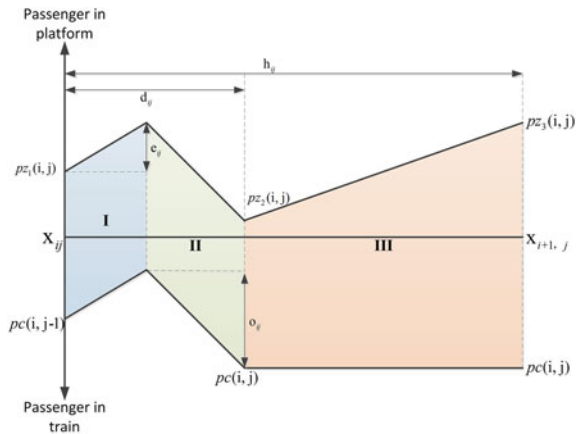
The horizontal axis of Fig. 2 refers to a long stop cycle, two directions of the longitudinal axis refer to the passenger flow in platform and in train. The two curves represent the passenger flow changing situation in the time of the stations and platforms. A cycle can be divided into three parts, Part I is the time for passenger to get off the train, in this period, the passenger is transferred to the platform from the train. Part II is the time for passenger to get on the train, in this period, the passenger is transferred to the train from the platform. Part III is the time for passenger flow accumulation after the train leaves the station, in this period, passenger flow in the platform continues to increase, while the passenger in the train will not change until the train arrives at the next station. In the diagram, we set up the passenger flow arrival speed a_j of the station j is a constant, but actually it is a time-varying variable $a_j = a_j(t)$. According to Fig. 2, passenger load and transfer equations are as Eqs. (5), (6) and (7):

$$pz_1(i, j) = pz_3(i - 1, j) \tag{5}$$

$$pz_2(i, j) = \int_{x_{ij}}^{x_{ij} + d_{i,j}} a_j(t) dt + pz_1(i, j) + e_{ij} - o_{ij} \tag{6}$$

$$pz_3(i, j) = \int_{x_{ij} + d_{i,j}}^{x_{i+1,j}} a_j(t) dt + pz_2(i, j) \tag{7}$$

Fig. 2 Load and transfer of passenger flow within a stop cycle



$$e_{ij} = \sum_{k=1}^{k < j} y_i(k, j) \tag{8}$$

$$o_{ij} = \sum_{l=j}^n y_i(j, l) \tag{9}$$

$$c_{ij} = e_{ij} + o_{ij} \tag{10}$$

$$pc(i, j) = \sum_{k=1}^j \sum_{l > j}^n y_i(k, l), \tag{11}$$

where $pz_1(i, j)$ refers to the platform passenger flows when the train i arrives at platform j ; $pz_2(i, j)$ refers to platform passenger flow when train i leaves at platform j ; $pz_3(i, j)$ refers to the platform passenger flow at the end of the stop cycle of train i at platform j ; $y_i(k, l)$ refers to passenger flow in the train i that get on at station k and get off at station l ; e_{ij} refers to passenger flow in the train i that get off at station j ; o_{ij} refers to passenger flow in the train i that get on at station j ; c_{ij} refers to exchange passenger flow between train i and station j in the stopping time $d_{i,j}$; $pc(i, j)$ refers to passenger flow in the train i when it leaves station j .

2.2 Line Operation Analysis

Figure 2 shows that we can select the density as a measure of operating platform security index. For the difference of passenger density in each waiting area on the platform, the uneven spatial distribution of passenger flow in the waiting area can be expressed by the unbalanced spatial distribution coefficient $uz_{i,j}$. The Eqs. (12) and (13) formula can be used to describe the distribution of passenger flow density on the platform:

$$\rho z_{i,j} = pz_1(i, j) + e_{ij} / sz_j \tag{12}$$

$$uz_{i,j} = \max \{ \rho z_{i,j}^1, \rho z_{i,j}^2, \dots, \rho z_{i,j}^k \} / \rho z_{i,j} \tag{13}$$

The passenger density is the most important index to restrict the safety of the waiting area. In the daily operating, if the passenger density in the effective waiting area is more than $4 P/m^2$, the platform is in a state of insecurity. So the overall operation safety indicators are as Eq. (14):

$$S_1 = \max_{i,j} (4 - \rho z_{i,j} * u_{i,j}) \tag{14}$$

For city rail transport operations, efficiency refers to the comprehensive benefits including passenger flow, economic, social and environmental aspects, objectives that can be reached within a period [7, 8]. The objective of the line operation coordination model, that is, the daily operation management purpose of urban rail transit system, is to maximize the line operation efficiency under the premise of ensuring the operation safety.

In the statistics of the carrying capacity of the line, the number of passengers on all the stations in the control time shall be counted, that is, the carrying capacity of the urban rail transit line in the target period:

$$S_2 = \sum_{T_1 < x_{ij} < T_2} o_{ij} \tag{15}$$

3 Cooperative Operation Control Model of Urban Rail Transit Line

In order to improve the availability and efficiency of the model, the following assumptions and settings are made for the passenger flow inside the platform.

- (1) The passenger demand control period and the outbound traffic are known, the specific data can be obtained from historical data analysis and forecast.
- (2) The departure intervals are identical within the same control period, and the running time and stopping time between stations have been determined. In general, no delay or unexpected situation is considered.
- (3) The passenger flow out of the station will be able to leave the station quickly, and will not be stranded on the platform for too long.

3.1 Objective Function

According to the analysis in 1.2.2, the control model has two objective functions. The first objective function is to calculate the passenger density of the stations within the line, and to control the safety of the largest passenger density platform in the most secure range. The second objective function is to calculate the carrying capacity of the line operation during the target control period.

$$\begin{cases} \max S_1 = \max_{i,j} (4 - \rho z_{i,j} * u_{i,j}) \\ \max S_2 = \sum_{T_1 < x_{ij} < T_2} o_{ij} \end{cases} \tag{16}$$

3.2 Constraint Condition

3.2.1 Interval Capacity Constraint

The most important cause of passenger flow retention and congestion is the finite capacity caused by the limited train capacity, and it is also an important constraint condition of passenger flow control model. The interval carrying capacity refers to the number of passengers that can be transported per unit time under the condition of certain vehicle type, fixed equipment and driving organization method. In general, the operating standard stipulates that the overload rate of the train under running condition shall not exceed 150%, so the restriction of the capacity of the interval transportation can be expressed in the lower form:

$$\frac{\forall i \in [1, m], \forall j \in [1, n], \sum_{k=1}^j \sum_{l>j}^n y_i(k, l)}{\text{Cap}} \leq 150\% \tag{17}$$

3.2.2 Platform Capacity Constraint

In order to ensure the normal operation of all parts of the station, the smooth flow of operation and the safety of operation, the passenger flow inside the station must be controlled within the capacity limit of the station. This paper mainly considers the passenger demand redistribution at each station, so we regard station as the nodes of the network. Therefore, the platform capacity constraint is expressed as platform passenger flow must not exceed its maximum capacity and can be expressed by the lower expression:

$$\forall i \in [1, m], \forall j \in [1, n], pz_1(i, j) + e_{ij} \leq 4 * sz_j \tag{18}$$

3.2.3 Train Safety Constraint

The safety of train operation is mainly controlled by the train headway, and the interval between trains is the time interval between two adjacent trains moving from the garage to the same direction. The dynamic changes of running intervals must comply with the safety requirements, which means to keep a safe distance between before and after the two train operation organization, so the security constraints can be expressed by the following train:

$$\forall i \in [1, m], \forall j \in [1, n], h_{ij} \geq t_s \tag{19}$$

$$\forall i \in [1, m], \forall j \in [1, n], \check{d}_{ij} \leq d_{ij} \leq \hat{d}_{ij} \tag{20}$$

$$\forall i \in [1, m], \forall j \in [1, n], \check{r}_{ij} \leq r_{ij} \leq \hat{r}_{ij} \tag{21}$$

4 A Case Study

Urban rail transit takes on more passenger pressure than other modes of public transportation in Beijing. Therefore, it is of great practical significance to analyze the cooperative control of urban rail transit line operation in Beijing.

4.1 Analysis of Passenger Flow and Operation Organization of Line 1

Through field investigation and data for each station inbound and outbound passenger flow diversion ratio and the downstream station downstream station waiting area is determined, and the transfer station transfer inflow and outflow coefficient; other related parameters specific values shown in Table 1.

4.2 Cooperative Control Analysis of Line 1 Operation

From the point of view of mathematics, model can finally be transfer into double targets linear programming problem, and can be transfer into a single objective linear programming problem using the ideal point method, and MATLAB

Table 1 Model parameter value

Parameter	Value	Explain
Departure interval/s	120	Use the shortest distance of Line 1
Train staff/p	1424	Standard load of train (B type, six section formation)
Maximum full load rate/%	150	The maximum proportion of passenger flow and train personnel during the peak period
Station carrying capacity/(p/h)	24,000	Design capacity of a station
Minimum limiting rate/%	125	Percentage of passenger demand for passenger traffic under maximum current limiting intensity

programming can be used to calculate this problem. To ensure accurate results, the passenger demand data of the target control period before participating in the operational period (6:20–7:00) and the target time after two control period (9:00–9:40) will participate in the calculation.

As shown in Fig. 3a, from the site distribution, limiting station is mainly located in the upstream line, six limiting site limiting strength is the biggest in Pingguoyuan station, whose limiting the average ratio reached 40%, passenger flow organization pressure. As shown in Fig. 3b, the important period in which limiting rate less than 80% is estimated at 7:20–8:40.

The chart Fig. 4a shows the average platform passenger density for each station in the control period, it means the model proposed brings a good result that the average passenger flow density can be controlled below 3 P/m², which is relatively safe for the operation.

In this section, we set the control time units length to 15 min as a control period, while the other parameters remain unchanged, and continue to calculate the

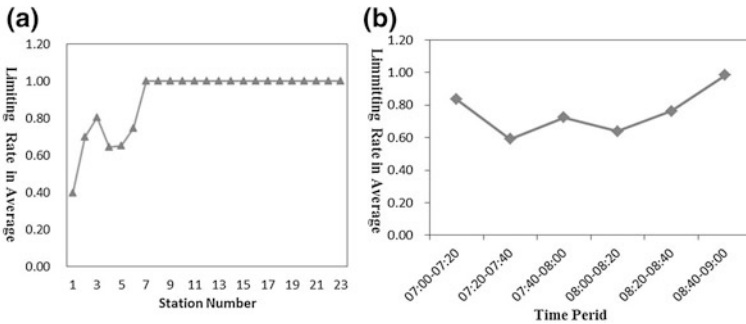


Fig. 3 Limiting rate in average

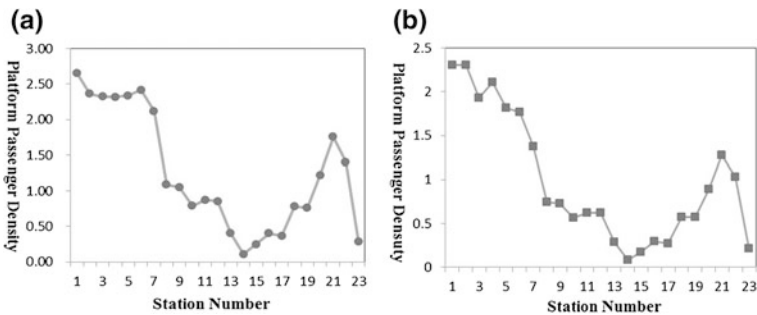


Fig. 4 Platform passenger density

passenger flow control strategy of the Line 1 in the early peak time (7:00–9:00). As shown in Fig. 4b, the platform passenger density is less than 3 P/m^2 , which belong to the safe operation range. The platform passenger density characteristic is basically the same as the result which is calculated with period of 20 min, it shows that the model is scientific. When the unit control time duration is shortened, the description of the traffic demand, characterization of passenger flow distribution, and dynamic characteristics of the traffic source data are more accurate of the system; on the other hand, the model result is more accurate, because of more fine details, the calculation results better.

5 Conclusion

Safety is the prerequisite and core competitive power of rail transportation operation. As the most sustainable urban transportation mode, urban rail transportation plays an irreplaceable supporting role in China's economic and social development, the improvement of people's livelihood and social security.

In this paper, a passenger flow coordination model which is designed to ensure the safety and efficiency of metro system operation is proposed. The purpose is to provide a theoretical method to alleviate the contradiction between passenger demand and line transportation capacity based on the cooperative relationship among train, station, and passenger flow. It is illustrated in the study case that the model proposed in this paper can provide some impactful guidance for the coordinated passenger flow control in the actual operation management.

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No.2016YFB1200402-002).

References

1. Xiao X (2014) Risk and safety assessment of urban rail transit networks operation—a topology based approach. Beijing Jiaotong University. (in Chinese)
2. Yao X, Zhao P, Qiao K (2012) Modeling dynamic passenger flow distribution of urban rail transit network. *J Transp Syst Eng Inf Technol* 12(6):52–59 (in Chinese)
3. Assis WO, Milani BEA (2004) Generation of optimal schedules for metro lines using model predictive control. *Automatica* 40(8):1397–1404
4. Castelli L, Pesenti R, Ukovich W (2004) Scheduling multimodal transportation systems. *Proc Zool Soc London* 119(3):603–615
5. Mignone A, Accadia G (2010) operations research models for programming support of cadenced timetables. *Ingegneria Ferroviaria*
6. Pan Y (2011) Study on optimization of train formation in urban rail transit. Beijing Jiaotong University. (in Chinese)

7. Li W (2009) Study on operation efficiency of transportation infrastructure in Jiangsu. Nanjing University of Science and Technology. (in Chinese)
8. Yao W, Zheng Y (2011) Measures to improve urban rail transit efficiency. Urban Plann Forum 5:73–79 (in Chinese)

Application of Standardization Training of Operational Services in Rail Transit Enterprises

Bin Xu and Yiming Shao

Abstract The quality of operational service can improve the market competitiveness of urban rail transit operators and improve the training level of rail transit service, which is conducive to the improvement of service level and the decrease of operating accident rate. In order to ensure the quality of operation and service training, promoting the construction of standardized training services is the only way. This paper analyzes the characteristics of standardized training and construction of operational service, puts forward the way of standardization construction, and analyzes the aspects of service management and other aspects, aiming at the standardization training of urban rail transit service and provides some ideas and reference.

Keywords Rail · Format · Operation services · Service standardization Training

1 Introduction

The rapid development of urban rail transit not only brings the promotion of operational services but also boosts the standardization of operating training system as well as standardization of urban rail transit operations training. Public service standards that: the provision of public services in the process of the use of standardized principles and the establishment and use of public service standards, to make the quality of public services, service methods standardized, the process of service procedures, which can get more high-quality public service [1]. Urban rail transit operation services belong to the public service category, the service standardization of the ultimate goal is to meet the needs of public travel under the premise of the rail transit business itself through the behavior and system, to pro-

B. Xu (✉) · Y. Shao

Chongqing City Construction Technician School, Shuangfu Avenue
on the 9th, Shuangfu New District, Jiangjin, Chongqing, China
e-mail: 35330505@qq.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_82

813

vide the best travel service. At the same time, modern management [2] pointed out that the standardization of management is the basis of fine management and personalized management, the rail transit operations to achieve refinement, first of all to achieve standardization.

In October 2013, China promulgated the national “Urban Rail Transit Operation and Management Code” for the first time. Although the standard has been going on for a long time, since the construction of urban rail transit in China, the urban rail transit operators have been committed to operating services standardized construction. Such as the Beijing Metro [3], Shanghai Rail Transit [4] has opened the implementation of standardized operations as a pioneer in the standardization of operational services within the industry, but also for the national industry standards and other local industries. The standard formulation provides valuable reference experience.

2 Standardized Training and Construction of Rail Transit Operation Service

2.1 Analysis on the Characteristics of Service Standard Training

Content coverage is wide: Operation training is comprehensive, involving driving, passenger, ticketing, staff organizations, and other aspects, and more for the job, for the types of jobs and jobs need to develop appropriate training or technical training standards. Therefore, all the requirements in standardization process should be met.

Technical standards are highly required since operation service is the fundamental of the entire rail transit service, the quality of its service quality is not only related to the efficiency of the operation of enterprises but also the safety of passengers security is essential. This makes the technical requirements for the operation of the relevant services more stringent, for the job content to be refined training technical points to ensure that key training links improve the training procedures and methods and supervision system.

Establish people’s consciousness: The fundamental object of the operation service is the passenger, for the direct contact with the passengers of the post, but also needs to pay attention to the feelings of passengers were served. Apart from strengthening the sense of service, normalization and standardization of staff behaviors are also stressed for improving customers’ consolation.

Attention to operability [5]: For the specification and assessment of operational service training processes or training quality, metrics can be used. In order to ensure the enforceability of standard training, it is necessary to select the indicators according to the actual work content, grasp the measurement reference value, and explain the key indicators in the standard documents, so as to ensure the operability of the training.

2.2 The Way to Improve the Construction

Safeguarding the standardization of construction Traffic mainly from three aspects to improve the construction approach.

2.2.1 Working Mechanism

Standard training system involved in administration, operation and supervision is established to standardize the operation service and improve service quality [6], involving management, operation, and supervision. The operation branch of the rail transit group Co., Ltd. is responsible for the publication of standardized training. Its subordinate committee has established the standardization committee as the main management unit of the standardization training work. It is responsible for the relevant standards for the operation of various departments (such as the passenger department, the vehicle department, etc.) Layer; city transportation committee, city passenger transport bureau is mainly responsible for operation and service supervision and management work. Thus, the formation of standardized training and construction of the endogenous mechanism to ensure a clear division of labor, clear responsibility.

2.2.2 Working Mechanism

Operation and service standards training and training should be supplemented by adequate technical and legal support [7], therefore, to establish a resource protection, technical support and legal protection as one of the security mechanism. Through the resource security to achieve operational services standard construction process of logistics and provide operational services in the necessary facilities and equipment; through a wealth of management experience, high-tech level, excellent professional quality of the technical team to standardize the preparation; through professional legal adviser, to ensure the effectiveness of standardized results and to protect intellectual property rights through legitimate means.

2.2.3 Training Standard System Construction

Based on “rail transit operation and management norms”, mainly contented with operation service training system, mass transit company has established relatively complete standardization system as a general guidance which involves traffic service, passenger service, ticketing service standards and many other, as shown in Fig. 1, a comprehensive summary of the contents of the operation services, Put forward the service standard system.

In addition, to specify management, work regulations, work criteria and work management are established to push the integrity and variety of standardization

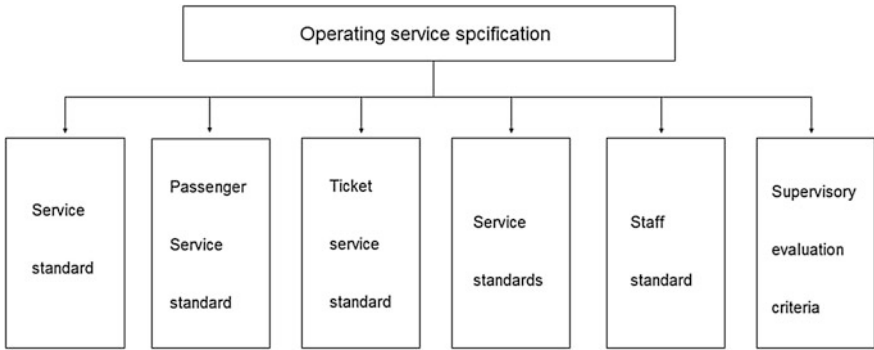
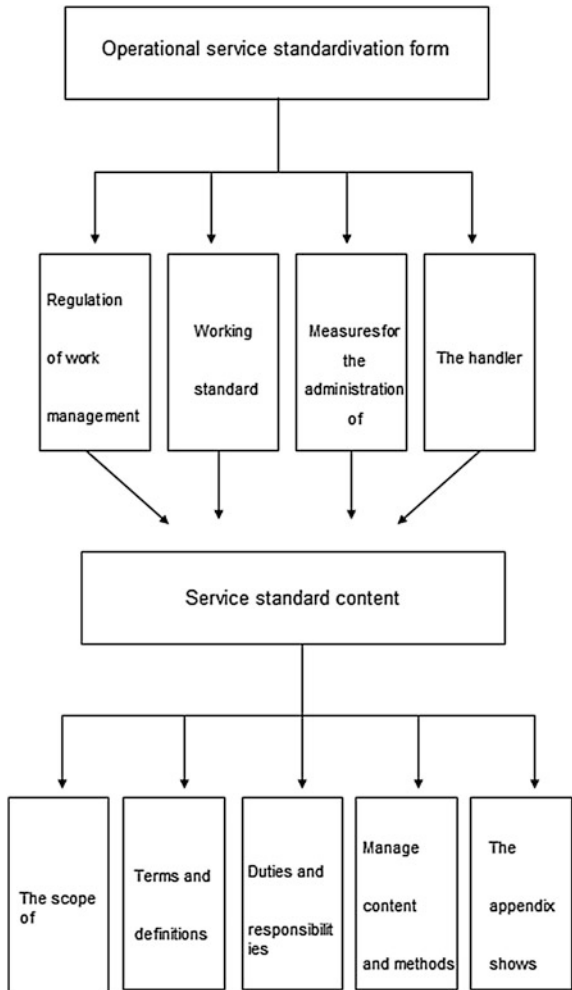


Fig. 1 Operational service specification category

Fig. 2 Operational service guarantee standard form and coverage content



system in the form of standardized documents (Fig. 2). In the content development of the service standard, the preparation of the work scope, service, and technology-related terms and definitions, job responsibilities, management contents and methods, etc., is made in accordance with the specific work of the operational services set up in the “norm”. The document has strong technical guidance and practical operability [8].

3 Standardized Training Construction of Rail Transit Operation Service

Standard training’s methods are various. However, to essentially push the application of standard training, the content of training should be especially stressed. The rail transit group has made a detailed, specific, relevant basis for the standardization of content, strict technical requirements and other requirements to ensure that the operation of high standards of service positioning. Combined with part of the operational services to do some of its training standards to do:

3.1 Service Work Management Training Regulations

It is responsible for the preparation of the passenger service department of the operating branch. The main contents are related to the management training of the passenger service management organization and the duties of the departments, the general management of services, the service management system and the contents, service inspection and evaluation. First of all, clear the service management structure, as shown in Fig. 3, set up the department of integrated, the ministry of vehicles, passenger and other departments with the coordination of various departments and responsibilities; service requirements in accordance with the safety first, operation service and creativity source are required to follow in “service outlines”. The service management training system includes the supporting service facilities, the guidance system, the broadcasting system, and the personal service; the service management training content is divided into the passenger organization, the passenger transaction processing, the passenger security guarantee, the emergency treatment, and so on. Inspection and evaluation of a clear multilevel inspection system.

3.2 Passenger Service Work Standard Training

The staff is responsible for the preparation of the passenger department, for all the windows (referring to the station, train, service hotline) staff to develop a passenger

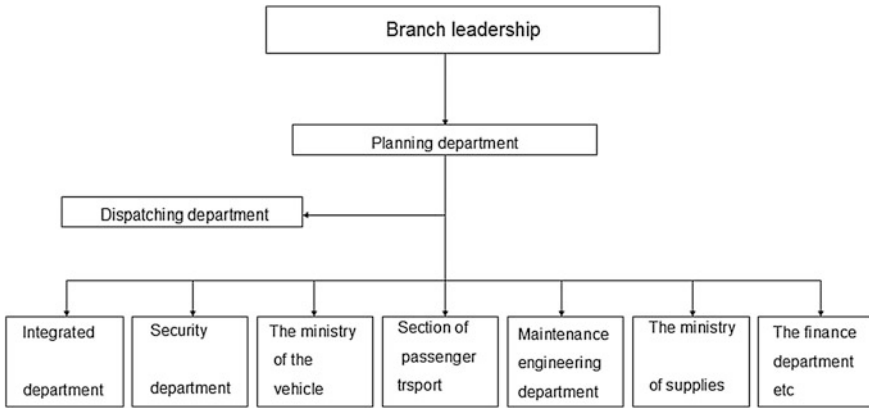


Fig. 3 Operational service management architecture

service work of the general training standards, job standards, service equipment and facilities standards, and common passenger transaction standards. Among them, the common standard should provide service awareness standards, instrument dress standards, behavior standards, service language standards, and environmental health standards. The post-training is for the training rules for the inspection of the inspection, posting, ticket sales, traffic attendants, passenger attendants, duty owners, crew members, service hotline, and other behavioral norms, job skills, service terms. Service facilities training stresses on guide system and its supporting facilities management criteria, such as account management, inspection, maintenance response. Common passengers transaction standard training refers to the passengers stop -ticket -entrance -waiting -ride -get off -exit -out. All these steps will be classified and summarized. Then some corresponding measures are provided with the standard. In addition, in order to fully guarantee the corresponding requirements for passenger demand and special events, it should also regulate the consultation and advising of passenger transaction processing procedures, commemorative transaction procedures, help transaction procedures and other non-complaint transaction procedures, etc. Making the classification of passenger transactions more specific, and the work content more procedural.

3.3 Ticket Management Training Regulations

The responsibility, content and requirement of ticketing System are explicitly stipulated by Ticket card income center, the ticketing section of Metro Operating Company. Among them, the ticketing management content, including ticket management and revenue management, ticket management standardize the fare policy, ticket use, ticket concessions, make up, overtime/overtime, refund/vote to do the provisions of the AFC normal operation mode, Downgrade operation mode and

emergency release mode were explained and provided in different modes of emergency treatment; income management is the provision of the ticketing revenue check, audit, confirmation, and so on. In this regard, the concept of cumbersome and complicated work of the station ticketing work are carried out on the concepts of cash management, ticket management, ticket sales management, station ticket preparation, key management, ticketing safety monitoring management, reporting and special case elaborated, the relevant normative procedures and work requirements to sort out. So as to achieve the strengthening of ticketing management, and the quality of the work. To ensure the stations and the full network ticketing system work normally and orderly.

3.4 Electric Train Quality Inspection Management Training

In addition to service management, passenger service processing, ticketing management, and the operation of the facilities related to the management of equipment are also an important part of the operation of the service. Taking the electric vehicle for example, the training regulations from the train vehicle section of Metro Operating Company state that the job duties, requirements, and rights of quality inspection personnel, and the basis of inspection and evaluation, the overhaul of electric vehicle, spare parts and the inspection of domestic products. To clarify the responsibility, it states that how to process the quality testing and give a working summary by recording “vehicle department technical notice”, “technical notice rectification record table” and accepting corresponding inspection and evaluation.

4 Standardized Future Development Trend of Rail Transit Operation Service

4.1 To Further Improve the Service Needs of Fine Management

With the diversification of passenger service demand, the standardization training of construction services will develop emphasis. Therefore, the standard will be guided by passengers' demand and classified the content of services and training. To meet the different types of services, the corresponding standardization regulations will be builded. Such as the demand for passengers to update the Internet communication, the preparation of operational services, information dissemination training standards, station communications equipment service training standards; to deal with the safety needs of the ride, the development of security requirements or emergency implementation of training procedures.

4.2 Innovation Service Quality Supervision and Evaluation Criteria

The rail transit industry has clearly defined the monitoring and evaluation mechanism. However, due to the lack of experience, the tracking feedback on operational services, traffic safety, and station facilities are still not perfect, and the use of satisfaction surveys in service and event tracking is still tacking and so on. The next step will be innovative supervision and evaluation methods and systems, the preparation of a more systematic operational service supervision and evaluation standards. Such as improving the quality of external service quality supervision and inspection standards, to accept the supervision of the community to improve the quality of service; network operators as the goal, the establishment of operational service quality evaluation system, standardize the evaluation process and methods to form based on the station level, line layer supervision and evaluation standards.

References

1. Ting X (2012) Our government public service standardization construction. China Ocean University, Qingdao (in Chinese)
2. Li X (2012) On the trend of fine management from the practice of standardization of public service—taking the practice of standardization construction of public service in Beijing. Chinese Stand 3:108–111 (in Chinese)
3. Beijing quality and technical supervision bureau. Code for management of urban rail transit operation service. DB11/T647–2009. (in Chinese)
4. Shanghai Metro Shanghai rail transit operation service code [EB/ OL]. (in Chinese)
5. Li L (1997) On the safety behavior science. Chinese J Safety Sci 7(2):10–13 (in Chinese)
6. Qin G, Chen Y, Zhang S Study on the compilation of passenger transport service standard for urban rail transit. Urban Transp. (in Chinese)
7. National Standardization Management Committee Germany and France to actively promote service standardization [EBOL]. <http://www.sac.gov.cn/templet/default/ShowArticle.jsp?Id=26572008>. (in Chinese)
8. Li B (2012) Highway engineering technical standards of economic jurisprudence. Chang'an University. (in Chinese)

Research on Urban Rail Transit Line Collaborative Safety Operation

Xiaowei Shi and Yanhui Wang

Abstract Based on analyzing the structure of urban rail traffic station, combined with moving patterns of train, the paper proposes security grade of passenger density and safety operation effected elements of line. Then, this paper refers to the collaborative management theory and moving patterns of train, and builds a collaborative optimization model of interval. Finally, an experiment is given to illustrate the effect of the proposed model, and the result of the experiment provides multiple plans for satisfying the different demands of the real operation. This research is significant to improve the efficiency of operation, provide guarantee to the safety of operation, and make reasonable traffic plans.

Keywords Urban rail transit · Safety collaborative operation · Passenger flow

1 Introduction

Passengers are the main service object and the trains and the stations are the accommodate space of the rail transit operation system. Therefore, it is significant to research the collaborative operation of the passengers, the trains and the stations. This paper is based on the passenger flow of urban rail transit line cooperative security operation optimization method, mainly including the departure interval optimization model of urban rail transit route and traffic control strategy, to improve

X. Shi (✉) · Y. Wang
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: shixiaowei365@bjtu.edu.cn

X. Shi · Y. Wang
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing 100044, China

X. Shi · Y. Wang
Beijing Jiaotong University, no 3, Shangyuancun Haidian District,
Beijing 100044, People's Republic of China

the efficiency of the train transport organization, maintenance operation safety, and operational management unit reasonable transport organization plan has important guiding significance.

Xu et al. [1] calculated the train departure time according to interval time of train running, and applied the result to planning. Yuhem et al. [2] constructed the lowest operating cost and minimum passenger travel time multi-objective integer programming model. Guo [3] built the different stop scheme optimization models which are based on the stable traffic demand. Zheng and Song [4] based on the minimum travel time and established integer programming model. Zolfaghari [5] calculated travel time which was based on delay time. Suh [6] proved the superiority of cross-site parking when the passenger flow is crowded.

Dai et al. [7] analyzed statistical the causes of the formation of large passenger flow based on Beijing metro passenger flow. Wang [8] formulated the corresponding transportation organization scheme of urban rail transit system under the condition of sudden large passenger flow. Meng et al. [9] transformed multiple objective to single objective through the way of setting weights.

This paper built the collaborative optimization model which is based on the collaborative management theory. Then the fuzzy analysis theory is employed to solve the model. This research is significant to improve the efficiency of operation, guarantee the safety of operation, and make reasonable traffic plans.

2 Operating Safety Influence Factor Analysis of Urban Rail Transit

2.1 The Density of Platform

As the main area getting on or off of passengers, the platform is the normal transport infrastructure of station which can make the organization smoothly. In general, the platform can be divided into waiting area and walking area (as shown in Fig. 1). Waiting area is close to the train carriages and is the edge of the platform, which is

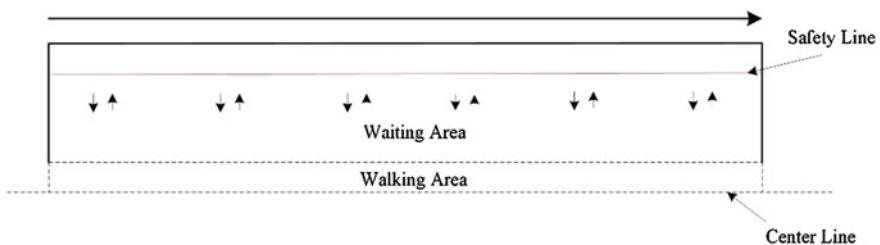


Fig. 1 Structure of platform

used for getting on or off. Walking area is the area which is for passengers from the channel, the staircase (escalator) platform get into or out of the car. In this paper, the platform density of passenger flow refers to the density of waiting area.

Passenger flow density of platform is defined as the number of passengers distributed in the platform, and it can be calculated by the formula 1.

$$\rho_s = Q_s/S \tag{1}$$

- ρ_s The passenger flow density of platform (waiting area) (p/m²)
- Q_s The passenger flow of platform (waiting area) (p)
- S The effective area of platform (waiting area) (m²).

2.2 The Density of Standing Area of Carriage

Passengers in the train can be divided into two states which are standing and seating. So, train carriages are divided into the seat area and standing area (as shown in Fig. 2). Seat area is in commonly set at close to the coach side. And standing area is at edge of seat for no seat passenger. The standing passenger density in the train reflects the aggregation degree of standing passengers.

Passenger carriage density refers to the ratio of standing passenger flow and the effective area of standing area in the carriage. (p/m²).

$$\rho_c = Q_c/S_c \tag{2}$$

- ρ_c Passenger car mat density(p/m²)
- Q_c Standing passenger traffic(p)
- S_c The stand area(m²).

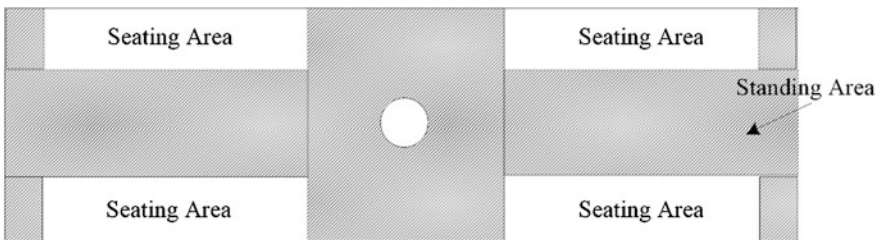


Fig. 2 Structure of carriage

3 Urban Rail Transit Lines Departure Interval Optimization Model

3.1 Description of Operation Process

The research line is shown in Fig. 3. For the line l , a total of m stations, $m - 1$ intervals. The station is divided into three categories, and red station is on behalf of the originating station, yellow station represents the middle stand, green station represents the terminal. In this paper, we study all stations belonging to the same line, so no transfer station in this line.

3.2 Train Departure Interval Collaborative Optimization Model

3.2.1 Parameter

(Table 1).

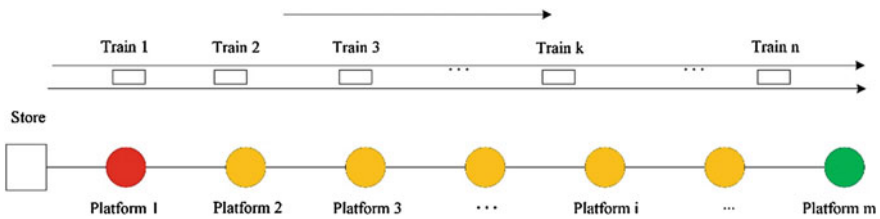


Fig. 3 Sketch of line unidirectional operation

Table 1 Definition of the variables in model

Symbol	Meaning
m	The number of stations in the line
n	The number of trains in the line
$C_{\max}(i)$	The capacity of the platform
n_t	The time period
$S(i)$	The effective waiting area
L_C	The number of seats in the train
C	The capacity of the train
δ_{\max}	The maximum full capacity

3.2.2 Objective

$$\min Z_1 = \sum_{i=1}^{m-1} (F_i - C)^2 \tag{3}$$

The objective aims to maximize the value of the transport capacity.

3.2.3 Constraints

Constraint 1: Train transportation security constraints

The real number in train F_{ij} cannot be larger than the capacity of different trains. When the train type is B type train, the maximum load factor δ_{max} is 1.33.

Constraint 2: Departure interval security constraints

The interval of the departure time can not be too short and it must satisfy the physical limitation of the design capability. According to the existing design capability, the smallest departure intervals for safety is $h_{min} = 90$ s.

3.2.4 Fuzzy Goal Programming Model

In order to solve the multiple objective function of optimal solution, this paper refers to a model based on fuzzy goal programming to solve the optimal solution method.

Assume that there are objectives like $Z_1, Z_2, Z_3, \dots, Z_i$ and the process of calculates the optimal solution as follows.

Step 1: Calculate each of the optimal value of objective function under the constraint condition

Step 2: Calculate the telescopic indicators that reflect the importance the objective function.

$$l_i = \max(z_i) - \min(z_i) \tag{4}$$

Step 3: Fuzzy goals set are given G_i

Step 4: Solve the fuzzy goals G_i .

4 Application Example

4.1 Introduction and Analysis of Lines

Based on data of the Beijing subway line 1 (ascending), this paper takes an experiment to demonstrate the model which is proposed. The line 1 of the Beijing subway is the first line of the Beijing subway. It passes through the city center and it is the central axis of the Beijing. Total length is 31.04 km, and there are 23 stations and two car depot. The train type B car with 6 marshalling plan, the minimum interval is 2 min, target design minimum interval as the 90 s (Table 2).

4.2 Model Calculation

The nonlinear programming professional software lingo 12.0 is used to solve the model. The optimal solution of the objective function is obtained under the constraint condition. Under constraint conditions, the optimal solution of the objective function is:

Plan 1: $n_t = 1$, the departure interval is 1 h.

$$h_1 = 174.98 \text{ s}, Z_1 = 7243766$$

Plan 2: $n_t = 2$, the departure intervals are h_1, h_2 .

$$h_1 = 324.73 \text{ s}, h_2 = 114.16 \text{ s}, Z_1 = 0.1986109E + 08$$

Plan 3: $n_t = 3$, the departure intervals are h_1, h_2, h_3 .

$$h_1 = 402.86 \text{ s}, h_2 = 129.28 \text{ s}, h_3 = 135.98 \text{ s}, Z_1 = 25952082.71$$

Table 2 Parameters of B-train

Length (m)	19	Width (m)	2.8
Marshalling	6	Seat	256
Edge standing area (m ²)	5.6	Mid standing area (m ²)	6.86
Specified capacity(p)	1460	Maximum of capacity (p)	1942

Table 3 Train departure scheme under different statistics length

$n_t = 1$	7:00–8:00			
	$h_1 = 174.98 \text{ s}$			
	$n_{c1} = 22$			
	$Z_1 = 7,243,766$			
$n_t = 2$	7:00–7:30		7:00–7:30	
	$h_1 = 324.73 \text{ s}$		$h_2 = 114.16 \text{ s}$	
	$n_{c1} = 6$		$n_{c2} = 16$	
	$Z_1 = 0.1986109E + 08$			
$n_t = 3$	7:00–7:20	7:20–7:40		7:40–8:00
	$h_1 = 402.86 \text{ s}$	$h_2 = 129.19 \text{ s}$		$h_3 = 135.98 \text{ s}$
	$n_{c1} = 3$	$n_{c2} = 10$		$n_{c3} = 9$
	$Z_1 = 25952082.71$			
$n_t = 4$	7:00–7:15	7:15–7:30	7:30–7:45	7:45–8:00
	$h_1 = 441.32 \text{ s}$	$h_2 = 256.85 \text{ s}$	$h_3 = 92.38 \text{ s}$	$h_4 = 142.05 \text{ s}$
	$n_{c1} = 2$	$n_{c2} = 4$	$n_{c3} = 10$	$n_{c4} = 6$
	$Z_1 = 0.3752250E + 08$			

Plan 4: $n_t = 4$, the departure intervals are h_1, h_2, h_3, h_4 .

$$h_1 = 441.32 \text{ s}, h_2 = 256.85 \text{ s}, h_3 = 92.38 \text{ s}, h_4 = 142.05 \text{ s}, Z_1 = 0.3752250E + 08$$

As you can see the departure schedule of trains under different statistics in Table 3, under the condition that the demand of 1 h cross section transportation is invariable, sending trains under different statistical time constant ($n_{c1} = 22.81$), the passenger flow in section increases first and then decreases in 1 h, that specify the statistical period, sending trains is directly proportional to the number and the section traffic this time.

5 Conclusion

Currently, the researches of the urban rail transit line collaborative operations are focused on using the traditional statistics method to predict or distribute the passenger flow. It cannot adapt to the complex and changeable operation environment. Therefore, this paper studies the density of the passenger flow which can reflect the cluster degree of the passenger flow. Then the train operation process is analyzed and the line cooperative security operations model is built. The model not only considers the interval of the departure time, but also the passenger flow control strategy.

A case study is carried out to demonstrate the computation efficiency and the practical significant.

Acknowledgements The authors gratefully acknowledge the support from “Key Research Project of Safety Assurance Technology of Urban Rail System” under China National “13th Five-Year Plan” (Grant No. 2016YFB1200402-002).

References

1. Hong X, Jian-jun M, Jian-cheng L, Yang H, Zhao L (2006) Study on model and method of train working diagram of urban rail transit. *J Beijing Jiaotong Uni* 30(3):10–14. (in Chinese)
2. Yuhern C, Chunghsing Y, Chingcheng S (2000) A multi-objective model for passenger train services planning: application to Taiwan’ s High-Speed Rail Line. *Transp Res Part B* 34(2): 91–106. (in Chinese)
3. Yu G (2009) Study on optimal problems of stop schedule plan for urban rail transit. Beijing Jiaotong University
4. Zheng L, Song R, He S, Hao-dong L (2009) Optimization model and algorithm of skip-stop strategy for urban rail transit. *J China Rail Way Soc* 31(6):1–8. (in Chinese)
5. Zolfaghari S, Jaber NAMY (2004) A model for holding strategy in public transit system with real-time information. *Int J Transp Manage* 2(2):99–101
6. Suh W, Chon KS, Rhee SM (2007) Effect of skip-stop policy on a Korean subway system. *Transp Res* 1793:33–39
7. Baoqian D, Tong W, Ya C (2009) Analysis and Countermeasures of large passenger flow safety in Beijing Metro. In: Proceedings of the 2009 academic annual conference of China association for occupational safety and health 223–226. (in Chinese)
8. Yinan W (2008) Study on operation organization of urban rail transit system under outburst mass passenger flow. Beijing Jiaotong University. (in Chinese)
9. Xuelei M, Limin J, Meng B, Tao W, Wei M (2012) Train operation design based on decision preference controllable. *J Railway Sci Eng* 9(1):46–50. (in Chinese)

Research on Train Operation Daily Schedule Based on Balanced Use

Bo Wu, Zongyi Xing and Yao Zang

Abstract The scientific arrangement of the train operation daily schedule allocation has important significance to improve the efficiency of train operation. This paper presents a new train number and trip number matching algorithm based on improved best–worst ant system. First, train operation daily schedule is analyzed and the mathematical model of the problem is described. Then, taking the highest degree of match between the train and the trip as the optimization target, uniqueness constraint, morning peak and designated train order constraint, turnout turn-off times minimum time constraint, and convenience constraint as the constraints, a model of the train operation daily schedule based on improved BWAS is established. Finally, the simulation experiment is carried out with the actual trains, trips, and tracks information, the experimental results show that the algorithm can converge in a short time and greatly improve the balanced use of train, and verify the effectiveness of the algorithm.

Keywords Schedule allocation · Balanced use · Improved best–worst ant system

1 Introduction

After selecting the second-day train operation schedule template, select the train in good condition as the corresponding train in the template, which is called the train operation day schedule allocation. At present, the traditional manual allocation method has the characteristics of low production efficiency and high security risks. Therefore, it is of great practical significance to plan the daily operation of train scientifically and rationally.

EMU (Electric Multiple Units) operation planning and planning on the subway operation provide a good state of the vehicle to meet the given task of train by road

B. Wu · Z. Xing (✉) · Y. Zang

School of Automation, Nanjing University of Science and Technology,
No. 200, Xiao Lin Wei, Xuanwu District, Nanjing 210094, China
e-mail: xingzongyi@163.com

transport or operation schedule in the field of railway transportation, so it has a certain significance in the study of EMU scheduling problem [1]. There are few studies currently on the daily schedule allocation of metro vehicles [2–5].

In this paper, the model of train operation daily schedule allocation is set up according to the maximum matching degree of train number and trip number. The model achieves the decoupling of train maintenance and improve the reliability of the train operation and ensure the safety of traffic.

2 Problem Description

A train operation daily schedule as shown in Table 1, includes the morning peak trips and the evening peak trips. Each trips has its departure direction, among them, 0102, 0302, 0402, 0602, 0902, 1102 for the up direction, the rest of the trips for the down direction.

The assignment problem of train operation daily schedule can be expressed as: the schedule/trips information, tracks information, train's information is given, allocating the train operation daily schedule. That is under the constraints of the

Table 1 Train operation daily schedule

Serial number	Trips number	Planned delivery time	Planned return time	Trips number
01	0102	5:37:22	20:40:41	0135
02	0202	5:40:30	9:55:56	0213
03	0302	5:42:42	0:22:26	0305
04	0402	5:47:40	10:40:47	0413
05	0502	5:47:59	0:05:16	0543
06	0602	5:53:44	23:48:00	0641
07	0702	5:55:27	10:10:53	0713
08	0802	6:02:56	23:24:27	0841
09	0902	6:04:00	20:10:47	0933
10	1002	6:10:24	10:25:50	1013
11	1102	6:18:57	20:25:44	1133
12	1202	6:20:32	21:45:24	1237
13	1302	6:28:43	19:55:50	1333
14	1402	6:38:51	22:10:41	1437
15	1502	6:53:55	21:21:39	1535
16	1602	7:08:30	21:37:29	1635
17	1702	16:13:20	0:13:46	1721
18	1802	16:28:17	23:32:22	1819
19	1902	16:38:15	23:40:17	1919
20	2002	16:53:12	23:56:46	2019

minimum interval constraint, the morning peak and the designated trips are required to be restricted, choosing good trains for special train trips, to balance the use of trains, ensuring traffic safety.

3 Design of Train Operation Daily Scheduling Algorithm Based on Improved BWAS

3.1 The Solution Construction Graph and Construction of Solution

Deconstruction construction is essentially a description of the solution space, usually the topological structure of the point-edge structure. Any solution to the optimization problem can be decomposed into construction blocks after processing and mapped into nodes in the graph. In the construction graph, the finite set of nodes and connections is combined with the topology of the network to combine the solution space of the problem [6].

The improved best–worst ant system (BWAS) is adopted in this paper [7]. It is first necessary to determine the cost matrix $[C_{ij}]$ of all the trains to be allocated and the number of trains to be allocated outside the evening peak, where the row represents the train number and the column represents the trips. When the plan is prepared, it is necessary to schedule the trains in order that all the nodes are connected from left to right in turn. The current node can only connect with all nodes in the next column on the right.

All ants start from the starting point and select the transfer node from the optional node in the first column according to the state transition strategy. After the selection is completed, the pheromone is updated and repeated until the last column to complete the construction of the solution. In the solution construction process, if there is a case that a set of available vehicle sets is empty, the current solution is discarded, and the first column is returned to rebuild the solution.

3.2 Objective Functions and Constraint Conditions

3.2.1 Constraint Conditions

The following analysis of the constraints of the allocation of operational planning table includes the unique constraints, turnout conversion trips minimum time

constraints, morning peaks and designated trips constraints, out of library convenience constraints, and so on.

$$\begin{aligned}
 & \text{Min} \left(\sum_{i=1}^{m-k} \sum_{j=1}^{n-k} C_{ij} X_{ij} \right), i = 1, 2, \dots, m - k, j = 1, 2, \dots, n - k \\
 & s.t. \begin{cases} \sum_{i=1}^m X_{ij} = 1, j = 1, 2, \dots, n - k \\ \sum_{j=1}^n X_{ij} \leq 1, i = 1, 2, \dots, m - k \\ P_m = \frac{MP_f}{MP_{total}} = 1 \\ \text{Time}_l > 60, l \in [2, 3, \dots, n] \\ A_{i2} = \begin{cases} 1, & ; T_{i1} \text{ is scheduled train} \\ 1, & \text{otherwise} \end{cases} \end{cases}, \tag{1}
 \end{aligned}$$

where MP_f for the number of completed tasks of the morning peak and designated trips and MP_{total} for the total number of tasks of the morning peak and designated trips. Time_l is the time interval between the current trips and the last trips, and the unit is second. l is the trains needed for switch turnout. A_i indicates whether train T_i can arrange trips. $A_i = 0$ represents for train T_i temporarily cannot arrange train number, $A_i = 1$ represents for T_i can arrange train number.

3.2.2 Objective Functions

The total cost of all trains should be the lowest when carrying out the operation daily schedule, that is, the total matching degree of the total trips and trains is the highest, and the objective function is as follows:

$$\min \left(\sum_{i=1}^{m-k} \sum_{j=1}^{n-k} C_{ij} X_{ij} \right), i = 1, 2, \dots, m - k, j = 1, 2, \dots, n - k, \tag{2}$$

C_{ij} represents the cost of train T_i takes train number F_j , that is, the smaller the value of C_{ij} , the higher the match degree between train T_i and train number F_j . The decision variable is X_{ij} , when $X_{ij} = 0$, the train T_i does not serve as the train number F_j . When $X_{ij} = 1$, the representative train T_i takes the train number F_j .

- Step 1 The number of trains with morning peak or designated trips to fill in the operating plan corresponding to the trips.
- Step 2 For all unsettled trips (except for evening peak trips), the corresponding mileage is sorted from small to large.

- Step 3 Trains with no scheduled are ordered according to the rules.
- Step 4 Match the trains and trips in order.
- Step 5 The matching degree C_{ij} of train T_i and trip F_j is calculated as follows:

$$C_{ij} = |R_j - R_{best}|, \tag{3}$$

where R_j stands for the rank of the trip F_j in all trips, R_{best} stands for the rank of the best matching trips of train T_i in all trains.

3.3 Representation, Initialization, and Updating of Pheromones

In the process of creating a solution, each ant chooses a node v_{ij} , which applies the local update rule of the next type to update the pheromone of the selected node.

$$\begin{cases} \tau_{ij} = (1 - \rho) \cdot \tau_{ij} + \rho \cdot \tau_{ij}(0) \\ 0 < \rho < 1 \\ \tau_{ij}(0) = K \end{cases}, \tag{4}$$

where the pheromone τ_{ij} is placed on each node to represent the expected degree of the train T_i serve as the trip F_j . ρ stands for the coefficient of volatile information. At the initial moment, the pheromone traces on each path are equal and K is constant.

In order to make the search process more instructive when all the ants complete a cycle, the field of ants concentrated in the current cycle which is the best path area so far, the need for the global optimal and the global worst path pheromone track update [8].

- Step 1 Find the global optimal path to the current loop.

$$\tau_{ij} = (1 - \rho) \cdot \tau_{ij} + \gamma \cdot \Delta\tau_{ij}, \tag{5}$$

where γ is a parameter, L_{best} is the global optimal path length for the current loop.

- Step 2 Find the path of the worst ants for the current loop, and update the pheromone of the node that belongs to the worst path but not the global optimal path by applying the global updating rule of the next type.

$$\tau_{ij} = \tau_{ij} - \varepsilon \cdot \frac{L_{best}}{L_{worst}}, \tag{6}$$

where ε is a parameter, L_{best} is the global optimal path length for the current loop. L_{worst} is the worst path length for the current loop.

3.4 State Transition Strategy

Unlike the ant state transition strategy in the classical TSP problem, the visibility of the path in this paper is determined by the matching degree between the current train and the trips. The higher the matching degree, the greater the visibility [9].

- Determining the optional train parking position according to the track constraints of current trip and determining the optional trains set allowed_{k1} in conjunction with the trains parking position table.
- Filter out the train set allowed_{k2} in the current state of 2 according to the current vehicle disabled table.
- This paper assumes that the ants ignore the existence of the stimulus if the stimulus on the path does not reach the ant’s sensory threshold. The train T_i served as the trip F_j with the following probability in the initial several generations of N .

$$s = \begin{cases} \arg \max_{s \in \text{allowed}_k} \{ \eta_{sj} \}, & \text{if } q > q_0 \\ \frac{\eta_{ij}(t)}{\sum_{s \in \text{allowed}_k} \eta_{sj}(t)}, & \text{otherwise} \end{cases} \quad (7)$$

- In the subsequent cycle, the appropriate train is selected as the next train based on the transmission probability. The transfer probability is as follows:

$$p_{ij}^k = \begin{cases} \frac{\tau_{ij}^\alpha(t) \cdot \eta_{ij}^\beta(t)}{\sum_{s \in \text{allowed}_k} \tau_{sj}^\alpha(t) \cdot \eta_{sj}^\beta(t)} & i \in \text{allowed}_k \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

4 Simulation and Analysis

In order to verify the effectiveness of the algorithm, Java is used as the development language. The operating environment is windows 10 system and MyEclipse is used as a development tool to establish a simulation platform for train operation daily schedule allocation.

First of all, collecting of algorithm data which mainly includes three parts, trips information, tracks information, trains information (Tables 2, 3, and 4).

In this paper, the improved best–worst ant system is used to allocate the daily operation schedule. Assume that the initial ant population was 50, the pheromone factor was 1, the expected inspiration factor was 5, and the volatility coefficient of

Table 2 Trips information

Id	Train number	Train direction	Planned travel time	Track constraint	Kilometers per hour	Morning/ evening peak
01	0102	Up	05:37:22	No	34	
02	0202	Down	05:40:30	No	12	Morning peak
03	0302	Up	05:42:42	No	4	
04	0402	Up	05:47:40	Yes	12	Morning peak
05	0502	Down	05:47:49	Yes	42	
06	0602	Up	05:53:44	No	40	
07	0702	Down	05:55:27	No	12	Morning peak
08	0802	Down	06:02:56	No	40	
09	0902	Up	06:04:00	No	32	
10	1002	Down	06:10:24	No	12	Morning peak
11	1102	Up	06:18:57	No	32	
12	1202	Down	06:20:32	No	36	
13	1302	Down	06:28:43	No	32	
14	1402	Down	06:38:51	No	36	
15	1502	Down	06:53:33	No	34	
16	1602	Down	07:08:30	No	34	
17	1702	Down	16:13:20	No	20	Evening peak
18	1802	Down	16:28:17	No	18	Evening peak
19	1902	Down	16:38:15	No	18	Evening peak
20	2002	Down	16:53:12	No	18	Evening peak

Table 3 Tracks information

Id	Track	Occupancy	Occupancy train number
01	4AG	Yes	8A159160
02	4BG	Yes	
03	5AG	Yes	6970
04	5BG	Yes	Engineering train
05	6AG	Yes	8A157158
06	6BG	Yes	
07	7AG	Yes	5960
08	7BG	No	
09	8AG	Yes	5556
10	8BG	Yes	
11	9AG	Yes	8A137138
12	9BG	Yes	
13	10AG	Yes	7374
14	10BG	Yes	6566
15	11AG	Yes	8788
16	11BG	No	
17	12AG	Yes	8384
18	12BG	Yes	Engineering train

Table 4 Trains information

Id	Train number	Condition	Assign task	Baseline total mileages	Days from baseline	Current total mileages	Daily mileages
01	5354	Unavailable		1,398,214	29	1,398,214	0
02	5556	Good		1,548,797	29	1,558,022	318.1
03	5758	Good		1,442,316	29	1,452,872	364
04	5960	Good		1,500,341	29	1,511,217	375.03
05	6162	Good	0502	1,473,398	29	1,473,557	729.62
06	6364	Good		1,550,668	29	1,560,963	355
07	6566	Good		1,495,392	29	1,504,471	313.07
08	6768	Good		1,568,621	29	1,578,227	331.24
09	6970	Good		1,565,201	29	1,572,193	241.1
10	7172	Good		1,538,759	29	1,548,880	349
11	7374	Good	Morning peak	1,538,872	29	1,544,027	177.76
12	7576	Good		1,439,209	29	1,446,799	261.72
13	7778	Unavailable		1,449,611	29	1,458,118	293.34
14	7980	Good		1,547,899	29	1,557,220	321.41
15	8182	Unavailable		1,440,500	29	1,448,463	274.59
16	8384	Good		1,406,149	29	1,410,736	158.17
17	8586	Good		1,447,421	29	1,456,618	317.14
18	8788	Good		1,392,956	29	1,396,922	136.76
19	8990	Unavailable		1,426,144	29	1,426,144	0
20	8A135136	Good		282,275	29	289,000	231.9
21	8A137138	Good		326,976	29	336,010	311.52
22	8A157158	Good		353,206	29	365,096	410
23	8A159160	Good		288,152	29	297,946	337.72
24	8A163164	Good		127,198	29	136,705	327.83
25	9192	Unavailable		1,459,058	29	1,471,137	416.52

the pheromone was 0.2, $K = 0.5$, $q_0 = 0.2$, the maximum cyclic algebra was 80. The convergence of the algorithm is shown Fig. 1.

The daily mileage traveled by the train during the end of the operation schedule is shown in Fig. 2. It can be seen that the balance of train operation has been improved slightly.

This paper assumes that the vehicle information, stock information, trips information, and other conditions remain unchanged, the schedule is repeated 19 times to better illustrate the trend of the trains' daily mileage as shown in Fig. 3.

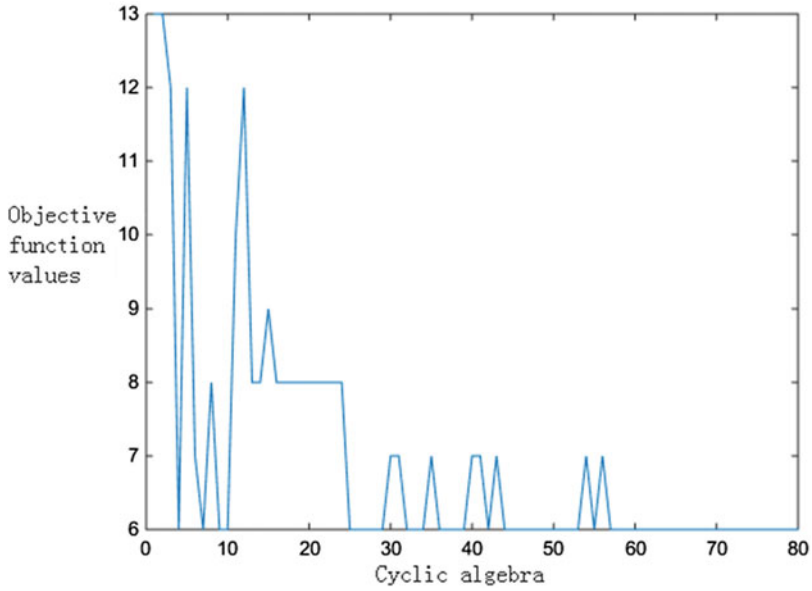


Fig. 1 The convergence of the improved optimal worst ant colony algorithm

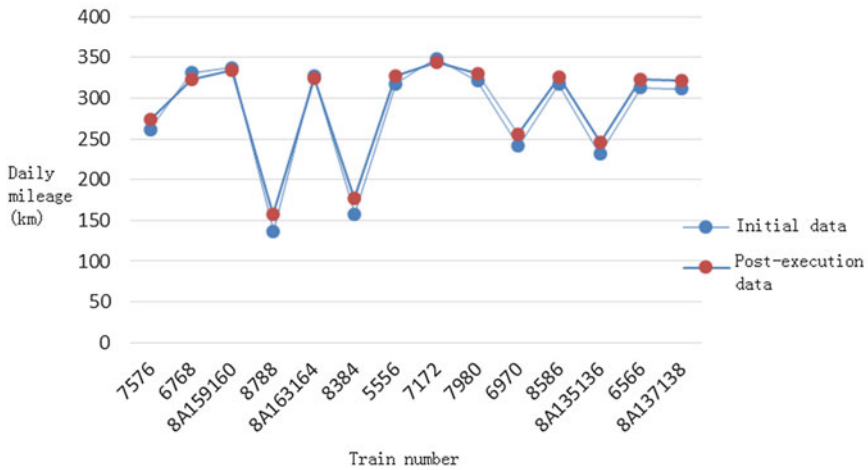


Fig. 2 The daily mileage traveled by the train

It is seen that in the application of the schedule for about 19 days, the daily traveling distance of all the trains tends to converge. In summary, the algorithm can be converged in a short time and the balanced use of train is improved significantly which verifies the effectiveness of the algorithm.

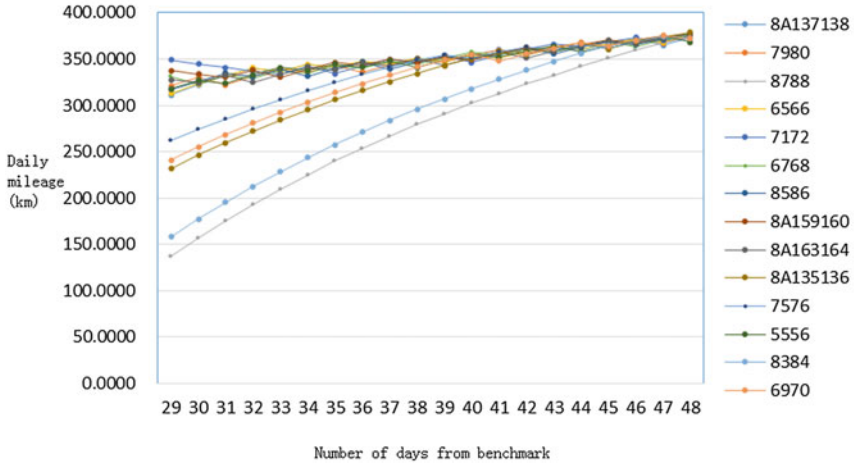


Fig. 3 The trend of the trains' daily mileage by repeated the schedule 19 times

5 Conclusion

This paper proposes a train matching algorithm based on the improved best-worst ant system to realize the balanced use of the trains. First of all, the actual situation of the vehicle is investigated, and the content of the research is refined. Then, the construction of the solution and the state transition strategy are designed, and the objective function and constraints of the problem are determined. And the design algorithm is designed and improved by combining the problem characteristics. Finally, the simulation experiment is carried out with the actual vehicle, trips, and stock information. The simulation results show that the algorithm can converge in a short time and improve the balanced use of train and verify the effectiveness of the algorithm.

Acknowledgements This work is supported by National Key R&D Program of China (2017YFB1201202).

References

1. Zhongkai W, Tianyun S (2012) Model and algorithm for the integrative scheduling of EMU utilization plan and maintenance plan. *China Railway Sci* 33(3):104–110 (in Chinese)
2. Lu T (2009) Application of ant colony algorithm in train-set scheduling problem. *J Transp Syst Eng Inf Technol* 9(6):161–167 (in Chinese)
3. Chen L (2009) Application of genetic algorithms to electric multiple unit scheduling. *J Transp Eng Info* 7(2):67–71 (in Chinese)

4. Hong SP, Kim KM, Lee K et al (2009) A pragmatic algorithm for the train-set routing: the case of Korea high-speed railway. *Omega* 37(3):637–645
5. Abbink E, Van den Berg B, Kroon L et al (2004) Allocation of railway rolling stock for passenger trains. *Transp Sci* 38(1):33–41
6. Wen Y (2014) Ant colony optimization method for complex multi-stage dynamic decision making and its application in traffic system control. Zhejiang University. (in Chinese)
7. Hongbing L (2010) Static clustering routing algorithm based on best-worst ant system for wireless sensor networks. *Telecommu Eng* 50(4):96–101 (in Chinese)
8. Dorigo M, Birattari M, Stutzle T (2006) Ant colony optimization. *IEEE Comput Intell Mag* 1(4):28–39
9. Ghoseiri K, Nadjari B (2010) An ant colony optimization algorithm for the bi-objective shortest path problem. *Appl Soft Comput* 10(4):1237–1246

Evaluating the Efficiency of Urban Public Transit Enterprises Based on DEA Approach with Preference

Rong Yu and Jiaqi Sun

Abstract Performance evaluation about urban public transport is one of the keys to ensure the developing of urban public transport. As the traditional evaluation model cannot reflect industry policy information and management ideas of administrators. This paper establishes DEA model with preference. And then we use triangular fuzzy numbers to calculate preference restraint cone. After establishing a set of evaluating indicator system, this paper uses DEA method with preference for public transport enterprises performance evaluation. By comparing the results of basic CCR model and the DEA method with preference, we get the advantages of DEA method with preference which can reflect more management information and the decision-maker preferences. Based on the analysis of the efficiencies, bus companies should utilize customer satisfaction as a breakthrough to improve operational efficiency.

Keywords Performance evaluation · Urban public transit · Preference restraint cone · Triangular fuzzy number · Data envelopment analysis (DEA)

1 Introduction

With the take-off of the national economy and the advancement of the urbanization, the urban aborigines and the floating population are increasing, and the urban vehicle ownership is increasing rapidly, but it also causes the traffic jam, the environmental pollution, and so on. This series of problems not only seriously affects the travel and health of residents but also restricts the steady growth of the economy and the sustainable development of cities, which has aroused the attention of all levels of government and relevant departments nationally. Nowadays, local governments transfer urban public transport to bus companies and subsidize them in

R. Yu (✉) · J. Sun

School of Management Science and Engineering, Nanjing University of Finance and Economics, No. 3Wenyuan Road Qixia District, Nanjing, Jiangsu, China
e-mail: 157868984@qq.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_85

841

response to the call of the state. Therefore, the key to promote the development of urban public transport is to improve the management and operation level of public transport companies. However, how to evaluate the management and operation level of public transport company scientifically and effectively and put forward the definite improvement opinion according to the result of the evaluation, so as to improve the operation efficiency of the public transportation company, becomes a difficult point of disturbing the development.

In recent years, scholars at home and abroad use different methods to evaluate the performance of public transport management. The study of the performance of urban public transport by foreign scholars mainly concentrates on the service quality of bus companies. Stuart et al. [1] explained the relationship between urban public transport service quality and passenger satisfaction by structural equation model. Laureshyn et al. [2] designed the survey plan of urban public transport service quality. The United States public transport capacity and quality of service manual (TCQSM, transit Capacity and quality of service manual, 2003) [3] studied index of evaluating the quality of public transport service, and the framework of service quality is constructed. Domestic scholars mainly focus on the selection of urban public transport evaluation indicators and improve urban public transport evaluation methods, such as Xu and Xi [4], Wang et al. [5], and Chao et al. [6].

Data Envelopment Analysis (DEA) [7] was proposed by a well-known research scholar, A. Charnes and W. Cooper in 1978. Cui and Li [8] used three-stage DEA model for performance evaluation of traffic energy efficiency; Murray et al. [9] carried out a new traffic demand management scheme. Lao and Liu [10] evaluated the performance of bus line in public transport system. Ding et al. [11] established performance evaluation index of China's transportation industry; Wang [12] used gini criterion to reduce the dimension of the DEA method. Triangular fuzzy number in fuzzy mathematics is to solve the problem of multiple decision-making [13].

2 The Operational Performance Evaluation Model of Urban Public Transport Companies Based on Preference DEA Method

2.1 Evaluation Index Selection of Operation Performance of City Bus Company

This paper chooses the evaluation index from the angle of input and output, respectively. From the point of view of input, the input of urban public transport is put into the cost item, which is embodied in fuel cost, labor-capital cost, vehicle depreciation cost, and other expenses. From the point of view of output, serving the people, convenient travel is the main purpose of the public transport industry, so it is the key to establish the whole performance evaluation system to evaluate the passenger's perception of public transport service as one of the output indexes.

In addition, because the most basic function of public transport is to realize people’s displacement, this article will be passenger volume, operating mileage as the public transport company operating output indicators. Finally, this paper establishes a set of urban public transport industry characteristics of the index system (as shown in Table 1).

2.2 The Operational Performance Evaluation Model of Urban Public Transport Companies Based on Preference DEA Method

In this paper, the CCR model is improved to the DEA model C^2WH with preference constraint cone on the basis of the original data envelopment analysis theory, and the model description is shown in (1). In the model U, V is the constraint cone which shows the importance degree of input and output indexes according to the preference of industry policy and decision-maker. Through the Charnes–Cooper transform, the original fraction model C^2WH is transformed into a linear programming model (P_{C^2WH}), and the model description is shown in (2).

$$\begin{aligned} \max h_{jo} &= \frac{\sum_{r=1}^s u_r y_{rjo}}{\sum_{i=1}^m v_i x_{ijo}} \\ \text{s.t. } \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} &\leq 1, \quad j = 1, 2, \dots, n; i = 1, 2, 3, 4; r = 1, 2, 3 \end{aligned} \tag{1}$$

Table 1 Index set of input–output and related index

	Indicators	Descriptions
Input	Fuel cost X_1	Petrol, diesel, and other oil charges for public transport
	Labor costs of employees X_2	Employees’ wages, bonuses, medical insurance, etc.
	Cost of vehicle depreciation X_3	Financial form of vehicle investment recovery
	Other costs X_4	Loss of traffic accidents, insurance costs, etc.
Output	Passenger volume (million people) Y_1	Total passengers transported
	Operating mileage (million kilometers) Y_2	The total mileage of the vehicle running for the purpose of operation
	Passenger satisfaction Y_3	Based on the data of the market questionnaire, using structural equation models (Structural Equation model, SEM) to calculate the obtained

$$\begin{aligned}
 \max h_{j_0} &= \mu^T y_o \\
 s.t. \begin{cases} w^T x_j - \mu^T y_j \geq 0, & j = 1, 2, \dots, n; i = 1, 2, 3, 4; r = 1, 2, 3 \\ w^T x_0 = 1 \\ w \in V = \{V|V \geq 0\} \\ \mu \in U = \{U|U \geq 0\} \end{cases} \tag{2}
 \end{aligned}$$

If the linear programming problem has the optimal solution U^*, V^* which satisfies $U^{*T} Y_0 = 1$ and $U^* \in U, V^* \in V$, then the decision unit DMU_{j_0} is called the DEA effective.

2.3 The Construction of DEA Model Preference Constrained Cone Based on Triangular Fuzzy Number

First, the questionnaire is used to collect the preference information of the decision-makers on evaluation metrics. The input and output indicators are graded by decision-makers on the basis of scale (Table 2), with full knowledge of the evaluation indicators. In terms of input indicators, if policymakers wish to see a significant reduction in the input of this indicator, the weighted score should be the highest and vice versa. In terms of output indicators, if policymakers wish to see a significant increase in the output of this indicator, the weights given to this indicator should be highest and vice versa.

Second, the preference data of the decision-makers is synthesized. This article assumes that all participating decision-makers have the same evaluation authority; so after collecting the preference data of all decision-makers, the results of the evaluation of input–output index preference by all decision-makers are combined with the arithmetic average method $\bar{A}_i = (A_{i1} + A_{i2} + \dots + A_{ik})/k, \bar{B}_i = (B_{i1} + B_{i2} + \dots + B_{ik})/k$, respectively, and the triangular fuzzy numbers corresponding to each data are further obtained according to Table 2.

Table 2 0.1–0.9 the meaning of scale

0.1–0.9 five scale	Corresponding triangular fuzzy numbers	Meaning
0.1	(0.1, 0.1, 0.2)	Indicator I is extremely unimportant in all indicator
0.3	(0.2, 0.3, 0.4)	Indicator I is generally unimportant in all indicators
0.5	(0.4, 0.5, 0.6)	Indicator I is relatively important in all indicators
0.7	(0.6, 0.7, 0.8)	Indicator I is generally important in all indicators
0.9	(0.8, 0.9, 0.9)	Indicator I is extremely important in all indicators

Then, based on the above-mentioned comprehensive index preference data, the preference constraint interval is calculated. $a = (a_l, a_m, a_u)$ represents decision-makers' preference for indicator X_1 , and $b = (b_l, b_m, b_u)$ represents the preference of decision-makers for the indicator X_2 . According to the formula (5) (6), the calculated $\frac{a}{b} = \left(\frac{a_l}{b_u}, \frac{a_m}{b_m}, \frac{a_u}{b_l}\right)$ is obtained, and the preference constraint interval of index 1 to index 2 is $\left(\frac{a_l}{b_u}, \frac{a_u}{b_l}\right)$. According to this method, the preference relation between input indexes and the preference relationship between output indices are calculated.

Finally, the constraint interval $\left(\frac{a_l}{b_u}, \frac{a_u}{b_l}\right)$ can be transformed into the constraint inequality $\begin{cases} -b_l a + a_u b \geq 0 \\ b_u a - a_l b \geq 0 \end{cases}$, and thus the preference constraint matrix $W_{1,2} = \begin{bmatrix} -b_l & a_u \\ b_u & -a_l \end{bmatrix}$ of the index X_1 and the index X_2 is obtained; the remaining input and the output index preference constraint inequality is calculated further according to the comprehensive index preference data. We can get input index preference constraint cone matrix W and output index preference constraint cone matrix V .

The preference constrained cone W, V are taken into the preference DEA model, and the DEA effectiveness with preference of each DMU is calculated, whose the discriminant ability is better than that of the traditional DEA method. The performance evaluation of City Bus Company is more objective using DEA model with preference.

3 Performance Evaluation of Public Transport Company in Nanjing

This paper chooses Nanjing Seven bus companies to carry out performance evaluation of each bus company by collecting their original data of operation in the first half of 2011 and the preference data of decision-makers, using the DEA model proposed above. The following data is derived from the data envelopment analysis software EMS1.3.

3.1 Data Collection

In the first half of 2011, Nanjing City Bus Company's fuel costs, workers' labor costs, vehicle depreciation costs, other costs, passenger satisfaction, passenger volume, operating mileage specific data reference from the Institute of Quality and safety of Jiangsu Province, *urban public traffic Management performance evaluation system and its subsidy policy* research report are shown in Table 3.

Table 3 Performance evaluation data of Nanjing Public transport industry

DMU	Input			Output			
	Fuel costs (million) X_1	Labor costs of workers (million) X_2	Vehicle depreciation costs (million) X_3	Other costs (million) X_4	Passenger volume (million people) Y_1	Operating mileage (million people) Y_2	satisfaction Y_3
Public Transport Corporation	32674.73	53715.45	13470.47	3368.88	50057.13	18581.94	60.87
North Central Bus	12614.8	17051.8	4792.8	1293.8	20852.12	7381.98	63.77
Accor Bus	5684.55	5897.24	2360.46	850.49	7591.78	2939.93	55.56
Metro Bus	7037.31	9621.09	2471.73	1399.18	8511.25	4754.78	50.48
Xinning Company	2786.8	3058.44	822.22	560.71	4364.7	1755.61	56.68
Pukou Passenger	2246.12	3119.24	620.48	288.79	3127.54	1878.91	61.83
Liuhue Passenger	3706.32	4265.04	760.27	273.74	2947.18	2938.86	61.93

According to the questionnaire survey data of the public transport management personnel and the transportation professional researchers, this paper applies the arithmetic average method $\bar{A}_l = (A_{i1} + A_{i2} + \dots + A_{ik})/k$ arithmetic averaging, $\bar{A}_l = (A_{i1} + A_{i2} + \dots + A_{ik})/k, \bar{B}_l = (B_{i1} + B_{i2} + \dots + B_{ik})/k$. The evaluation data of decision-makers' preference for each index is calculated comprehensively in Table 4.

3.2 Performance Evaluation of Urban Bus Companies Based on Preference DEA Method

This paper calculates the input index preference cone matrix W and output index preference cone matrix V according to the data provided in Table 4, as shown in the following formula.

$$w = \begin{bmatrix} -4 & 9 & 0 & 0 \\ 3 & -4 & 0 & 0 \\ 0 & -1 & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -1 \\ -2 & 0 & 9 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & -1 & 0 & 6 \\ 0 & 1 & 0 & -2 \\ -1 & 0 & 0 & 9 \\ 1 & 0 & 0 & -4 \end{bmatrix} \quad v = \begin{bmatrix} -2 & 9 & 0 \\ 1 & -2 & 0 \\ 0 & -3 & 2 \\ 0 & 4 & -1 \\ -2 & 0 & 3 \\ 1 & 0 & -1 \end{bmatrix}$$

In this paper, the input-oriented CCR model and the preference DEA model are taken as examples to calculate the data in Table 3; compare the results of the performance evaluation of each enterprise based on the CCR model and the preference DEA model, and the final efficiency results are described in Table 5.

3.3 Data Analysis of Performance Evaluation Results

From Table 5, we can see that the efficiency value of performance evaluation is obtained using DEA model with preference. Only the Xinning company for the DEA effective, while the efficiency value of more than 90% of the only North China bus and Pukou two companies, shows that most of the public transport companies in the allocation of resources are more or less unreasonable.

We can further compare the difference between an input-type CCR model and a preference DEA model based on the histogram of each efficiency value, as shown in

Table 4 Index preference comprehensive evaluation data

DMU	Input				Output			
	Fuel costs (million) \bar{A}_1	Labor costs of workers (million) \bar{A}_2	Vehicle depreciation costs (million) \bar{A}_3	Other costs (million) \bar{A}_4	Passenger volume (million people) \bar{B}_1	Operating mileage (million people) \bar{B}_2	Satisfaction \bar{B}_3	
Comprehensive data	(0.8, 0.9, 0.9)	(0.4, 0.5, 0.6)	(0.2, 0.3, 0.4)	(0.1, 0.1, 0.2)	(0.8, 0.9, 0.9)	(0.2, 0.3, 0.4)	(0.6, 0.7, 0.8)	

Table 5 Efficiency data calculated by traditional CCR model and preference DEA model

DMU	Efficiency		Efficiency ranking	
	CCR model (%)	DEA model with preference (%)	CCR model	DEA model with preference
Public Transport Corporation	94.51	83.58	3	5
North Central Bus	100.00	97.19	1	2
Accor Bus	95.25	85.95	2	4
Metro Bus	84.61	75.63	4	6
Xinning Company	100.00	100.00	1	1
Pukou Passenger	100.00	90.93	1	3
Liuhe Passenger	100.00	65.43	1	7

Fig. 1. It can be seen from Fig. 1 that the difference between the two models is large, and the biggest difference is that the results of the CCR model are significantly higher than those of the DEA model. In the efficiency value calculated by the CCR model, the efficiency is 1, that is, the number of bus companies with effective DEA exceeds half the total number, to 57.14%. In the efficiency value of the DEA model, only one company is effective for DEA. This is due to the result of the improvement of the index weights of the underlying model. In the calculation of the CCR model based on input, all indexes have equal weights without advance weight preset and limitation, and the preference DEA model in this paper is based on the policy of the decision-maker and the background-oriented; the calculation method of weight constraint is considered synthetically. We can analyze the improvement

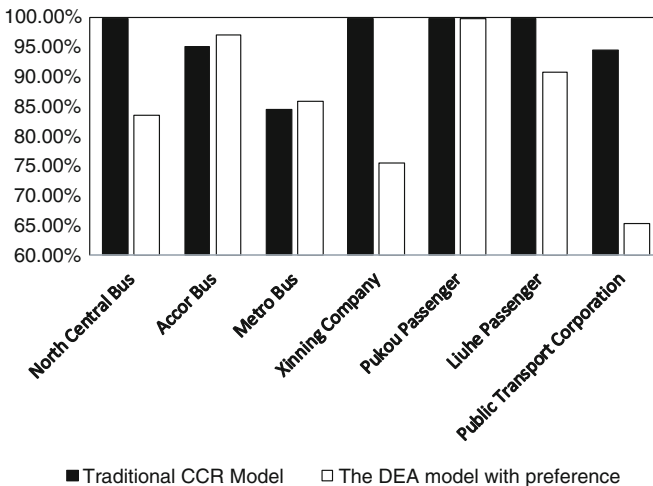


Fig. 1 Comparison between traditional CCR efficiency and DEA efficiency with preference

of the preference DEA model relative to the standard DEA model by comparing the DEA effectiveness in the CCR model with the DEA effective unit.

The calculated results of the preference DEA model contain more information than the calculated results based on the CCR model, which not only emphasizes the influence of the industry policy on the target enterprise evaluation but also emphasizes the preference of the decision-makers to the relative importance of each index. The DEA model with preference has two advantages: one is to be able to put forward more strict management and operation requirements to the bus company, and the second is that the model is more in line with the actual situation of the whole industry development and can better guide the development of the bus company industry.

3.4 Suggestions for Improving the Operation of Nanjing Bus Company

Combined with Tables 3, 4, 5, and 6, we can make further suggestions for the various bus companies.

As the largest company in the seven companies, the efficiency is not high but in terms of customer satisfaction and the maximization of passenger volume to do a good job, the next need to find the higher cost reasons and take measures to reduce fuel costs and labor costs. By contrast, Yagao Bus and Xincheng Bus are medium-sized bus companies, on the one hand they have to find ways to significantly reduce operating costs, on the other hand improve the operational strategy to improve customer satisfaction and passenger traffic. As a large-scale bus company in the north, the overall operation of the high level then needs to work in customer service to improve the enterprise's passenger recognition and market competitiveness. As a small-scale bus company, the management and operation cost of Pukou passenger transport is too high, and then the utilization efficiency of vehicle and cost should be improved, and the related investment should be reduced scientifically. As a small-scale bus company, the operating cost of Linhe Bus is too high while the operating efficiency is too low, so in the next period of time, Liuhe passenger traffic due to seize the time to find all aspects of the problem, particularly in the cost of investment and customer service, with a view to early rectification, While the most efficient of the xinning passenger transport, although the effectiveness of the DEA is still low, the customer satisfaction is required to improve customer satisfaction to find a breakthrough, and pay attention to other aspects of the problem, timely detection, timely solution to consolidate their market performance and enterprise recognition.

Table 6 Analysis of input–output projection of bus companies based on DEA model

Decision unit	Input redundancy				Insufficient output			
	Fuel costs (million)	Labor costs of workers (million)	Vehicle depreciation costs (million)	Other costs (million)	Passenger volume (million people)	Operating mileage (million people)	Satisfaction	
Public Transport Corporation	4493.01	6119.45	0	219.22	0	529.31	0	
North Central Bus	371.11	0	0	0	0	0	77.02	
Accor Bus	648.01	104.31	415.45	0	1439.11	405.35	43.03	
Metro Bus	1059.81	0	377.36	407.46	1004.40	0	35.54	
Xinning Company	0	0	0	0	0	0	0	
Pukou Passenger	645.49	0	711.86	0	0	351.44	0	
Litue Passenger	1044.3	573.56	0	0	2044	0	97.04	

4 Conclusions

This paper takes the operation data of Nanjing Bus Company in the first half of 2011 years as an example, on the basis of establishing a set of index system which takes into consideration the public transport commonweal and commercial of the city, constructs the preference constraint cone through the triangular fuzzy number method, and obtains the considerable performance appraisal result by the DEA method where the results show that the preference DEA model embodies the preference and policy orientation of decision-makers, which can reflect more information than the standard DEA model. To improve the performance direction of urban public transport enterprises, it is helpful to improve the service quality of public transport companies, to ensure the smooth implementation of the strategy of *giving priority to the development of urban public transport*, and to improve the market performance of bus companies.

Acknowledgements Financial support from the project of natural science research of Jiangsu Province—RongYu and Jiaqi Sun, are gratefully acknowledged by the first and second authors, under the contract 14KJB580006.

References

1. Stuart KR, Mednick M, Bockman J (2000) Structural equation model of customer satisfaction for the New York City subway system. *Transp Res Rec* 1735:133–137
2. Laureshyn A, Ardö H, Svensson Å, Jonsson T (2009) Application of automated video analysis for behavioural studies: concept and experience. *IET Intell Transp Syst* 3(3):17–22
3. TCRP. Transit Capacity and Quality of Service Manual. Transportation Research Board, Washington DC, pp 73–154
4. Qun X, Xi C (2006) Index system of urban public transport service level. *City Traffic* 06:42–46 (in Chinese)
5. Wang H, Tang R, Zheng YJ (2011) Research on performance evaluation system of urban public transport industry. *China's Indust Economy* 03:68–77. (in Chinese)
6. Yang C, Zhongbin X (2007) Evaluation model of public transit network based on macro-simulation simulation. *Syst Eng* (12). (in Chinese)
7. Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Operat Res* 2(6):429–444
8. Cui Q, Li Y (2014) The evaluation of transportation energy efficiency: an application of three-stage virtual frontier DEA. *Transp Res Part D Transp Environ* 6(29):1–11
9. Zhao Y, Triantis K, Murray-Tuite P et al (2011) Performance measurement of a transportation network with a downtown space reservation system: a network-DEA approach. *Transp Res Part E* 47:1140–1159
10. Lao Y, Liu L (2009) Performance evaluation of bus lines with data envelopment analysis and geographic information systems. *Comput Environ Urban Syst* 33:247–255
11. Ding Xiaodong X, Xian YZ (2011) Performance evaluation of China's transportation industry based on DEA method. *J Wuhan Uni Sci Res* 03:77–81 (in Chinese)
12. Wang H, Zheng YJ, Tang R (2012) Application of Dea-Gini criterion in performance evaluation of urban public transport enterprises. *Syst Eng Theory Pract* 33(5):1083–1090. (in Chinese)
13. Zhang D, Wang G (2005) Money in China. The weighted average of triangular fuzzy numbers and its application in evaluation decision. *Operat Manage* 02:5–9. (in Chinese)

Research on BIM Application in High-Speed Railway OSC Operation Maintenance Management

Fei Meng, Tianyun Shi and Beisheng Liu

Abstract OSC is an important part of high-speed railway system. And it plays a key role in keeping the stable operation of high-speed railway. But the traditional pattern of OSC operation maintenance management is low efficiency, high cost, and heavy workload. In view of these problems, this article introduces the situation of BIM applications in China's railway industry. Based on demands of OSC operation maintenance management, this paper takes a deep research on OSC engineering information model encoding, modeling, and data integration. Then, this paper discusses four typical scenarios such as facility management, emergency management, operations maintenance training, and operation history management. The research suggests that BIM plays an important role in improving OSC operation maintenance management standard. This paper may provide some references for the research on related field.

Keywords Building information modeling (BIM) · High-speed railway Overhead contact system (OSC) · Operation maintenance management

1 Introduction

As the artery of high-speed railway, overhead contact system (OCS) plays a vital role in ensuring the stable operation of high-speed railway. All the time OCS operation maintenance is the difficult point of high-speed railway operation. Because of the long distance, the large amount of parts and the high failure rate, OCS operation maintenance is not only heavy, but also dangerous. High-speed railway OCS operation maintenance in China is divided into sections for every 60 km, each of which is a worksite. For example, there are more than 200 worksites in Beijing-Shanghai high-speed railway. Therefore, it requires a great deal of manpower and material resources. According to incomplete statistics, the cost of

F. Meng (✉) · T. Shi · B. Liu
China Academy of Railway Sciences, Hai Dian District, No.2 Daluoshu Road, Beijing, China
e-mail: 896858242@qq.com

OSC operation and maintenance accounts for more than 80% of its whole life cycle cost.

Building information modeling (BIM) is a new construction management concept, which proposed by Dr. Charlie Eastman in 1975, and was first introduced to China in 2002. In 2012, China Railway Company considers BIM as an innovative technology and brings BIM to railway industry as to improve railway whole life cycle management. Restricted by many factors, the current application of BIM in high-railway industry is just at the early stage [1–3]. BIM has characteristics such as information visualization, information completeness, information relevance, information consistency, coordination, simulation, and optimization. Therefore, it will play an important role in improving the level of railway operation and maintenance management.

2 OCS Engineering Information Modeling

2.1 OCS Engineering Information Model Coding for Operation Stage

According to “Railway Engineering Information Model Classification and Coding Standard” (1st edition), OCS information model contains more than 40 kinds of components and nearly hundred kinds of parts. Due to the long distance of the OCS, the number of parts and components is huge. Referring to the railway IFD coding system, in the respect of OCS operation, maintenance, and management, an information model coding method for railway OCS is proposed.

- (1) Consider OSC worksites as the first level of coding, and reserve three bits, i.e., 000–999;
- (2) Consider OCS components as the second level, and reserve seven bits, in which the first two bits are abbreviations of OSC component names, and the last five bits are used as component numbers;
- (3) Consider OCS parts as the third level, and reserve four bits, in which the first two are parts name abbreviations, the latter two bits are part numbers.

This coding rule optimizes the method of hierarchical division of traditional OCS components. By removing unnecessary intermediate levels to reduce the total encoding layer series and encoding bits, the encoding is flatter. The encoding usability is enhanced effectively. In order to improve the accuracy rate of encoding effectively, the combination of letters and numbers is used to distinguish names and code of OSC components and parts. Research suggests that the number of components and parts in this coding rule can meet the accuracy and usability requirements of information model.

Based on above rules, for example, the code representation of the fifth wrist arm of No. 2033 supports structure in the 125th worksite as shown in Fig. 1.

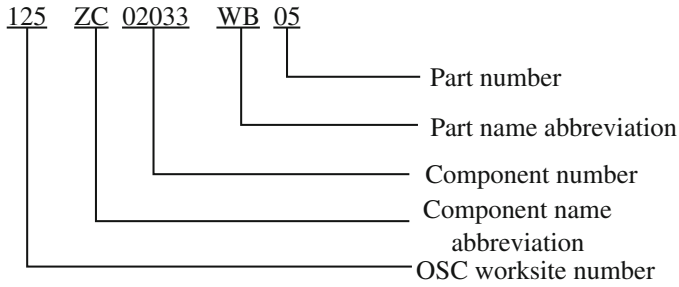


Fig. 1 The code structure of OSC engineering information model

2.2 OCS Modeling

BIM model is a concrete manifestation of BIM visualization function. It can provide OCS spatial structure information for users, as shown in Fig. 2.

According to the state and demands of high-speed railway OCS operation management, there are three methods can be used for OCS modeling:

- (1) Based on 2D drawings, equipment account, and other document information of OCS to build OCS 3D model with common modeling software, such as Autodesk Revit, Bentley ABD, and Dassault Catia.
- (2) Based on engineering entities, OCS modeling with laser scanning point cloud, and oblique photography [4].
- (3) Transformation of OCS model submitted by construction into operation maintenance model.

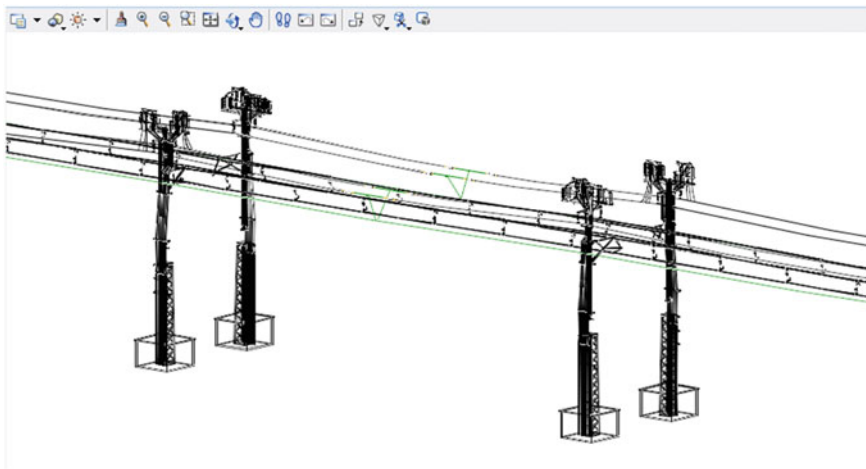


Fig. 2 High-speed railway OCS 3D model

The first and second are used for existing high-speed railway lines that constructed by traditional method. Their data is transferred and stored by 2D drawings and document information. The third is for new lines constructed with BIM, as in the construction stage, BIM management pattern was used.

In addition, precision and volume of OCS operation maintenance models need to be balanced. Due to the long distance and the large amount of parts, the volume of OCS model should be enormous in case of that high precision is required. The rendering of the model will be very time consuming on the current hardware platform, which will seriously affect the practicability and operability. However, the low precision will also limit its application. Therefore, in order to improve the display and rendering efficiency of OSC models and optimize the model application, the model needs to be transferred into a light volume format to meet the application requirements.

3 OCS Operation Maintenance BIM Information Integration

Information is the core of BIM. The operation maintenance management data of high-speed railway OCS include the basic information of OCS equipment itself, status data of equipment during operation and maintenance, and external data that affect equipment status. According to different sources, OCS operation and maintenance data are composed of six main parts: basic data for construction period, monitoring data of OCS equipment condition, monitoring data of OCS external environment, and maintenance data of OCS facility, as shown in Fig. 3.

- (1) Basic data for construction period, such as facility location, geometric attributes, structural features, accounts, materials, as well as completion of acceptance, joint commissioning and test operation data [5].
- (2) Monitoring data of OCS equipment condition. It mainly includes online monitoring data of OCS, 6C system, power SCADA system, high-speed railway RTU system, and static detection data obtained by men inspection on spot [6].
- (3) Monitoring data of OCS external environment. It mainly includes wind, rain, snow, mudslides, landslides, and other natural disasters, as well as invasion data obtained mainly by various sensors and video monitoring equipments.
- (4) Maintenance data of OCS facility, such as maintenance records, maintenance logs, frequency, effect, fault information, staff information, and so on, mainly managed by OCS maintenance dispatching system [7].

At present, these data are scattered in different systems, such as 6C system, SCADA system, OA system, integrated monitoring system, maintenance scheduling system, etc. The data are isolated and difficult to be utilized comprehensively. Through storing the above information via a dedicated channel in an OCS operation

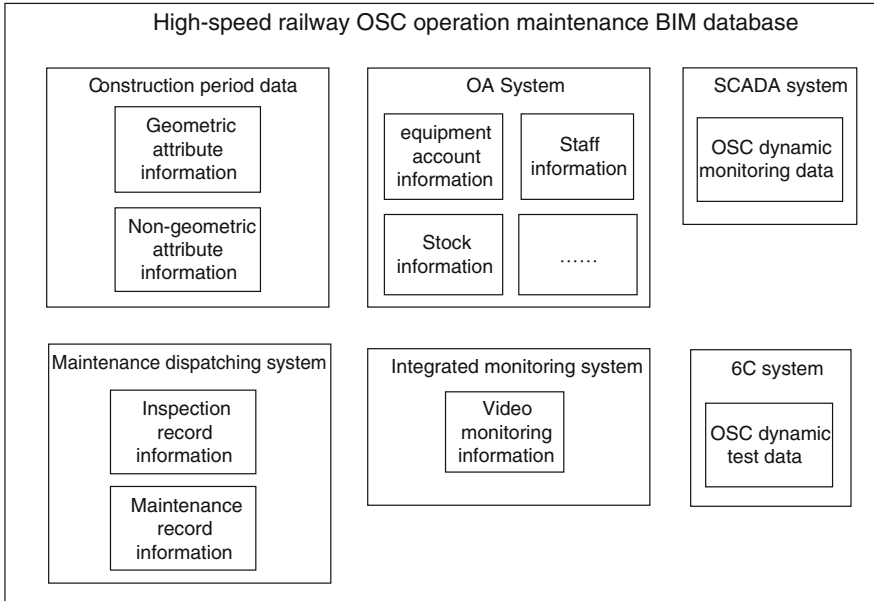


Fig. 3 High-speed railway OCS operation and maintenance BIM information structure

and maintenance BIM database, and connecting these data with the OCS operation model, unified, comprehensive utilization of management data will be achieved. It will provide decision support for the operation and maintenance management.

4 BIM Application in OCS Operation Management

BIM application in OCS operation management includes facility management, emergency management, operation and maintenance training, and operation record management, as shown in Fig. 4.

4.1 OSC Facility Management

Nowadays, 2D engineering drawings and equipment ledger are used to manage OCS facility. However, these methods have some shortages. For example, due to the scattered information, the equipment cannot be found in time and managed uniformly. Thus, RFID and QR code make the methods of solving these problems come true. During the maintenance work, staffs can get the information of equipment, such as production date, installation method, maintenance method, purchase

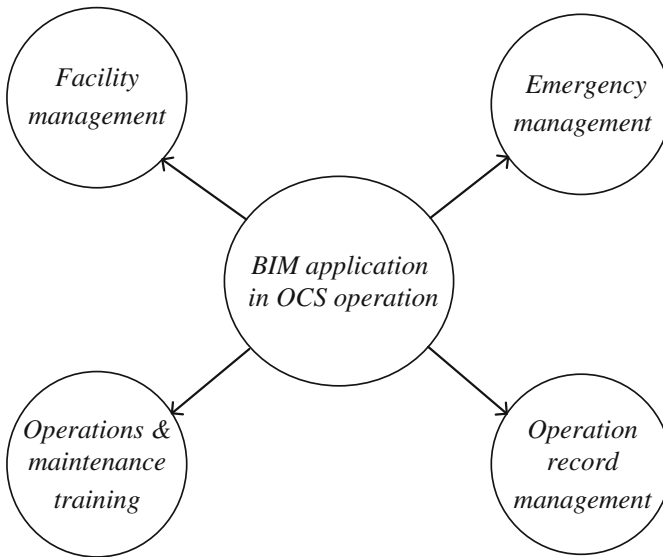


Fig. 4 BIM application in OCS operation management

information, and stock information. Moreover, in order to ensure the consistency and veracity of the information, the information should be updated timely [8, 9].

Meanwhile, the life and status of OCS equipment can be managed using these information as well. For instance, in order to guarantee the normal operation of equipment, we can maintain the parts in curing period on time and change the parts out of life cycle in time. Furthermore, scientific and reasonable maintenance plans can be set using the life and status data of equipment, and in this way, it not only can be the useless site tour and workload be reduced but also the pertinence of inspection is enhanced.

4.2 Emergency Management

Because of the short-time development of emergency management of high-speed railway in China, the study of OCS emergency management is shortage. Hence, a new technology needs to be sought. With the development of information technology, BIM, GIS, and 3D simulation used in emergency management are gradually taken seriously by many research groups. Using these novel technologies, we can simulate and analyze the natural calamities, such as fire disaster, lightning stroke, snow disaster, the terrorist attack, and the frequent fault of OCS. In this way, the rational utilization of emergency resources, the reduction of response time, the enhanced efficiency of rescue, and the support of emergency command can be achieved by means of the exercise of emergency plan.

Moreover, if there is a fault which needs to be handled in time, the management of geometric parameter in OCS, quick positioning of the fault location, fault repair procedure providing, navigation, information flow showing, and equipment logical relationship and network topology searching can be used to solve the accident.

4.3 Operations Maintenance Training

Nowadays, 2D drawing and document information is usually used in traditional operations maintenance training of OCS system. Because of the shortage in space structure, the result of this training method is not very ideal. Thus, based on some three-dimensional simulation technology, such as VR and MR, participating trainees can make some exchanges with the stereopsis of OCS system.

Moreover, with 3D simulation technology, the facility and operational details could be displayed vividly. Based on above, the workload and risk of training can be reduced, and the effect can be improved as well.

4.4 Operation History Management

The routine inspection and maintenance work of OCS are relatively hard. Because of the absence of effective supervision, employees could neglect their duties. This will cause OCS facility cannot get maintenance timely, which makes the failure rate increases.

The inspection and maintenance history can be easily traced based on BIM. When an OCS equipment fails, the inspection history and relevant workers could be rapidly traced by checking the information of the equipment on the BIM model. Clear responsibility could effectively eliminate carelessly work in routine inspection.

5 Conclusion and Prospect

According to the research and realization of code, modeling and data integration of OSC engineering information model, and exploration to the typical scenarios in the paper, it is gotten that BIM can provide help for OSC operation maintenance management achieving lightweight and refinement.

In recent years, there were few researches on BIM application in railway operation management. A lot of work remains to be done to achieve BIM mature application. Nowadays, BIM has been accepted by railway industry, and its value has received recognition. The next stage of the main emphasis should be altered to railway operation management BIM application and combined it with railway

construction management BIM application. Then, BIM application in railway whole lifecycle will come true, and BIM will certainly become the key tool that promotes the overall level of China's railway.

References

1. Meng F, Shi T, Gao G (2017) Research on current situation and promotion program of Chinese railways BIM. *Project Manage Technol* 15(2):24–28 (in Chinese)
2. Jun X, Li A, Liu H, Ye Mingzhu, Zhang J (2014) Application and risk analysis of BIM in railway. *System* 3:129–133 (in Chinese)
3. China Railway Corporation (2014) Technology development and application framework of china railway BIM. *Railway Tech Inno* 2:8–11 (in Chinese)
4. Cheng Z, Xiaojuan W (2016) Methods of railway infrastructure geo-spatial data management based on Open Source Database Engine. *Railway Comput Appl* 25(4):6–9 (in Chinese)
5. Xiangjin Li, Jun Xu, Zhao Liu (2016) BIM-based construction information management model. In: *The 4th international technical symposium on the application of bim technology in design, construction and real estate enterprise collaboration work.* (in Chinese)
6. Wang W, Guihong X, Liu J, Zhang W, XingChinese X (2015) Application and development of railway infrastructure big data. *Chinese Railways* 5:1–6 (in Chinese)
7. Ken D (2010) Railways infrastructure lifecycle management In: *Beijing: 7th world congress on high speed rail 1: 172–175*
8. Feng H, Ma X (2015) Facility management at operation and maintenance phases based on BIM and RFID technology. *Build Elect* 34(11):57–60 (in Chinese)
9. Yao B, Ma X (2015) Application of BIM and RFID in open-building whole life cycle information management. *Constr Technol* 44(10):92–95 (in Chinese)

Integrated Optimization Model on Maintenance Time Window and Train Timetabling

Lingyun Meng, Ce Mu, Xin Hong, Ran Chen, Xiaojie Luan and Tao Ma

Abstract Most of previous studies optimize maintenance time window scheduling problem under a given train schedule, leading to a relatively poor quality of maintenance time window schedule, increasing the influence on traffic assignment. In order to reduce the negative effects on maintenance schedule and improve the utilization of railway resources, we consider integrating maintenance time window scheduling and train timetabling. In this way, more reasonable maintenance time window schedule can be obtained. We propose a mixed integer programming model and in particular we focus on the characteristics of the problem, including the speed limits affected by maintenance tasks on a double-track railway line. The benefits of the proposed integrated optimization model are demonstrated by numerical experiments.

Keywords Maintenance time window · Train timetabling · Integrated optimization
Mixed integer programming · Speed limits

L. Meng (✉) · C. Mu · X. Hong · R. Chen
School of Traffic and Transportation, State Key Laboratory
of Rail Traffic Control and Safety, Beijing Jiaotong University,
Beijing 100044, China
e-mail: lym@bjtu.edu.cn

X. Luan
Section Transport Engineering & Logistics, Department of Maritime
and Transport Technology, Faculty of Mechanical, Marine & Materials Engineering,
Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

T. Ma
Data Center of China Railway Shenyang Group Company,
China Railway Shenyang Group Company, 4 Taiyuan North Street,
He-ping District 110001, Shenyang, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_87

1 Introduction

Train timetabling is a classical problem in railway operations management. A train timetable specifies a physical network route and detailed arrival time and departure time stamps for each train at passing stations. Maintenance time window scheduling is defined work-space and work-time possession for each maintenance task. The former aims at delivering railway services to customers, and the latter has the role of supporting railway services by preventing infrastructure failures. In practice, train timetable and maintenance time window are usually designed separately by different departments/planners. However, the interaction between those two is critical, as they take possession of infrastructure (utilizing capacity) competitively. Operating more trains leads to less time slots available for performing maintenance, and vice versa. In addition, the maintenance operation has a certain dangerous nature, so we must limit the speed of the trains to guarantee the safety of the workers and trains. Different maintenance operations have different conditions of the restrained speed, which will affect the train timetabling. It is hard to find a timetable with efficiently utilized capacity for trains and maintenance tasks.

The research on maintenance time window and train timetabling could be divided into three aspects: (1) The purpose of research on train timetabling without considering maintenance time window is to minimize the total traveling time of trains or to minimize the cost of transportation [1–5]; (2) The purpose of research on maintenance time window without considering train timetabling is to minimize the cost of maintenance [6–9]; and (3) The purpose of research on train timetabling with considering maintenance time window is to minimize the cost of maintenance, meanwhile maximize the transportation efficiency [10–14].

In summary, previous research on planning maintenance time window which considered the restrained speed is little. The paper adopts the mixed integer programming to build the integrated optimization model on maintenance time window and train timetable, including the speed limits affected by maintenance tasks on a double-track railway line. The model aims at exploiting potential infrastructure capacity by better deployment and coordination. It has important practical significance to improve the efficiency of the train timetable and railway transportation efficiency.

2 Problem Statement

2.1 *Speed Restriction or Reduction Rules*

According to the rules of railway safe construction, the safety management departments set different rules to test the quality of the maintenance and guarantee the safety of the train operation and the constructors. The speed restriction can be divided into two categories: one-line speed restriction and adjacent-line speed reduction.

- One-line speed restriction (OLSR) means that after the completion of maintenance tasks, the first, second, third, or other trains running on the maintenance segments should run at restricted (maximum) speeds (e.g., for the first train 30 km/h, the second train 60 km/h, the third train 80 km/h, etc.).
- Adjacent-line speed reduction (ALSRT) means that in a double-track segment, when one segment belonging to one track is contained the maintenance task, trains running on the other whole track (a sequence of cells) in the same segment (if there are) should slow down (e.g., running at 50 km/h) for safety reasons.

In this paper, the OLSRT and ALSRT represent one-line and adjacent-line speed restriction trains. The FSRT, SSRT, and TSRT, respectively, represent the first, second, and third speed restriction train of OLSRT.

2.2 The Input and Output Data

The input data of the model includes the topology of rail network, the number of trains, the maximum running speed that do not be disturbed by the maintenance time window for trains, the speed reduction or restriction that is caused by the maintenance, the trains' earliest departure time and the latest departure time at origin station, the stop scheme and the minimum dwelling time for trains, the maintenance segments, the earliest starting time and the latest starting time of maintenance tasks, and the minimum time duration of each maintenance task.

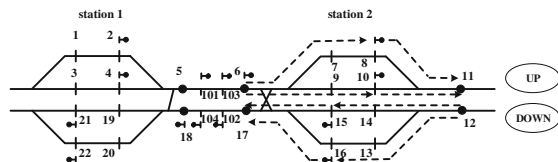
The output data of the model includes the timetable of each station, and the maintenance tasks starting time and ending time.

3 The Integrated Optimization Model for Maintenance Time Window Scheduling and Train Timetabling

3.1 Definition of the Railway Network

In order to description of the problem, we abstract the railway network as shown in Fig. 1:

Fig. 1 Sketch map of the rail networks and the train working resource



- Node: A node can be viewed as a point where tracks join together. It can be used to model a beginning/ending point of a cell or a point inside a cell. Numbers 1, 2, 3, 4 ... indicate nodes.
- Link: A link is the track between two nodes. 1–2, 3–4, 2–5, 4–5 ... indicate links.
- Cell: A cell is a group of links. A cell is allowed to be occupied by at only up to one train at any given time. Please note that once one link in the group is occupied, it means all other links in the same group are occupied at the same time, e.g., cell [1–2], [3–4], [5–101], [6–7, 6–9], [7–8], [9–10], [8–11, 10–11]...
- Block section: A block section means a sequence of cells. In a station, it refers to the track between one home signal and one departure signal or between one departure signal and one boundary point. In a segment, it refers to the track between two passing signals, e.g., block sections ([6–7, 6–9], [7–8]), ([8–11, 10–11]), ([5–101]), ([101–103]), ([103–6])...
- Track maintenance task: A set of jobs which are conducted within a time duration to ensure tracks and infrastructure are in good states for running trains. This is specified at the level of single cell involved.

3.2 *Problem Assumptions*

1. One train is represented by a virtual dot for simplicity, and train acceleration and deceleration processes are not necessary to be considered (i.e., they are infinite).
2. For a double-track railway segment between two stations, each track is modeled as a sequence of block sections, and for a single-track railway segment, the only track between two stations is modeled as one block section
3. In order to describe the influence between maintenance time window and train timetable, as well as guarantee the quality of scheduling maintenance time window, we only discuss the FSRT, SSRT, and TSRT among OLSRT and ALSRT; the others do not be considered temporarily in this paper.

3.3 *Definition of the Symbol*

All of the fundamental sets are corresponding to those sets defined in the definition of the problem. And the sets, parameters, and decision variables used in our research are described in Tables 1, 2, and 3.

Table 1 Definition of sets

Symbol	Definition
T	Set of all trains, index by t , i.e., $t \in T$
T_c	Set of trains whose routes contain the cell c
T_s	Set of trains whose routes contain the station s
C	Set of all cells, index by c , i.e., $c \in C$
C_t	Set of cells which are used by the train t
C_s	Set of cells which are constituted by arrival–departure tracks at station s
C_{mot}	Set of cells which make up the maintenance tasks mot
C_{adj}	Set of adjacent cells which the maintenance tasks mot
S	Set of stations in the rail networks, index by s , i.e., $s \in S$
MOT	Set of maintenance tasks, index by mot , i.e., $mot \in MOT$

Table 2 Definition of decision variables

Symbol	Definition
$z_{t_1,t_2,c}$	0–1 variable, equal 1 if train t_1 is scheduled earlier on cell c than train t_2 , 0 otherwise
$y_{t,c}$	0–1 variable, equal 1 if train t uses the cell c , 0 otherwise
$\sigma_{1,t,c}$	0–1 variable, equal 1 if train t is the FSRT of traveling through cell c after track maintenance task on cell c is finished, 0 otherwise
$\sigma_{2,t,c}$	0–1 variable, equal 1 if train t is the SSRT of traveling through cell c after track maintenance task on cell c is finished, 0 otherwise
$\sigma_{3,t,c}$	0–1 variable, equal 1 if train t is the TSRT of traveling through cell c after track maintenance task on cell c is finished, 0 otherwise
$\sigma_{adj,t,c}$	0–1 variable, equal 1 if train t is the adjacent train traveling through adjacent track of cell c in the same segment, 0 otherwise
$e_{t,o}^{start}, e_{t,d}^{end}$	The train t departure time at origin station o and the arrival time at the end station d
$e_{t,c}^+, e_{t,c}^-$	Entry time and exit time for train t at cell c
$e_{t,c,mot}^+, e_{t,c,mot}^-$	Entry time and exit time for train t uses the cell c which makes up the maintenance task mot during the maintenance task is proceeding
$w_{mot}^{start}, w_{mot}^{end}$	Beginning time and ending time of each maintenance task mot
$w_{mot,c}^{start}, w_{mot,c}^{end}$	Beginning time and ending time of each cell c in maintenance task mot

Table 3 Definition of parameters

Symbol	Definition
$\alpha_{t,o}^{min}, \alpha_{t,o}^{max}$	Earliest departure time and latest departure time of train t at the origin station o
$\alpha_{mot}^{min}, \alpha_{mot}^{max}$	Earliest starting time and latest starting time of maintenance task mot
d_{t,c_s}	Minimum dwelling time of train t on arrival–departure track cell c of station s
$x_{t,s}$	0–1 variable, equal 1 if train t stop at station s , 0 otherwise
$\tau_{t,c}$	Minimum running time of the train t passes the cell c

(continued)

Table 3 (continued)

Symbol	Definition
τ_{mot}^{\min}	Minimum time duration of maintenance task mot
k_1, k_2, k_3	Speed restriction ratio of the FSRT, SSRT, and TSRT which run on the maintenance cell after maintenance task
k_{adj}	Speed reduction ratio of trains running on the adjacent track in the same segment when maintenance task is being operated
l_c	Length of the trains pass the cell c
$v_{t,c}$	Velocity of the trains pass the cell c
M	One constant that is big enough, and in this paper it equals 10,000
m	One constant that is small enough, and in this paper it equals 1/10,000

3.4 Mathematical Model

3.4.1 Objection Function

The model is aimed to minimize total running time of all trains which can intuitively measure the optimizing effect.

The objection function is the following:

$$\min \sum_{t \in T} e_{t,d}^{\text{end}} - e_{t,o}^{\text{start}}$$

3.4.2 Constraints

In order to specifically describe the model, we divide the constraints into three categories, trains operation constraints, maintenance operation constraints, and train speed restriction constraints, respectively.

1. Trains operation constraints

$$\alpha_{t,o}^{\min} \leq e_{t,o}^{\text{start}} \leq \alpha_{t,o}^{\max} \quad \forall t \in T \tag{1}$$

$$\tau_{t,c} = \frac{l_c}{v_{t,c}} \quad \forall t \in T \forall c \in C_t \tag{2}$$

$$e_{t,c}^- - e_{t,c}^+ \geq \tau_{t,c} \cdot y_{t,c} + d_{t,c_s} \cdot x_{t,s} \cdot y_{t,c_s} \quad \forall t \in T \forall c, c_s \in C_t \forall s \in S \tag{3}$$

$$e_{t,c}^+ = e_{t,c-1}^- \quad \forall t \in T \forall c \in C_t \tag{4}$$

$$e_{t,c}^- = e_{t,c+1}^+ \quad \forall t \in T \forall c \in C_t \quad (5)$$

$$\sum_{c \in C} y_{t,c} = 1 \quad \forall t \in T \quad (6)$$

$$e_{t_1,c}^+ - e_{t_2,c}^- \geq M(z_{t_2,t_1,c} - 1) \quad \forall t_1, t_2 \in T_c \cap t_1 \neq t_2 \forall c \in C_t \quad (7)$$

$$e_{t_1,c}^- - e_{t_2,c}^+ \geq M(y_{t_1,c} + y_{t_2,c} + z_{t_1,t_2,c} - 3) \quad \forall t_1, t_2 \in T_c \cap t_1 \neq t_2 \forall c \in C_s \quad (8)$$

$$\frac{y_{t_1,c} + y_{t_2,c} - 1}{2} \leq z_{t_1,t_2,c} + z_{t_2,t_1,c} \leq \frac{y_{t_1,c} + y_{t_2,c}}{2} \quad \forall t_1, t_2 \in T_c \cap t_1 \neq t_2 \forall c \in C_s \quad (9)$$

Constraint (1) ensures that trains' departure is within given time windows at origin stations; constraints (2)–(3) ensure that trains running time in cell c should not exceed the minimum running time; constraints (4)–(5) ensure the spatio-temporal continuity during the whole trip of trains; constraint (6) ensures that trains must only choose one and only route during the whole trip; constraint (7) ensures that one train occupies one cell c at any time; constraints (8)–(9) ensure that if two or more trains have to stop at the same station, they not allow occupying the same arrival–departure tracks at the same time.

2. Maintenance operation constraints

The method of integrating optimization model of maintenance time window and train timetable considers each of the maintenance tasks as one “virtual train”. The “virtual train” departs from the origin station means the maintenance tasks begin, and “virtual train” arrivals at the destination means the maintenance tasks end. The duration of maintenance tasks is the traveling time or occupying time on maintenance tracks. The method not only guarantees to plan the maintenance tasks and trains timetable globally but also achieve the purpose of judging the trains whose speed must be restricted or redacted.

$$\alpha_{mot}^{\min} \leq w_{mot}^{\text{start}} \leq \alpha_{mot}^{\max} \quad \forall \text{mot} \in \text{MOT} \forall c \in C_{\text{mot}} \quad (10)$$

$$w_{mot}^{\text{end}} - w_{mot}^{\text{start}} \geq \tau_{mot}^{\min} \quad \forall \text{mot} \in \text{MOT} \forall c \in C_{\text{mot}} \quad (11)$$

$$w_{\text{mot},c_1}^{\text{start}} \leq w_{\text{mot},c_2}^{\text{end}} \quad \forall \text{mot} \in \text{MOT} \forall c_1, c_2 \in C_{\text{mot}} \quad (12)$$

$$w_{\text{mot},c_1}^{\text{end}} \geq w_{\text{mot},c_2}^{\text{start}} \quad \forall \text{mot} \in \text{MOT} \forall c_1, c_2 \in C_{\text{mot}} \quad (13)$$

$$e_{t,c,\text{mot}}^+ - w_{\text{mot},c}^{\text{end}} \geq M(z_{\text{mot},t,c} - 1) \quad \forall \text{mot} \in \text{MOT} \forall t \in T_c \forall c \in C_{\text{mot}} \quad (14)$$

$$w_{\text{mot},c}^{\text{start}} - e_{t,c,\text{mot}}^- \geq M(z_{t,\text{mot},c} - 1) \quad \forall \text{mot} \in \text{MOT} \forall t \in T_c \forall c \in C_{\text{mot}} \quad (15)$$

$$z_{\text{mot},t,c} + z_{t,\text{mot},c} = 1 \quad \forall \text{mot} \in \text{MOT} \forall t \in T_c \forall c \in C_{\text{mot}} \tag{16}$$

Constraint (10) ensures that maintenance tasks can be operated within the prescribed period, whose purpose is to reduce its influence on train timetable; constraint (11) ensures that the duration of maintenance to be no less than minimum maintenance time which is 120 min in traditional railway and 240 min in high-speed railway at least; constraints (12)–(13) ensure that maintenance tasks could be operated continuously, whose purpose is to reduce the influence on train operation; constraints (14)–(16) ensure that trains must not occupy maintenance tracks when maintenance tasks are being operated.

3. Speed restriction constraints

We introduce the principle of identifying trains' sequence of occupying the maintenance segments. We assume that there are five trains, and they occupy the maintenance segments according to the order of Fig. 2. Therefore, we know that $T_c = \{t_1, t_2, t_3, t_4, t_5\}$, $\forall c \in C_{\text{mot}}$. We introduce variable φ and $\varphi = \sum_{u \in T_c} z_{\text{mot},u,c} - \sum_{u \in T_c} z_{t,u,c}$, $\forall t, u \in T_c \cap t \neq u$, $\forall c \in C_{\text{mot}}$, $\forall \text{mot} \in \text{MOT}$. In Fig. 2, it is easy to know that the FSRT is t_3 , the SSRT is t_2 , and the TSRT is t_1 . We can also know that $t = t_1$, $\varphi = 3$; $t = t_2$, and $\varphi = 2$. And by this analogy, $t = t_3, t_4, t_5$, and φ equals 1, 0, -1.

Therefore, $\forall c \in C_{\text{mot}}$, if $t = t_3$, then $\varphi = 1$, $\sigma_{1,t,c} = 1$; t equals others, $\varphi \neq 1$, $\sigma_{1,t,c} = 0$. Similarly, $\forall c \in C_{\text{mot}}$, if $t = t_2$, then $\varphi = 2$, $\sigma_{2,t,c} = 1$; t equals others, $\varphi \neq 2$, $\sigma_{2,t,c} = 0$. $\forall c \in C_{\text{mot}}$, if $t = t_1$, then $\varphi = 3$, $\sigma_{3,t,c} = 1$; t equals others, $\varphi \neq 3$, $\sigma_{3,t,c} = 0$. In this way, we can identify trains' sequence of occupying the maintenance segments.

The specific constraints are as follows:

$$\sigma_{1,t,c} \geq z_{\text{mot},t,c} \cdot m \cdot (1.5 - \varphi) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \forall \text{mot} \in \text{MOT} \tag{17}$$

$$\sigma_{1,t,c} \leq z_{\text{mot},t,c} - z_{\text{mot},t,c} \cdot m \cdot (\varphi - 1) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \forall \text{mot} \in \text{MOT} \tag{18}$$

$$\mu_{2,t,c} \cdot M \geq \varphi - 1.5 \quad \forall t \in T_c \forall c \in C_{\text{mot}} \tag{19}$$

$$(1 - \mu_{2,t,c}) \cdot M \geq -(\varphi - 1.5) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \tag{20}$$

$$\sigma_{2,t,c} \geq \mu_{2,t,c} \cdot m \cdot (0.8 - (\varphi - 1.5)) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \tag{21}$$

$$\sigma_{2,t,c} \leq \mu_{2,t,c} - \mu_{t,c} \cdot m \cdot ((\varphi - 1.5) - 0.5) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \tag{22}$$

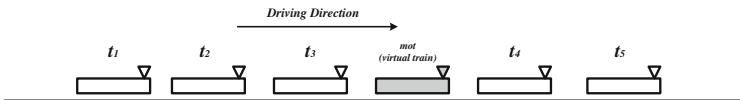


Fig. 2 Sketch map of the trains pass the maintenance segment

$$\mu_{3,t,c} \cdot M \geq \varphi - 2.5 \quad \forall t \in T_c \forall c \in C_{\text{mot}} \quad (23)$$

$$(1 - \mu_{3,t,c}) \cdot M \geq -(\varphi - 2.5) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \quad (24)$$

$$\sigma_{3,t,c} \geq \mu_{3,t,c} \cdot m \cdot (0.8 - (\varphi - 2.5)) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \quad (25)$$

$$\sigma_{3,t,c} \leq \mu_{3,t,c} - \mu_{3,t,c} \cdot m \cdot ((\varphi - 2.5) - 0.5) \quad \forall t \in T_c \forall c \in C_{\text{mot}} \quad (26)$$

$$e_{t,c,\text{mot}}^- - e_{t,c,\text{mot}}^+ \geq \sigma_{1,t,c} \cdot k_1 \cdot \tau_{t,c} + (1 - \sigma_{1,t,c}) \cdot \tau_{t,c} \quad \forall t \in T_c \forall c \in C_{\text{mot}} \forall \text{mot} \in \text{MOT} \quad (27)$$

$$e_{t,c,\text{mot}}^- - e_{t,c,\text{mot}}^+ \geq \sigma_{2,t,c} \cdot k_2 \cdot \tau_{t,c} + (1 - \sigma_{2,t,c}) \cdot \tau_{t,c} \quad \forall t \in T_c \forall c \in C_{\text{mot}} \forall \text{mot} \in \text{MOT} \quad (28)$$

$$e_{t,c,\text{mot}}^- - e_{t,c,\text{mot}}^+ \geq \sigma_{3,t,c} \cdot k_3 \cdot \tau_{t,c} + (1 - \sigma_{3,t,c}) \cdot \tau_{t,c} \quad \forall t \in T_c \forall c \in C_{\text{mot}} \forall \text{mot} \in \text{MOT} \quad (29)$$

$$\sigma_{\text{adj},t,c} \geq m \cdot (w_{t,c}^{\text{end}} - e_{t,c}^-) \cdot (e_{t,c}^+ - w_{t,c}^{\text{start}}) \quad \forall t \in T_c \forall c \in C_{\text{adj}} \quad (30)$$

$$\sigma_{\text{adj},t,c} \leq 1 + m \cdot (w_{t,c}^{\text{end}} - e_{t,c}^-) \cdot (e_{t,c}^+ - w_{t,c}^{\text{start}}) \quad \forall t \in T_c \forall c \in C_{\text{adj}} \quad (31)$$

$$e_{t,c,\text{mot}}^- - e_{t,c,\text{mot}}^+ \geq \sigma_{\text{adj},t,c} \cdot k_{\text{adj}} \cdot \tau_{t,c} + (1 - \sigma_{\text{adj},t,c}) \cdot \tau_{t,c} \quad \forall t \in T_c \forall c \in C_{\text{adj}} \quad (32)$$

Constraints (17)–(18) are used to determine the FSRT. Constraints (19)–(22) are used to determine the SSRT through introducing auxiliary 0–1 variable $\mu_{2,t,c}$. Constraints (23)–(26) are used to determine the TSRT through introducing auxiliary 0–1 variable $\mu_{3,t,c}$. Constraints (26)–(29) ensure that FSRT, SSRT, and TSRT satisfy the requirement of speed restriction. Constraints (30)–(31) are used to determine the ALSRT, in which we consider the only one situation that $e_{t,c}^+ > w_{t,c}^{\text{start}}$ and $w_{t,c}^{\text{end}} < e_{t,c}^-$; the others are not to be taken into account. Constraint (32) ensures that ALSRT satisfies the requirement of speed reduction.

4. Model transformation

Constraints (30)–(31) are nonlinear, which increase the difficulty of the model. Therefore, we need to transform the constraints to reduce the difficulty and improve the rate of solving. In this article, we judge the ALSRT by the way of introducing the 0–1 auxiliary variables $\mu_{\text{adj}1,t,c}$, $\mu_{\text{adj}2,t,c}$, which are used to judge the occupying maintenance segments successively between trains and maintenance tasks (i.e., virtual trains).

$$\mu_{\text{adj}1,t,c} \cdot M > e_{t,c}^+ - w_{t,c}^{\text{start}} \quad \forall t \in T_c \forall c \in C_{\text{adj}} \quad (33)$$

$$(1 - \mu_{adj1,t,c}) \cdot M \geq -(e_{t,c}^+ - w_{t,c}^{start}) \quad \forall t \in T_c \forall c \in C_{adj} \tag{34}$$

$$\mu_{adj2,t,c} \cdot M > w_{t,c}^{end} - e_{t,c}^- \quad \forall t \in T_c \forall c \in C_{adj} \tag{35}$$

$$(1 - \mu_{adj2,t,c}) \cdot M \geq -(w_{t,c}^{end} - e_{t,c}^-) \quad \forall t \in T_c \forall c \in C_{adj} \tag{36}$$

$$\sigma_{adj,t,c} = \mu_{adj1,t,c} \cdot (\mu_{adj1,t,c} + \mu_{adj2,t,c} - 1) \quad \forall t \in T_c \forall c \in C_{adj} \tag{37}$$

Constraints (33)–(37) ensure that when $e_{t,c}^+ > w_{t,c}^{start}$ and $w_{t,c}^{end} < e_{t,c}^-$, the $\sigma_{adj,t,c} = 1$, and the other situations, $\sigma_{adj,t,c} = 0$.

4 Numerical Case

We utilize the business solution software IBM ILOG CPLEX 12.3 to test the model. The runtime environment of the CPLEX is a computer with Intel Core i5-2430 M, CPU (4) 2.40 GHz, 4 GB RAM.

We choose a certain of railway line to verify the feasibility of the model. The railway line contains six stations, and the circuit diagram is shown in Fig. 3. In the planning time, there are eight trains whose speed ratings are same. The normal speed limit is 150 km/h. Four trains are in the up direction. Three trains are in the down direction. They all need to stop at stations B, C, D, and E. The minimum stop times are 5, 8, 7, and 4, respectively. The another one train is from station D to station A, which need to stop for 2 min at station C, and stop for 5 min at station B. The segments which are between station B and station C of up direction, and between station E and station D of down direction need to be maintained. The maintenance starting time windows are [0:20, 0:50] and [0:00, 0:50], respectively. And the duration of the maintenance tasks is 120 min at least. In order to guarantee the safety of train operation and examine the quality of maintenance tasks, the trains need to satisfy the speed restricted requirements which are $v_{FSRT} \leq 60\text{km/h}$, $v_{SSRT} \leq 100\text{km/h}$, $v_{TSRT} \leq 120\text{km/h}$, and $v_{ALSRT} \leq 60\text{km/h}$.

The model spends 33.67 s to get the optimal solution by the CPLEX. The specific timetable is shown in Fig. 4. Based on the result, the B–C maintenance tasks are operated on 0:49–2:49 and the E–D maintenance tasks are operated on 0:00–2:00 (the shaded part in Fig. 4). The speeds of trains 2, 3, 4, 5, 6, 7, and 8 need to be restricted.

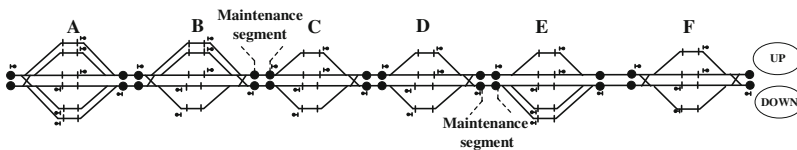
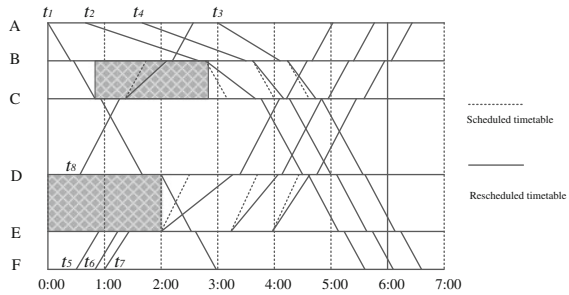


Fig. 3 Sketch map of a line’s stations and maintenance segments

Fig. 4 The train timetable with the maintenance time window



5 Conclusion

In this paper, first, the paper utilizes a method that considers the maintenance tasks as “virtual trains”; then, we take advantage to the relationship between trains and “virtual trains”, to judge the FSRT, SSRT, and TSRT of OLSRT and ALSRT. Afterward, we adopt mixed integer programming to build the integrated optimization on maintenance time window and train timetabling. In view of the hard constraints, we introduce the auxiliary variables to reduce the difficulty and improve the rate of solving. The benefit of the proposed integrated optimization model is demonstrated by numerical experiments.

Acknowledgments This work is partially supported by the Fundamental Research Funds for the Central Universities (No.2017JBM029) and National Natural Science Foundation of China (No.71571012).

References

1. Szpigel B (1973) Optimal train scheduling on a single line railway. *Operat Res* 72:344–351
2. Jovanović D, Harker PT (1991) Tactical scheduling of rail operations: the SCAN I System. *Transp Sci* 25(25):46–64
3. Fischer F, Helmberg C, Janben J et al (2008) Towards solving very large scale train timetabling problems by Lagrangian relaxation. *Atmos*
4. Meng L, Zhou X et al (2014) Simultaneous train rerouting and rescheduling on an N-track network: a model reformulation with network-based cumulative flow variables. *Transp Res Part B Methodolo* 67(3):208–234
5. Goverde RMP, Bešinović N, Binder A et al (2016) A three-level framework for performance-based railway timetabling. *Transp Res Part C Emerg Technol* 67:62–83
6. Budai G, Huisman D, Dekker R (2006) Scheduling preventive railway maintenance activities. *J Oper Res Soc* 57(9):4171–4176
7. Vale C, Ribeiro I, Rui C (2012) Integer programming to optimize tamping in railway tracks as preventive maintenance. *J Transport Eng* 138(1):123–131
8. Fokkert ZD, Verhoeven JHM (2007) The Netherlands schedules track maintenance to improve track workers’ safety. *Interfaces* 37(2):133–142
9. Higgins A (1998) Scheduling of railway track maintenance activities and crews. *J Oper Res Soc* 49(10):1026–1033

10. Cheung BSN, Chow KP, Hui LCK et al (1999) Railway track possession assignment using constraint satisfaction. *Eng Appl Artif Intell* 12(12):599–611
11. Peng F, Kang S, Li X et al (2011) A heuristic approach to the railroad track maintenance scheuling problem. *Comput-Aided Civil Infrastruct Eng* 26(2):129–145
12. Borraz-Sánchez C, Klabjan D (2012) Strategic gang scheduling for railroad maintenance. *Maintenance*
13. Bahramian Z, Bagheri M (2015) A simulation-based optimization approach for passenger train timetabling with periodic track maintenance and stops for praying. *J Modern Transport* 23(2):148–157
14. Xiaojie L, Jianrui M, Lingyun M et al (2017) Integrated optimization on train scheduling and preventive maintenance time slots planning. *Transport ResPart C Emerg Technol* 80(1): 329–359

Controllability of Urban Rail Transit Network

Lu Zeng, Yong Qin, Jun Liu and Li Wang

Abstract Due to the rapid development of complex network research, the controllability research of different kinds of large-scale actual network has become a research hotspot nowadays. The purpose of this study is to reveal the controllability of urban rail transit network. First, it built the urban rail transit network topology by analyzing the characteristics of urban rail transit network structure. Then, a controllability model of subway is set up based on control theory of state equation. In the model, we defined the departure train as driver node. Whether the system is controllable is determined by driver nodes. At last, we analyzed urban rail transit network structure and compared the controllability of the city of Beijing, Guangzhou, and Dongjing. The result shows that the existing subway network is uncontrollable because of the complex structure of subway. We can increase the number of driver nodes in order to make it controllable.

Keywords Urban rail transit network · Controllability · Drive node
State space

1 Introduction

Urban rail transit becomes the backbone of the world's urban transport with the characteristics of large, fast, and safe. The physical environment of urban rail transit is special, technical composition is complex, system coupling is strong, and so on.

L. Zeng · J. Liu · L. Wang

School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

L. Zeng

College of Applied Science Jiangxi University of Science and Technology,
Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi Province, China

Y. Qin (✉)

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

e-mail: qinyong2146@126.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_88

873

The impact of the natural environment and other characteristics presents a serious challenge to the safety of rail transit. Therefore, the research on the controllability of urban rail transit network is of great significance to the transportation organization and planning of urban rail transit.

The study of complex network of urban rail transit started relatively late. Angeloudis et al. [1–3] applied the complex network to the subway network. The studies focus on the robustness and reliability of the subway network. But the control theory is only limited to system control, not applied to the network. The control of large-scale system is controlled from early flocking control to traction control. Pinning control is the representative of complex network control. Wang et al. [4, 5] combined the idea of pinning control with flock control. Chen [6] studied the pinning control application of complex network, controllability of dynamic complex networks and “network of network” modeling, and so on. Liu [7] studied the controllability problem of directed networks. The judgment of the state space equation in the control theory is applied to the controllability of the network at the same time. Based on Liu’s research, Yan [8] analyzed the relationship between controllability and energy consumption of different types of networks by the angle of energy consumption. Nepusz [9] considered the dynamics of the edge of the network and converted the network to edge-based model. Ferrarini [10] extended the concept of driver node to drive edge and pointed out that the node is time-varying and the edges are usually constant. Jia [11] calculated the control ability of nodes by random sampling method. It is found that the control ability of the node is determined by the entry degree of the node; the greater the degree of the node is, the harder the driver node become. Lombardi [12] applied controllable matrix to the network. Meng [13] studied the controllability of the railway train service network. It defines the drive node based on the immune transmission and cascade failures by train service network. An improved LB model is constructed, which proposed a dual chart of the train service network.

The research of network controllability is mostly based on the theoretical research of complex networks. Until recently, there are few examples of network controllability. Most of the existing research is on complex network features. Even analyzing different types of network also returns to the indicators of network.

The controllability model of urban rail transit network is constructed by analyzing the urban rail transit network. The controllability and the characteristics of urban rail transit network are analyzed by examples. It expects to change the input signal so that the state of the network achieves our desired state. The entire urban rail transit network can be within the scope of our control.

2 Network Control Theory

2.1 System Controllable Model

We hope that we can make the network system achieves our desired state by inputting the signal for the system and adjusts the performance of the network

system appropriately, in order to achieve full control of the network. But there are some difficulties, such as the complex network system size is very large, the traditional control theory is difficult to apply directly, the calculation is complex, and the computational complexity is also very high. Most complex network systems in reality are nonlinear. It is difficult to control the whole network because of destabilizing factors.

If there exists a segmented continuous input $u(t)$, it can make the system from an initial state $x(t_0)$ to any specified terminal status $x(t_f)$ in finite time interval $[t_0, t_f]$. It is said that the state is controllable. If all the state of system is controllable, the system is completely controllable.

The input–output model of linear time-invariant system can be represented as

$$\dot{x} = Ax(t) + Bu(t), \quad (1)$$

where the vector $x(t) = (x_1(t), \dots, x_N(t))^T$ captures the state of a system of N nodes at time t ;

The input signal $u(t) = (u_1(t), \dots, u_M(t))^T$, $M \leq N$;

A is the state matrix, $A \in R^{N \times N}$;

B is the input matrix, $B \in R^{N \times M}$;

The linear time-invariant system $\sum(A, B)$ is said to be controllable if and only if the $N \times NM$ controllability matrix has full rank. It can be shown as

$$C = (B, AB, A^2B, \dots, A^{N-1}B) \quad (2)$$

Urban rail transit system is a typical complex system. The complexity of the system structure and equipment facilities enhance the system's instability. If it can be controlled by changing the input signal, the efficiency of emergency measures can be improved.

2.2 Control Theory of Complex Network Structure

In the past, people mainly applied the theory of complex networks to different types of object entities. The nature of the complex network reveals some of the inherent characteristics of the object. In recent years, more and more attentions have been paid to the control of complex network system. The research of complex network control has become the hot topic in today's research. The ultimate goal of understanding the various attributes of entities is to be able to control them in accordance with established goals. The control theory indicates that the appropriate input signal can affect the characteristics of the dynamic system, so that the output of the system reaches the final state of expectation.

In complex network control, Liu [7] proposed a control complex networks method of minimizing the node N_D . The simulation results show that the number of

driver nodes is mainly determined by the network degree distribution. Sparse heterogeneous networks tend to be harder to control. Dense homogeneous network requires smaller number of driver nodes to be controlled. Sparsity means that the average degree of network is much smaller than the maximum possible connectivity N (number of network nodes), such as urban rail transit networks. Heterogeneous networks not only consider the topology but also the information of nodes and edges in the network, for example, social network, blog network, gene regulation network, power grids, language concept network, etc.

Liu [7] combined structure control with graph theory. It can determine the network whether is controllable by the method of edge and node matching. In order to determine whether a network structure is controllable, it must meet the following three conditions:

- (1) Linear control system (A, B) is structure controllable;
- (2) Directed graph (A, B) does not contain unreachable nodes and extended structure; and
- (3) Directed graph (A, B) generated by the palm structure.

If the system is controllable, the input signal can reach all paths. How to convert an uncontrollable network into a controllable network and determine the minimum input signal changes to a matching problem.

We can propose a node matching method of binary graph. Matching edge means that any two directed edges do not have common vertex (head or tail node). The matching node is the head node of matching edge. The method of edge and node matching can be represented as shown in Fig. 1.

If all the edges of the whole network do not share head node or tail node, the whole network is maximum match, as $N_D = 1$. If not maximum match, the value of the drive node N_D is equal to the number of unmatched node, as shown in Eq. 3. For a unmatched node, the given input signal can reach all match points to control

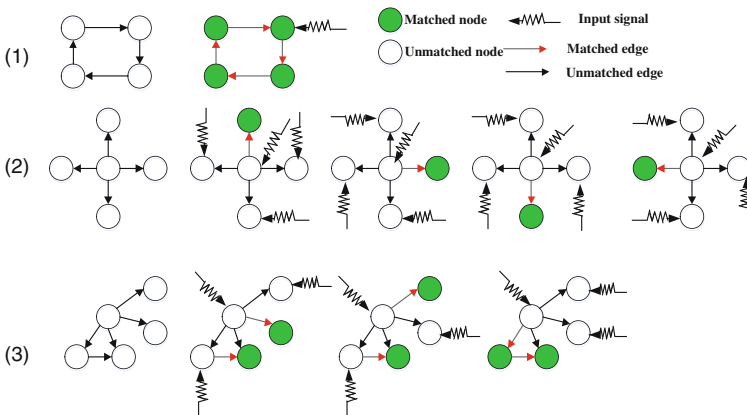


Fig. 1 Example of edge and node matching

the entire network. The minimum drive nodes can be obtained by the maximum matching method.

$$N_D = \max\{1, N_{\text{unmatched}}\} \tag{3}$$

The maximum matching algorithm is usually based on improved Hopcroft and Karp algorithm. The main idea of the algorithm is converting the network to bipartite graph. It selects a matching edge L , and then finds its augmenting path S . If the augmenting path exists, it can be judged that there is a greater match L' than L . This method is repeated until the augmenting path does not exist; it can judge that the maximum match in the whole network is found.

3 Example of Controllability Network

Beijing subway has 18 operating lines (17 subway lines and 1 airport line) until 2015. It has 319 stations (including the transfer stations of repeated calculation).

The node corresponds to the station, and link corresponds to the rail transit line of topology of urban rail transit network. According to Eq. (1) of urban rail transit network, we can define the parameters as follows:

In the element a_{ij} of the state matrix A , i is the departure station, j is the destination station, i and j are the adjacent stations;

If two adjacent stations have orbits, $a_{ij} = 1$, otherwise $a_{ij} = 0$.

The element b_k of the input matrix B is the number k of departure station. If the station has departure train, the control input factor is 1, otherwise 0.

We apply the theory to the entire Beijing subway network with both direction conditions. In order to calculate easily, we build a basic database as shown in Fig. 2.

Each station of the network is numbered by line. Such as line 1: Pingguoyuan 0101, Gucheng 0102, Bajiaoyouleyuan 0103, Babaoshan 0104, Yuquanlu 0105, Wukesong 0106, and so on. Line 2: Xizhimen 0201, Chegongzhuang 0202, Fuchengmen 0203, and so on. The first two digits are the line number. The last two digits are the station number, as shown in Table 1. According to the subway topology, we can know that the matrix A is sparse.

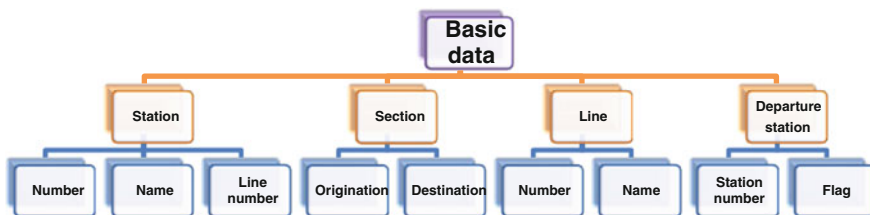


Fig. 2 Basic database

Table 1 Information of Beijing urban railway network

Line	Number	Number of stations	Station number
Line 1	01	23	0101–0123
Line 2	02	18	0201–0218
Line 4/Daxing	04	36	0401–0435
Line 5	05	23	0501–0523
Line 6	06	26	0601–0626
Line 7	07	19	0701–0719
Line 8	08	17	0801–0817
Line 9	09	13	0901–0913
Line 10	10	45	1001–1045
Line 13	13	16	1301–1316
Line 14	14	17	1401–1417
Line 15	15	19	1501–1519
Batong	16	13	1601–1613
Fangshan	17	11	1701–1711
Changping	18	7	1801–1807
Yizhuang	19	13	1901–1913
Airport Line	20	4	2001–2004

The edges between the node are directed. The direction is the same as train. The whole network is two-direction network. The state matrix A is 319×319 composed of 0 and 1 which is based on basic data by adjacency matrix-analytic method.

$$A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & \cdots & 1 & 0 \end{bmatrix} \tag{4}$$

The matrix B is an input matrix which is composed of departure stations. The so-called departure station is the empty train from the depot to the station. The departure station of Beijing urban railway network can be obtained by line traffic data and timetable as shown in Table 2. The matrix B is the matrix of 319×51 composed of 0 and 1.

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & & 1 & 0 \\ \vdots & \vdots & \vdots & & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \tag{5}$$

Table 2 Information of train departure station of Beijing urban railway network

Line	Departure station
Line 1	Pinguoyuan Gucheng Sihui Sihuidong Fuxingmen
Batong	Sihui Tuqiao
Line 2	Xizhimen Jishuitan Beijingzhan
Line 4/Daxing	Gongyixiqiao Anheqiaobei Tiangongyuan
Line 5	Tiantongyuanbei Songjiazhuang
Line 6	Haidianwuluju Caofang Lucheng
Line 7	Beijingxizhan Jiaohuachang
Line 8	Zhuxinzhuang Pingxifu Huilongguandajie Nanluoguxiang
Line 9	Guojiatushuguan Guogongzhuang
Line 10	Bagou Chedaogou Songjiazhuang Chengshousi Suzhoujie Cishousi
Line 13	Huoying Dongzhimen Xizhimen
Line 14	Shangezhuang Jintailu Zhuangguozhuang Xiju
Line 15	Maquanying Fengbo Qinghuadongluxikou
Fangshan	Guogongzhuang Suzhuang
Changping	Nanshao Zhuxinzhuang Xierqi
Yizhuang	Songjiazhuang Ciqu
Airport line	Dongzhimen Terminal 2

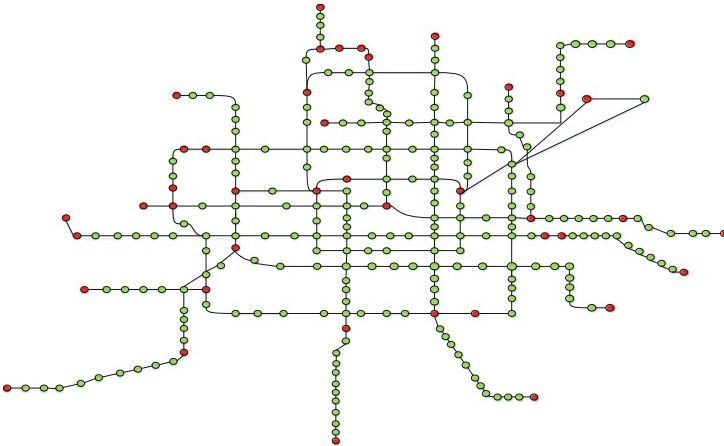


Fig. 3 The driver nodes of Beijing urban railway network

We established the drive node network topology of Beijing urban rail transit according to the actual operation in 2015, as shown in Fig. 3. The red circles in the figure are drive node stations, representing the departure stations. The two-direction arrow represents that there are trains in two directions between the stations. The calculated result is $rank(C) = 68$. It means the network is uncontrollable.

Besides that, we do have the same controllability verification of Guangzhou subway. Guangzhou urban rail transit has 160 stations until 2016 (including transfer stations of the repeated calculation, remove unopened stations and APM lines). The control matrix C can be obtained by analyzing the basic data. The result is $rank(C) = 37$. The entire Guangzhou subway network is also uncontrollable.

Subway network is different from other networks. It has a large number of transfer stations. The transfer node is at least crossed by two lines, and even more lines are interlaced together. According to the maximum matching theory, there are more nodes that do not match around the transfer station. The number of driver nodes is not enough. So the network is uncontrollable.

4 Analysis of Controllable Characteristics of Typical Subway Network

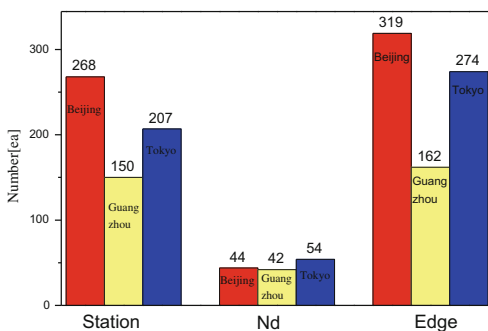
Beijing urban rail transit operation scale ranks first in the world. Annual passenger volume has reached 3.25 billion until 2015. Tokyo is the first city in Asia which has subway. It has 13 lines and 287 stations in 2016. The operating mileage is 304.1 km, and the average daily transport capacity is 8.5 million people. Guangzhou urban rail transit is the fourth city of China which has the subway. Guangzhou subway has 9 operation lines and 172 stations (including transfer stations of the repeated calculation).

The Tokyo subway is more complex than Beijing by topology and transport organization mode. Beijing is relatively standard, and Guangzhou subway is simple. The index contrast of urban rail transit of three cities is shown in Fig. 4.

The network topology of Tokyo subway presents radial ring structure. It combined loop and radiation structure in order to avoid excessive concentration of passenger flow and enhance the connectivity of the entire network. In addition, Tokyo subway transport organization mixed varieties of operation modes.

The network topology of Beijing urban rail transit is based on the distribution of city planning. The center of subway network structure is Tiananmen Square.

Fig. 4 Index contrast of urban rail transit of three cities



The loop is also more square. The city line is chessboard shape, and the end of the line is radial. The speed of urban rail transit trains in China is relatively monotonous. There are fewer opportunities to cross lines.

Guangzhou subway network topology mainly composed of the line of “traffic guidance” and “planning guide” type. There is no loop line. The network structure is both centripetal and interwoven.

Most of the large networks are often sparse. According to the definition of sparse network, we can see that urban rail transit network belongs to sparse network. The driver nodes and nodes contrast of urban rail transit of three cities are shown in Fig. 5.

From Fig. 5, Guangzhou has the highest driver node ratio of all nodes, followed by Tokyo and Beijing. Compared to Guangzhou, Beijing has more drive nodes. But the driving nodes accounted for the proportion of nodes are the least. Guangzhou has the least number of driver nodes. But it has the largest proportion of the drive node. Because Guangzhou Metro has no loop line, it is a radiation state, less transfer stations, node degrees, and clustering coefficients are relatively low. So it is relatively sparse. The more sparse the network is, the more driver nodes are needed.

Tokyo has fewer nodes than Beijing; the proportion of drive nodes is greater than Beijing. This is because Tokyo’s network structure is more complex than Beijing. There are a variety of train operation plan and transport organization. It is more flexible because of roundabout line. Beijing has more stations than Tokyo but the plan is more standardized. Degree distribution contrast of urban rail transit of three cities is shown in Fig. 6.

In Fig. 6, Beijing has the largest number of stations. But its line planning is more standardized. The degrees of drive nodes and transfer stations are more average. The degree distribution of the driving nodes presents a power law distribution. The above analysis can be drawn that the number of driver nodes is related to the degree distribution of the network.

Fig. 5 Driver nodes and nodes contrast of urban rail transit of three cities

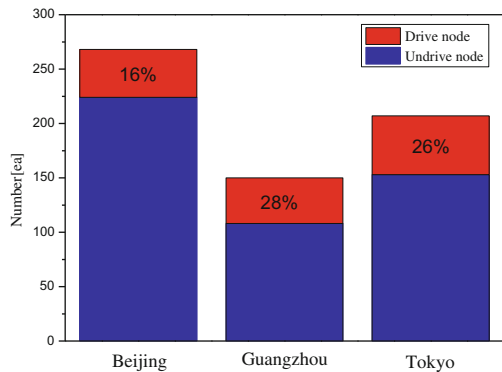
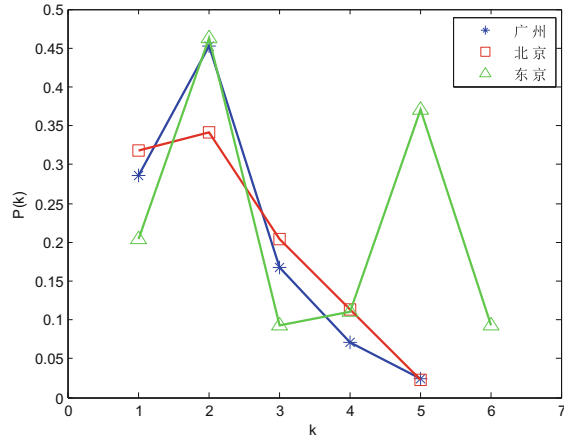


Fig. 6 Degree distribution contrast of urban rail transit of three cities



5 Conclusion

This paper mainly studies the controllability and the characteristics of urban rail transit network. The topology of urban rail transit is constructed. According to the structure controllable model, network controllability is verified by Beijing urban rail transit example. The structural controllable characteristics are analyzed by the characteristics of the complex network. Through the analysis of the controllability of urban rail transit network, the results show that the physical subway network is uncontrollable under the current conditions. Compare to Guangzhou and Tokyo, Beijing's driving nodes account for a lower percentage of all nodes by analyzing the controllable characteristics of typical subway network.

Acknowledgements This work was supported in part by the National Key Research and Development Program of China (2016YFB1200401), National Natural Science Foundation of China (61374157, 71701010, 61763016), National Key R&D Program of China No.2017YFB1201105-12 and the Science and Technology Project of the Education Department of Jiangxi Province (GJJ151524).

References

1. Angeloudis P, Fisk D (2006) Large subway systems as complex networks. *Phys A* 367:553–558
2. Santiago del Río PM, Hernández JA, Aracil J et al (2010) A reliability analysis of double-ring topologies with dual attachment using P-cycles for optical metro networks. *Comput Netw* 54(8):1328–1341
3. De-Los-Santos A, Laporte G, Mesa JA et al (2012) Evaluating passenger robustness in a rail transit network. *Transp Res Part C* 20(1):34–46
4. Wang XF, Chen G (2002) Pinning control of scale-free dynamical networks. *Physica A* 310(3):521–531

5. Wang XF, Su HS (2008) Recent progress in control of complex dynamical networks. *Adv Mech* 38(6):751–765
6. Chen GR (2013) Problems and challenges in control theory under complex dynamical network environments. *Acta Automatica Sinica* 39(4):312–321
7. Liu YY, Slotine JJ, Barabási AL (2011) Controllability of complex networks. *Nature* 473(7346):167–173
8. Yan G, Ren J, Lai YC et al (2012) Controlling complex networks: how much energy is needed? *Phys Rev Lett* 108(21):218703
9. Nepusz T, Vicsek T (2012) Controlling edge dynamics in complex networks. *Nat Phys* 8(7):568–573
10. Ferrarini A (2011) Some thoughts on the control of network system. *Netw Biology* 1(3–4): 186–188
11. Jia T, Pósfai M (2014) Connecting core percolation and controllability of complex networks. *Sci Rep* 4:5379
12. Lombardi A, Hörnquist M (2007) Controllability analysis of networks. *Phys Rev E* 75(5): 056110
13. Meng XL, Xiang WL, Wang L (2015) Controllability of train service network. *Mathe Probl Eng* 1–8

A Parallel Passenger Flow Management System for Outside Subway Stations Based on ACP Approach

Huijuan Zhou, Meijie Jia, Yu Liu, Qiang Zhang and Manrong Yuan

Abstract During peak time period, the entire subway is overloaded for a large amount of passengers. To solve this problem, there are many strategies to reduce or limit the number of passengers entering the subway stations. This paper proposes the parallel passenger flow management system (PPFMS) based on ACP theory to manage the inbound passenger flow. The ACP theory, including artificial societies, computational experiments, and parallel execution, is playing an essential role in modeling and control of complex systems. For a parallel passenger flow management system, realistic artificial scenes are used to model and represent complex real scenes; computational experiments are utilized to train and evaluate a variety of estimating models; and parallel execution is conducted to optimize the passenger flow system and achieve perception and understanding of complex environments.

Keywords ACP theory · Parallel passenger flow management system (PPFMS) Passengers crowd index

1 Introduction

Supply and demand imbalance in passenger transportation system especially during peak time period is the key problem of urban rail transit operation and management. With more and more new lines open, urban rail traffic has shifted from single-line operation mode to the network operation mode. And because it is fast, convenient, and punctual, urban rail transit becomes the main traffic to commute. According to

H. Zhou (✉) · M. Jia · Y. Liu · Q. Zhang
Beijing Key Lab of Urban Road Transportation Intelligent Control Technology,
North China University of Technology, No.5 Jinyuanzhuang Road,
Shijingshan District, Beijing 100144, China
e-mail: zhouhuijuan@ncut.edu.cn

M. Yuan
Kunming Municipal Public Security Bureau Traffic Police Detachment,
Kunming 650011, China

the statistical data of Beijing subway, the existing 15 subway lines have sent 3.025 billion passengers 1 year and there are 8.264 million passengers volume a day on average. In order to balance the subway capacity, there are 75 stations implementing strategies to control passenger to enter subway station at the rush hour. Despite all this, more passengers are crowded in entrance, passageway, stairway, escalator, platform, and train. Overloaded passengers make people uncomfortable and bring potential safety hazard.

In order to manage the passengers outside subway station, there are many strategies such as setting up isolated bar outside subway station, boarding limit, and skip-stops. It usually depends on the experience or the auto fare collection (AFC) data to decide the limit time because the station would not acquire the real-time incoming and inbound passenger volume accurately.

In fact, the hysteresis of AFC data and subjectivity of experience result in the limitation of these strategies. Although significant effort has been made over the years, the passenger flow in subway is still overloaded during the rush hours. It turns out that the passenger flow problems cannot be solved by exclusively focusing on the inbound passengers. The time distribution of passengers arriving at station must be considered. Specifically,

- (1) Passengers arriving at the station are the source of inbound passenger, which is the key part of passenger flow management. In the process of balance supply and demand climatic factor, holiday factor and passenger behaviors must be considered in building effective passenger flow management system. Based on artificial passenger flow management system, artificial system is effective tool for this purpose [1].
- (2) Since the rules of passenger which arrive at subway stations are different every day, there is no once-for-all solution for coordinated and sustained changing status of passenger flow. The passenger flow system contains much random factors, and the method of computational experiments can be used to solve the difficulty of experimenting [2].
- (3) After building the artificial passenger flow management system, the sensitivity of the artificial system must be tested. By comparing the difference between the artificial passenger flow management system and the actual passenger flow system of the same strategy, the artificial system needs to be unceasingly revised. In that case, the concept of parallel system can be applied in seeking effective limit time [3].

Based on above considerations, the paper is aimed to develop a framework of Parallel Passenger Flow Management System (PPFMS) between artificial passenger flow management system and actual passenger flow system. This effort can be considered as an important application of ACP theory [4].

The rest of this paper is organized as follows. In Sect. 2, we introduce the basic concepts and research issues of the PPFMS method. In Sect. 3, we propose the modeling and experiments of developing PPFMS to achieve real-time passenger flow control. In Sect. 4, control strategies are given to refine the artificial system. In Sect. 5, concluding remarks are given.

2 Theory Introduction and Research Issues

The ACP theory [4, 5], including artificial systems, computational experiment, and parallel execution, is universal to solve all kinds of complex system problems. It has been applied in many areas such as urban intelligent parking system [6], safety and reliability of nuclear power plants [7], fire evacuation [8], public health emergency management [9], accelerated GPU(Graphic Processing Unit)computing technology [10], and botany [11]. Especially, as a successful application of the ACP theory in the intelligent transportation system (ITS) [12, 13], parallel management provides a new perspective in solving passenger flow control issues.

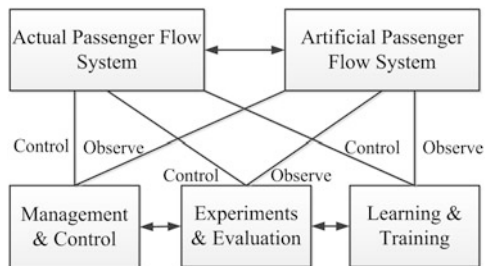
Based on the ACP theory, the process of developing artificial system is as follows: first of all, combined with the existing simulation control method, the artificial system which is similar to the actual system is built according to bottom-up approach. Second, computational experiments are conducted and optimal control strategies are explored. Finally, the actual system is controlled and managed by interacting information and parallel execution. The artificial system can be trained through learning from the actual system, and then, in turn, the actual system can be controlled and managed by means of experiments and analyses from the artificial system.

The framework of PPFMS is illustrated in Fig. 1, including the actual system and the artificial system. The interactions between the two systems advance the learning of actual system. By analyzing the differences between the two systems, we can predict the behaviors of actual system and improve the management strategies.

The key issues to be addressed in the building of artificial passenger flow system are the following:

- (1) Modeling: Currently, video processing method is often used to establish model for passenger flow, which is real-time and mature. In fact, when formulating strategies to limit passenger crowding, relevant departments cannot trial-and-error in actual passenger flow system according to the video data. So it is useful to build an agent passenger flow system to bear the risk of trial-and-error.

Fig. 1 Framework of PPFMS



- (2) Experimenting: When human and societal subjects are involved, it is difficult to get the passenger flow distribution once-for-all. The fixed limit time strategy is formulated by the frequent passenger flow distribution. Mechanical strategies will fail when emergency occurs. Artificial passenger flow management system can be used for solving emergency situations. Based on computational experiments, artificial passenger flow management system can be more sensitive and flexible.
- (3) Decision-making: The outstanding role of artificial system selects the optimum strategy among the multiple strategies. The artificial passenger flow management system can be used to test different strategies and decisions, and then evaluate the effects of them. When the bottom-up system is developed, it can be used for decision-making.

3 Modeling and Computing the Artificial PFMS

The goal of research is to refine the restricted strategies. By building a parallel system, which uses artificial systems as alternatives for the purpose of making decisions, the restricted strategies can be tested. The artificial passenger flow management system consists of three parts: passenger flow management system—artificial societies (PFMS-A), passenger flow management system—computational experiments (PFMS-C), and passenger flow management system—parallel execution (PFMS-P). The framework of artificial PEMS is given in Fig. 2.

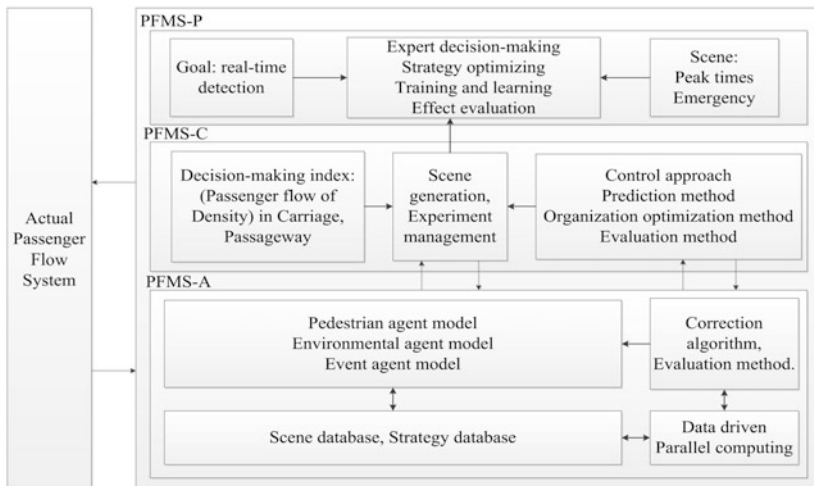


Fig. 2 Framework for artificial PFMS

3.1 Modeling the Artificial PFMS

(1) Artificial PFMS based on agent

The artificial PFMS, which is based on the agency model, consists of three parts (PFMS-A, PFMS-C, and PFMS-P). The artificial system consists of pedestrian agent model, environmental agent model, and event agent model. Pedestrian agent model is built based on pedestrian characteristics such as route choice, queuing process, and pedestrian avoidance. Environment agent model includes inside and outside the station. The inside environment includes density of passenger in carriage and passageway. The outside environment includes transit time, weather condition, etc. Event agent model covers emergencies events to show the evolution process of passenger flow.

(2) Support conditions

Scene database: Putting the scene of actual passenger flow management system into scene database is the basis of computational experiments. The scene database includes passenger flow state, environmental state, transfer state, and facility status. The significance of constructing the scene database is offline optimizing strategies of different scenes by computational experiments. The artificial passenger flow management system can provide the optimal strategy at the first time when the actual passenger flow system matches the artificial passenger flow management system.

Correction algorithm: For different scenes, computational experiments can get the corresponding strategies. While the passenger flow is dynamic, the artificial strategies need to be corrected constantly. The purpose of correction algorithm is to make the artificial system match the actual system as similar as possible. The correction includes models and their calibrations. The applicability of model is depended on the difference between actual system and artificial system. The model's parameters and related regulations are specified for calibration.

Evaluation method: The evaluation method is the key of testing artificial system. It includes static verification and dynamic verification. Static verification covers infrastructure situation, operation rules, etc. Dynamic verification is to validate the process of environmental change, such as the distribution of passenger flow during peak hours and the distribution under emergency.

3.2 Computational Experiments of Artificial Passenger Flow System

The computational experiments of artificial system contain scene design, algorithm design, optimization design, emergency plan design, etc. The scene design is the basic of computational experiments. For a complex system, the scene needs to

include all the influence factors. These rules of scene design should be considered: (1) Reflecting the real scene objectively; (2) Highlighting the scene characteristics; and (3) Considering the facility structure to the full. Based on that, the scene design is divided into static scene and dynamic scene. Static scene covers space structure of the station, infrastructure, passenger passage, road signs and traffic marking, etc. Dynamic scene includes passenger flow generation, weather and seasonal factors, emergency, equipment failure, etc. Static scene is the basic of dynamic scene design.

Based on the scene design, the other parts of computational experiments can be operated efficiently. For offline experiments, the new strategy should be experimented constantly to check the reliability. The modified strategies are also putting into strategy database. Online experiments are to search the similar scene from the scene database and find the relevant strategy from the strategy database. If the scene is new, a proper strategy according to experience should be selected.

Reinforce learning is a computational approach to automating decision-making. In the process of evaluating the strategies, hierarchy analysis and expert decision can be used to select the optimal decision. Meanwhile, the results of offline experiments can be solved as basic data and the online results can be selected as the final strategies which are the basic steps of reinforce learning.

4 Parallel Execution of Artificial Systems

To realize the real-time control, the actual passenger flow system and artificial passenger flow management system should be executed parallel. Artificial passenger flow management system consists of three parts:

- (1) The parallel execution of artificial system includes executed constantly in peak time and common time, evaluated and updated constantly, formulated strategies under emergency, adaptability analysis with weather changing or in holidays.
- (2) The strategies of artificial passenger flow system decided by the facility operation status, passenger crowd index including carriage, passageway and the area in front of the station, the distribution of passenger flow arriving at station, and the change of overflow time before entering the station. Integrating all these factors, the situation of passenger flow is estimated and then the corresponding strategies would be put forward.
- (3) After parallel execution, the actual passenger flow system would produce a series of results. The artificial system can receive the differences between the two systems and evaluate the results of strategies. On the one hand, according to the differences, artificial passenger flow management system can be improved further; on the other hand, the strategies of actual passenger system can be optimized too. The framework of the three parts is given in Fig. 3.

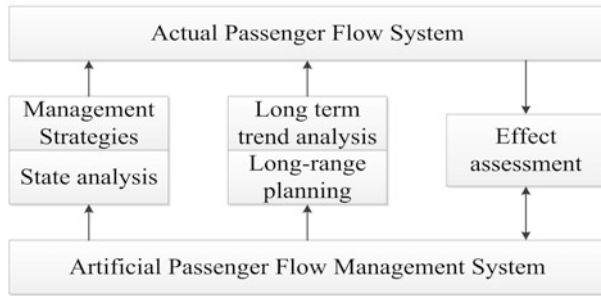


Fig. 3 Framework of parallel execution for PFMS-P

In the process of parallel execution, millions of data will be sampled. Data-driven parallel systems in cyberspace are the key to solve the trial-and-error challenge. To overcome the challenge, PDP (parallel dynamic programming) is introduced to the parallel execution, which means “artificial societies for descriptive analytic, computational experiments for predictive analytic, and parallel execution for prescriptive analytic” [14]. The principle of PDP is data driven for system and to utilize big data [15]. In the artificial system, data are collected for further descriptive analytics and predictive analytics during the parallel execution. The process of parallel execution based on the optimal principle of dynamic programming and evaluations is selected in the computational experiments. According to the feedback results of evaluating, the artificial system interaction can be conducted by observing the errors. Then, the decision-making system of artificial passenger flow management system can be updated. Finally, no matter how many possible strategies, we can only choose one to implement in the actual passenger flow system. So, after analyzing the results of different strategies, the rules can be extracted and the artificial passenger flow system can be improved.

In summary, the artificial passenger flow management system is based on the passenger flow management system—computational experiments (PFMS-C). First, it realizes the real-time monitor and evaluation by receiving information of environments, passenger flow, and infrastructure. Second, it makes the real-time calculation, selection, conduct, and evaluation come true by analyzing and managing the perceptual information of actual system. By experimenting as the closed loop, the reliability, safety, and efficiency of the two systems can be improved.

5 Concluding Remark

The artificial passenger flow management system is more feasible and flexible for refining the restricted strategy. For complex system, the ACP theory shows the outstanding adaptability of complex system. Based on that, the paper develops a framework of parallel passenger flow management system. Through the evaluation

results of artificial passenger flow management system, the operational efficiency of actual system can be improved. In the process of establishing parallel system, we found some major findings:

- (1) In the artificial passenger flow management system, the unpredictable of human behaviors and complexity of actual passenger flow system are considered in the process of developing the artificial system, which makes the strategies more flexible and humanizing. The support of ACP theory makes the refine management come true.
- (2) Many real-world systems which involved human and societies (Cyber-Physical-social system) [14] cannot be “tried” sufficiently for the fact of security and cost. The ACP theory provides a new way to test the strategies without destroying the actual passenger flow system, which saved human resources and materials.
- (3) Combined with the big data, the artificial passenger flow management system can be more dynamic and flexible. Combing the virtual data and historical data, the artificial passenger flow management system can be built perfectly which models the actual system and the objectives of agents, which can refine the restricted strategies to balance the supply and demand when passenger ballooned.

Acknowledgements The authors would like to acknowledge the support of National Key R&D Program of China (No.2016YFB1200402) and Scientific Research Project of Beijing Education Committee (PXM2017-014212-000031, PXM2017-014212-000033).

References

1. Wang FY, Lansing SJ (2004) From artificial life to artificial societies: new methods in studying social complex systems. *J Complex Syst Complex Sci* 1(1):33–41 (in Chinese)
2. Wang FY (2004a) Computational experiments for behavior analysis and decision evaluation of complex systems. *J Syst Simulat* 16(5):893–898 (in Chinese)
3. Wang FY (2004b) Parallel systems for management and control of complex systems. *J Control Decision* 19(3):485–490 (in Chinese)
4. Wen D, Yuan Y, Li XR (2013) Artificial societies, computational experiments, and parallel systems: an investigation on a computational theory for complex socioeconomic systems. *IEEE Trans Serv Comput* 6(2):177–185
5. Wang FY (2008) *Social computing: fundamentals and applications*
6. Wang FY, Yang LQ, Yang J et al (2016) Urban intelligent parking system based on the parallel theory. In: *International conference on computing, networking and communications*, IEEE, 1–5
7. Hu Y, Liu X, Wang F et al (2012) An overview of agent-based evacuation models for building fires. In: *IEEE international conference on networking, sensing and control*, IEEE, 382–386
8. Zhu F, Wang FY, Li R et al (2011) Modeling and analyzing transportation systems based on ACP approach. In: *International ieee conference on intelligent transportation systems*, IEEE, 2136–2141

9. Duan W, Cao Z, Wang Y et al (2013) An ACP approach to public health emergency management: using a campus outbreak of H1N1 influenza as a case study. *IEEE Trans Syst Man Cybernetics Syst* 43(5):1028–1041
10. Cui F, Cheng C, Wang F, et al (2010) Accelerated GPU computing technology for parallel management systems. In: *Intelligent control and automation*, IEEE, 5343–5347. (in Chinese)
11. Kang M, Wang FY (2017) From parallel plants to smart plants: intelligent control and management for plant growth. *IEEE/CAA J Automatica Sinica* 4(2):161–166
12. Wang FY, Tang S (2005) A framework for artificial transportation systems: from computer simulations to computational experiments. In: *Intelligent transportation systems, Proceedings*, IEEE Xplore, 1130–1134
13. He F, Miao Q, Li Y, et al (2006) Modeling and analysis of artificial transportation system based on multi-agent technology. In: *Intelligent transportation systems conference, ITSC'06*, IEEE, 1120–1124
14. Wang FY, Zhang J, Wei Q et al (2017) PDP: parallel dynamic programming. *IEEE/CAA J Automatica Sinica* 4(1):1–5
15. Wang FY (2012) A Big-data perspective on AI: Newton, Merton, and analytics intelligence. *Intell Syst IEEE* 27(27):2–4

Application of Triboelectric Nanogenerator in the Railway System

Laixin Geng, Sen Bian, Teng Li, Zhao Si and Zijun Wei

Abstract The triboelectric nanogenerator (TENG) is a new type of durable, clean, maintenance-free, low-cost, and high-reliability power generation device; so it is very suitable for energy harvesting in rail transit. This paper analyzed the power generation principle of the TENG, and the power storage mode for its pulse output; by considering the disadvantages of the standard interface circuit, an improved interface circuit has been designed, which can be used in the field of railway system to provide the power for the condition and security monitoring of the rail transportation. It expands the application range of the TENG.

Keywords TENG · Interface circuit · Energy harvester

1 Introduction

With the rapid increase of wireless devices, implanted medical devices, and low-power portable electronic devices, the harvest of mechanical energy from the surrounding environment to drive these devices sustainably has attracted worldwide attention. To date, various methods have been demonstrated, such as using piezoelectric [1, 2] and electromagnetic [3, 4] effects to harvest environmental mechanical energy. However, the study of harvesting rail track vibration energy is not as much as the road. Energy harvest method based on electromagnetic effect has the disadvantages of low energy density, low level of integration, and high cost. Under the large strain, piezoelectric materials will appear electrical fatigue, and its electrical properties will decline [5, 6]. While the process is irreversible, it will cause the energy harvester lose its efficacy. In order to ensure the normal output

L. Geng · T. Li (✉)
Beijing Jiaotong University, Beijing 100044, China
e-mail: liteng@bjtu.edu.cn

S. Bian · Z. Si · Z. Wei
Research Center of Green Communication Technology, Research Institute
of China Mobile Communication Co, Beijing 100032, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_90

895

power of the harvester, the above problems need to be solved. The recent invented triboelectric nanogenerator (TENG) may be one solution because of its simple structure, high reliability, no pollution, high energy density, and low cost [7–10] for harvesting vibration energy for a long term in the natural environment.

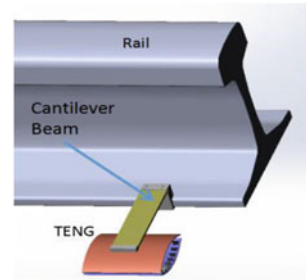
The TENG harvests energy through the friction between the electrification and electrostatic induction coupling, achieving electromechanical conversion. There are a variety of forms for TENG's packaging process; how to make the harvester maintenance-free gradually become a research hotspot. So far, there are four types of TENGs according to their different power generation modes: the contact type, sliding type, single-electrode type, and compartment type. Based on these modes, the harvester can basically realize the harvesting of all kinds of mechanical energy; however, the contact-separation mode is more suitable for the energy harvesting of the rail track. Here, we introduce one TENG design, which can be used for harvesting track vibration energy. By combining the cantilever beam, which is one of the structures often used in the piezoelectric harvester [11]; the TENG is installed on the sleeper. The cantilever can amplify the vibration displacement, thus improving the efficiency of TENG. This design can effectively harvest vibration energy with great strain without disturbing the routine check and maintenance of the track, and the impact on the traffic safety is less. This design also has the advantages of flexible installation and high reliability without the need of the rail modification.

At present, most of TENGs' interface circuits are standard interface circuit [12]. Even though the rectifier in the interface circuit can rectify somehow the shape of voltage pulses from the TENG, the output voltage of the interface circuit is still very high with a very fast rise rate, making the energy storage difficult. But the power of each pulse is too low to drive the follow-up devices directly. Some scholars used the pulse charging technology to charge Li battery [13]; however, this technology may shorten the lifetime of the Li battery in addition to decrease the energy conversion efficiency. In this paper, based on the standard circuit, an improved interface circuit has been designed with the improved output performance. Its feasibility has been proved through experiments. The parallel multi-machine power generation can also be achieved based on the new circuit.

2 An Improved Interface Circuit

The installation location of the cantilever beam and the TENG on the track is shown in Fig. 1. The copper cantilever is clapped at the base edge of the track; the length of the cantilever beam L is 100 mm, width B is 15 mm, and the thickness T is 1.38 mm; the mass M of the mass block on the end of the beam is 24 g. The first-order natural frequency f of the cantilever is around 16.8 Hz after the modal analysis simulation using the finite element software ANSYS. The maximum displacement D_{\max} of the end of the beam is 7.3 mm. The TENG harvester is installed under one end of the beam, and the distance between the beam and the TENG is

Fig. 1 The installation diagram of the TENG



1 mm. The cantilever beam can convert the vibration of the track to a larger displacement at the end of the beam, and then apply a force to the TENG. Because the inertia of the mass block is large and the TENG is a flexible material with the plastic characteristics, here we neglect the influence of the TENG to the movement of the beam.

Because the voltage output of the TENG is a pulsed AC form, the full bridge rectifier is generally used to rectify the AC output into DC, and its schematic is shown in Fig. 2a; then the rectified DC energy will be stored in the capacitor. The equivalent capacitor of the TENG is in the range of tens of picofarads. To maximize the harvested energy, the capacitor's value after the rectifier should also be in the range of tens to hundreds of picofarads to be charged quickly to follow the variation of the vibration energy; However, the voltage peak value will be high using a small capacitor; the large voltage variation also makes the energy storage difficult. In order to reduce the capacitor output voltage, while improving the energy harvesting efficiency, a strategy of capacitors in series when charging and parallel when discharging is used here; the schematic of the improved interface circuit with two capacitors is shown in Fig. 2b. With this new design, capacitors can be charged quickly and then provided a relatively stable power to the load.

The four working states I–IV of the newly designed circuit connecting the rectifier and the load are shown in Fig. 2c. I–IV correspond four states in one working cycle of the TENG: the state starting to release two plates of the TENG, the state releasing them to the vertex, the state starting to compress them, and the state compressing them to the bottom point. In the first three states I–III, capacitors are kept connecting with the TENG output, and they are in the charging status by collecting the power from the TENG; they are connected in series during these three states. When the TENG reaches the state IV, four mechanical switches 1–4 are switched on, and thus the interface circuit is disconnected from the generator TENG and the capacitors are connected in parallel, starting to discharge to provide the power to the load.

The capacitance of the capacitors cannot be exactly the same due to the limit of the production process, resulting in the uneven terminal voltage of the capacitors. The internal current circulation caused by the uneven terminal voltages will cause the damage of the capacitors. Using two Schottky barrier diodes, the circulation will

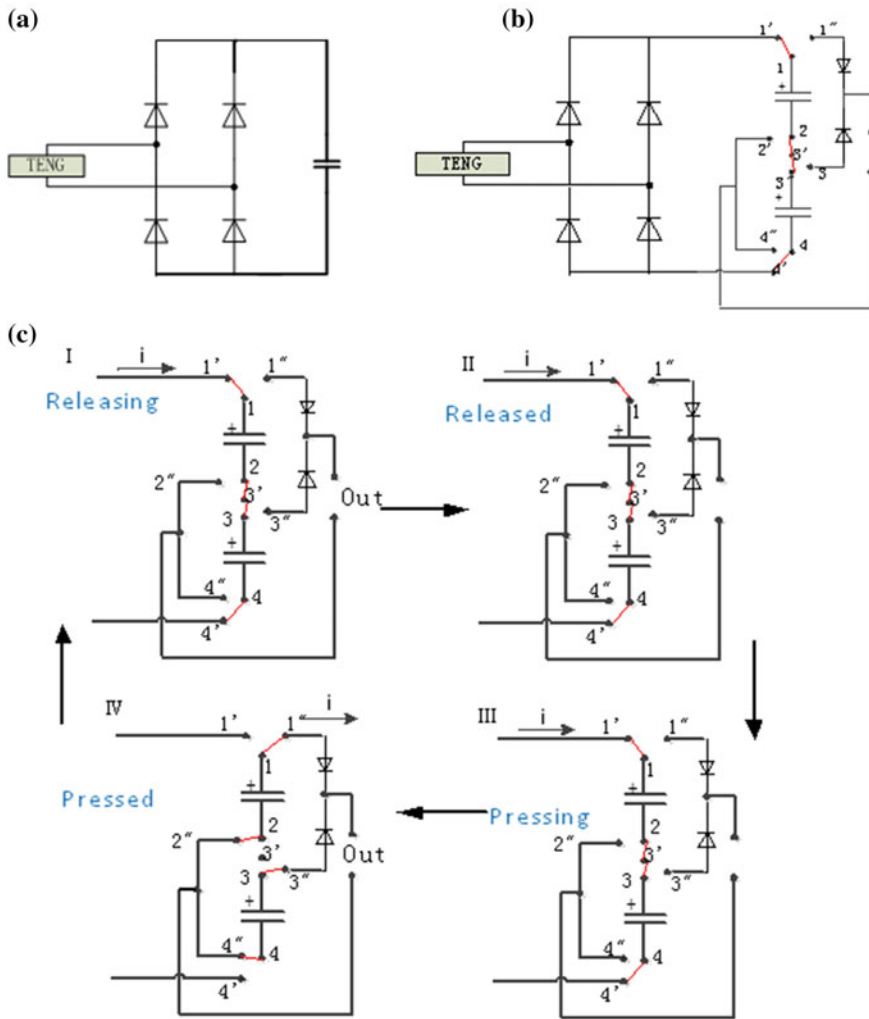


Fig. 2 The schematic of the TENG and interface circuit in different topologies. **a** Standard interface circuit. **b** Improved interface circuit. **c** The complete working cycle of the improved circuit

be prevented. Furthermore, all the switches are mechanical limit switches, and there is no extra power loss caused by external electrical components.

Figure 3 shows the output waveforms of the TENG, the standard interface circuit, and the improved interface circuit, respectively. The improved circuit has completed the dual roles of the rectification and reducing the voltage steepness, thus increasing the DC component of the output to the load, and optimizing the output of the TENG.

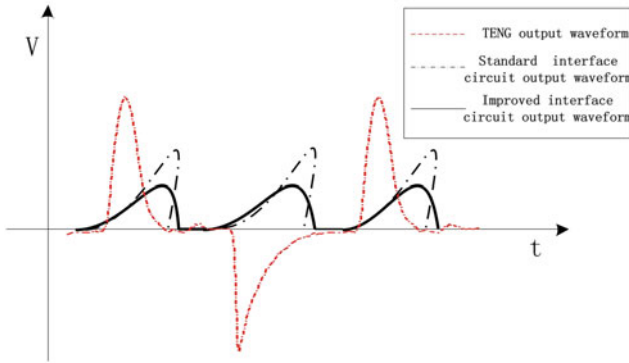


Fig. 3 Output waveforms of the TENG, the standard interface circuit, and the improved interface circuit

By connecting more capacitors in the circuit, it can reduce the amplitude of the output voltage and slow down voltage steepness further. It is assumed that the output voltage across the N capacitors connected in series is V_{os} , the voltage across the N capacitors connected in parallel is V_{op} , and V_N is the voltage of a single capacitor. We have

$$V_{os} = NV_N \tag{1}$$

$$V_{op} = V_N \tag{2}$$

By combining Eqs. (1) and (2), we have

$$V_{op} = \frac{1}{N} V_{os} \tag{3}$$

From Eq. (3), we can see that the voltage amplitude when capacitors in parallel will be $1/N$ of the voltage when capacitors in series; voltage amplitude will decrease because of the parallel connection of capacitors.

Suppose the equivalent capacitor when capacitors in series are C_{es} , the parallel equivalent capacitor is C_{ep} , and the equivalent resistor of the circuit is R , then we have

$$\tau_p = RC_{ep} = N^2RC_{es} = N^2\tau_s \tag{4}$$

From Eq. (3), we can see that the time constant when capacitors in parallel is N^2 times larger than capacitors in series. The voltage steepness is greatly reduced.

3 Experimental Phenomena and Conclusions

3.1 Track System Vibration Causes and Train Load

The disturbance of the wheel–rail system is the main reason for the vibration of the rail system. Therefore, it is necessary to clarify the excitation characteristics of the wheel–rail system.

The seamless rails are used in the contemporary rail transportation system to reduce the number of joints; meanwhile, the wheel has worn-type tread; so the impact caused by the unevenness of the wheel can be ignored. Therefore, the main reason for the vibration of the rail system is the geometrical irregularity of the track.

In general, the geometric irregularities of a track can be approximated by a single or multiple simple harmonic waves. For example, due to the unevenness of the joint, the disturbance usually is a single harmonic disturbance. Another example is that most rails in the world have wavy wear, resulting the generated disturbance has typical continuous harmonics. It is therefore reasonable to use sine and cosine functions to describe the disturbance.

As shown in Fig. 4, for a single harmonic, the track shape can be described by Eq. (5),

$$Z_0 = \frac{1}{2}a[1 - \cos(\omega t)] \quad \left(0 \leq t \leq \frac{L}{v}, \omega = \frac{2\pi v}{L}\right), \tag{5}$$

where a is the uneven wave depth and L is the uneven wavelength.

Three typical wavelengths of the rail uneven disturbance are considered:

- (1) the vehicle itself is uneven, here we take the caused uneven wavelength L_1 as 10 m, the wave depth a_1 as 5 mm;
- (2) the unevenness caused by the additional dynamic load applied on the track, taking the wavelength L_2 as 1 m, the wave depth a_2 as 0.3 mm;
- (3) the unevenness caused by the waveform loss, taking the wavelength L_3 as 0.5 m, the wave depth a_3 as 0.1 mm.

From the above three kinds of unevenness, we can get the displacement model caused by the multi-wave simple harmonic disturbance, namely

Fig. 4 Single harmonic excitation

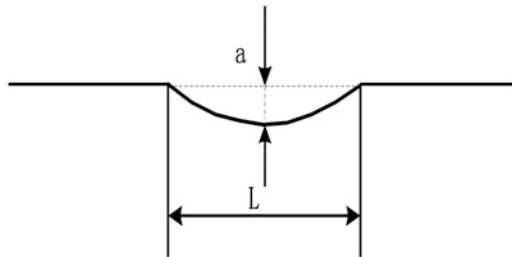


Table 1 Train load frequency components

Frequency/Hz	Amplitude/kN	Phase (°)
0	157.36	90.0
0.111	1.63	67.8
1.667	2.28	106.3
16.667	2.83	200.9
33.222	2.28	214.0

$$Z_0 = \frac{1}{2}a_1[1 - \cos(\omega_1 t)] + \frac{1}{2}a_2[1 - \cos(\omega_2 t)] + \frac{1}{2}a_3[1 - \cos(\omega_3 t)] \quad (6)$$

Taking the rail transit as an example, here we suppose the locomotive speed v is 60 km/h, and then take it into Eq. (6); we can get

$$Z_0 = 2.7 - 2.5 \cos(10.47t) - 0.15 \cos(104.75t) - \cos(209.44t) \quad (7)$$

According to the D'Alembert's principle, the frequency components of the train can be obtained by FFT analysis of Eq. (7) [14], and results are shown in Table 1.

Through Table 1, we can see there are mainly five kinds of frequency components in the disturbance. The physical meanings of these frequencies are explained here, where 0 Hz corresponds to the amplitude of the locomotive body when it is at stationary; 0.111 Hz corresponds to the vibration impact of the locomotive on the load, and the amplitude is 1.63 kN. 16.667 Hz and 33.222 Hz correspond to the dynamic responses of the train under track harmonics. The simulated train load expression using these five frequencies is shown in Eq. (8):

$$\begin{aligned} P(t) = & 157.36 + 1.63 \sin(0.697t + 1.183) \\ & + 2.28 \sin(10.474t + 1.855) \\ & + 2.83 \sin(104.722t + 3.506) \\ & + 2.28 \sin(208.740t + 3.735) \end{aligned} \quad (8)$$

It can be seen that the amplitude reaches a maximum value of 2.83 kN at the frequency of 16.667 Hz.

The output performance of the contact-separation mode TENG is closely related to the vibration frequency. Studies have shown that TENG has a higher energy harvesting efficiency when the excitation is in the frequency range of 7–30 Hz [15]. So the TENG is suitable for the energy harvesting of the track vibration.

3.2 Test of the Improved Interface Circuit

To characterize the output performance of the TENG, we measured its output voltage. The TENG was connected to the oscilloscope for the voltage measurement,

where the internal resistance of the oscilloscope probe was 100 MΩ. The output voltage waveform of the TENG is shown in Fig. 5. Output voltage waveforms corresponding to different loads are shown in Fig. 6 when there are two 100pF capacitors in the improved circuit. As shown in Figs. 5 and 6, the peak value of the output voltage was reduced from 720 to 60 V when the load was 100 MΩ, and the voltage pulse duration time was increased from 10 ms to 1 s. The output performance was improved using the improved interface circuit.

Experiments have been performed to see the influence of the number of connected capacitors on the output. Here, we keep the parallel equivalent capacitor C_{ep} the same to be 200pF all the time. Figure 7a, b, and c shows output voltage waveforms with different loads when the number of capacitors was 2, 4, and 8, respectively. Take the waveforms with 100 MΩ load as the example, when 2 capacitors were connected, the voltage peak was 60 V, and the pulse duration was 1 s; when four capacitors were connected, the voltage peak was 42 V and the pulse duration was 6 s; when eight capacitors were connected, the voltage peak was 15 V and the pulse duration was 6 s. Therefore, the more the capacitors connected to the output in the new circuit, the larger effect will be produced to reduce the output

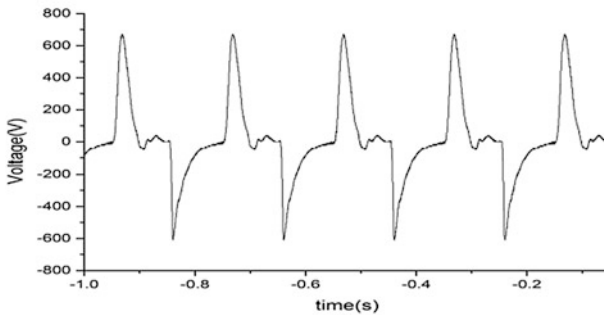
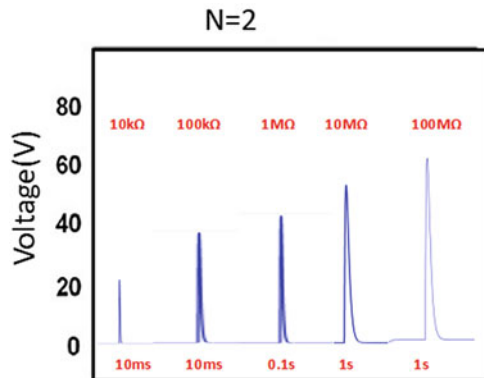


Fig. 5 TENG voltage output waveform

Fig. 6 Output voltage waveforms from measured data with two connected capacitors



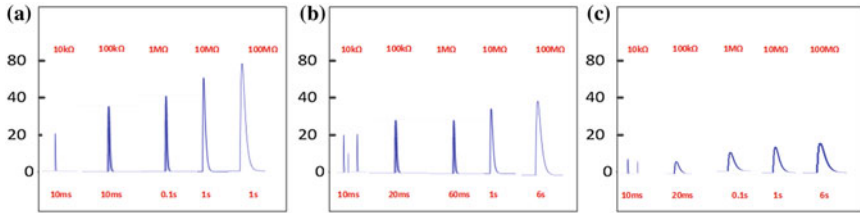


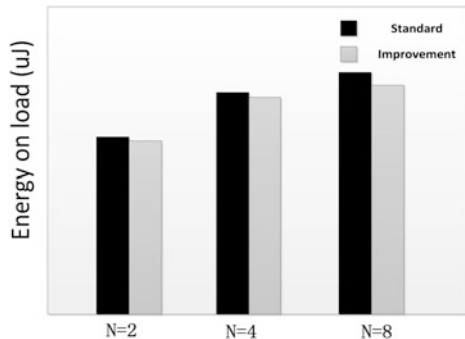
Fig. 7 Output voltage waveforms with different number of connected capacitors. **a** waveforms with two capacitors. **b** Waveforms with four capacitors. **c** Waveforms with eight capacitors

voltage amplitude and voltage steepness, and also the voltage DC component will be increased.

3.3 Interface Circuit Efficiency

Improved interface circuit introduces mechanical switches; although the power loss of this kind switch is small, its contact resistance loss is inevitable; there is also the diode power loss in the circuit, resulting the efficiency of the improved circuit is slightly smaller than the standard interface circuit without the mechanical switches. As shown in Fig. 8, when the number of capacitors was 2, the harvested energy of the standard interface circuit was 1.5 uJ; however, the energy harvested by the improved interface circuit was 1.45 uJ. So the efficiency was 96.7%. When the number of capacitors was 4 and 8, respectively, the harvested energy was 2.3 and 2.45 uJ, and the efficiency was 96 and 95.3%, respectively. It can be seen that the efficiency was decreased with the increase of the number of capacitors, the number of switches, and the number of diodes. However, the efficiency of the improved interface circuit was still above 95%, which satisfies the power supply requirement of subsequent sensors.

Fig. 8 Energy conversion with and without switch interface circuit



4 Conclusion

For the new technology of TENG, we improved its output characteristics on the basis of its original standard interface circuit and verified its feasibility of collecting energy in the rail system. The power quality of the improved interface circuit, including the voltage amplitude and the rising rate, is improved greatly compared with the standard interface circuit, and the efficiency is over 95%.

Acknowledgements This work is supported by Fundamental Research Funds for the Central Universities (2015JBM086, 2017JBM062).

We would like to thank the Research Center of Green Communication Technology, Research Institute of China Mobile Communication Co., Ltd. for the support of this study.

References

1. Wang J, Shi Z, Xiang H, et al (2015) Modeling on energy harvesting from a railway system using piezoelectric transducers. *Smart Mater Struct* 105017–105029
2. Yang R, Qin Y, Dai L et al (2009) Power generation with laterally packaged piezoelectric fine wires. *Nat Nanotechnol* 4(1):34–39
3. Beeby SP, Torah RN, Tudor MJ et al (2007) A micro electromagnetic generator for vibration energy harvesting. *J Micromech Microeng* 17:1257–1265
4. Koukharenko E, Beeby SP, Tudor MJ et al (2006) Microelectromechanical systems vibration powered electromagnetic generator for wireless sensor applications. *Microsyst Technol* 12 (10–11):1071–1077
5. Yang G, Yue Z, Li L (2007) Study on field fatigue properties and mechanism of piezoelectric ceramics. *Acta Metallurgica Sinica* (01):1–6. (Ch). (in Chinese)
6. Liu Haifeng, Tian Tian, Xie Jun (2000) Progress in electrical fatigue of piezoelectric ceramics. *Aerospace Mater Technol* 30(06):1–4 (in Chinese)
7. Fan FR, Tian ZQ, Wang ZL (2012) Flexible triboelectric generator. *Nano Energy* 1(2):328–334
8. Wang S, Lin L, Wang ZL (2012) Nanoscale triboelectric-effect-enabled energy conversion for sustainably powering portable electronics. *Nano Lett* 12(12):6339–6346
9. Lin L, Wang S, Xie Y et al (2013) Segmentally structured disk triboelectric nanogenerator for harvesting rotational mechanical energy. *Nano Lett* 13(6):2916–2923
10. Zhu G, Lin ZH, Jing Q et al (2013) Toward large-scale energy harvesting by a nanoparticle-enhanced triboelectric nanogenerator. *Nano Lett* 13(2):847–853
11. Wang ZL (2013) Triboelectric nanogenerators as new energy technology for self-powered systems and as active mechanical and chemical sensors. *ACS Nano* 7(11):9533–9557
12. Gao MY, Wang P, Cao Y et al (2016) A rail-borne piezoelectric transducer for energy harvesting of railway vibration. *J VibroEng* 18(7):4647–4663
13. Pu X, Liu M, Li L et al (2016) Efficient charging of Li-Ion batteries with pulsed output current of triboelectric nanogenerators. *Adv Sci* (Weinheim, Baden-Wurttemberg, Germany) 3 (1):1500255
14. Zhang Y, Hou Y, Yang Z (2005) Effects of track structure parameters on wheel/rail interaction force. *J Hebei Transport Poly* (2):1–3. (Ch). (in Chinese)
15. Wang X, Niu S, Yi F et al (2017) Harvesting ambient vibration energy over a wide frequency range for self-powered electronics. *ACS Nano* 11(2):1728–1735

Some Research on Suspension Gap Sensor of High-Speed Maglev Train

Jun Li, Jun Wu, Junyuan Tang and Shengjun Huang

Abstract Aiming at the special temperature environment problems faced by the suspension gap sensor of high-speed maglev train, the paper studies the circuit model analysis and software compensation respectively. Through the design of differential circuit structure and temperature correction table, we solve the problem of sensor temperature drift. At the same time, using the tracking differential device to deal with the change of the vertical gap signal twice, compared with the existing acceleration signal, we get a simple fault self-detection method for the acceleration sensor.

Keywords Gap measurement · Temperature drift · Fault self-detection
High-speed maglev train · Tracking differentiator

1 Introduction

High-speed maglev train is a kind of “zero altitude flight” fast transport, which has no contact with the track. During the rapid operation, the vehicle’s suspension control system controls the vehicle to remain in a suspended gap of about 10 mm. The suspension gap detection is performed by an inductive displacement sensor, which is important for stable suspension control [1].

As shown in Fig. 1, the suspension gap sensor is an inductive gap sensor, which produces two independent gap signals, an acceleration signal, a speed signal, and a diagnostic signal. It uses the structure with constant frequency, adjustable amplitude.

As shown in Fig. 2, the sensor is mounted in close contact with the suspension electromagnet, and the detection coil is embedded between the poles of the suspension electromagnet. The solenoid coil is in the form of aluminum foil.

J. Li (✉) · J. Wu · J. Tang · S. Huang
School of Mechatronics and Automation, National University
of Defense Technology, Changsha, China
e-mail: NUDT_lijun2016@163.com

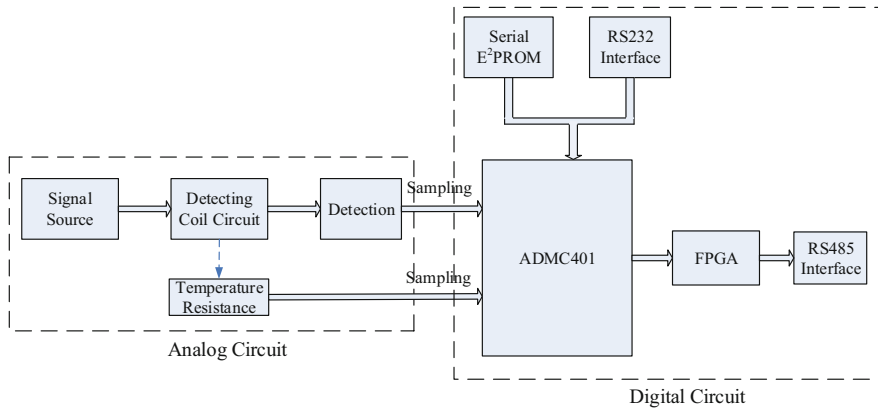


Fig. 1 The circuit structure of suspension gap sensor

Meanwhile, under the normal operating conditions, the coil current density is large, which leads the high temperature of electromagnet, especially when the vehicle is in static suspension or low-speed operation, because the heat dissipation is not good, the temperature of the electromagnet rises fast, the heat generated will be quickly transmitted to the suspension gap sensor detection coil and processing circuit, The maximum temperature that can be reached is 85 °C. If there is no effective measure, the sensor output is easy to result in a large temperature drift. This brings the temperature of the problem. Therefore, the gap sensor must reduce the temperature drift to improve the stability of the gap output.

In addition, because of the suspension gap sensor’s special application conditions, when the vehicle is in the low-speed operation, suspension electromagnet may product large heat, which will accelerate the aging of the sensor. Besides in the high-speed operating conditions, the vibration of the lower part of the vehicle is large, which is also easy to cause the sensor circuit device loosening and lead to sensor fail.

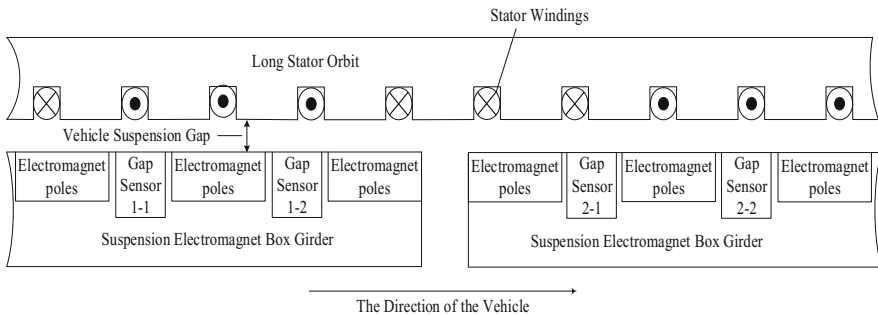


Fig. 2 Structure of suspension system in high-speed maglev

In a word, the suspension gap sensor is very important for the maglev train's normal operation, but the application environment is harsh, which leads to the large possibility of failure, so the study of the suspension gap sensor fault self-test has an important practical significance for high-speed maglev vehicle production and operation and maintenance.

2 Solution to the Problem of Temperature Drift

The suspension gap sensor of the high-speed maglev train has a certain particularity [2], it is embedded in the electromagnet between the poles, and its installation space is limited. In addition, for fault detection considerations, the gap sensor uses the redundant structure with two detection coils, the possibility of increasing the compensation coil is small, the paper attempts to study the detection coil to solve the problem.

Most of the literature [3, 4] think that the main reasons, which lead to the inductance sensor's temperature drift, are the resistance of detection coil, detection circuit parameters, and oscillation signal source. The compensation methods for the coil resistance factor are as follows: detection coil differential compensation, detection coil self-compensation, circuit parameter compensation, and compensation based on temperature model [5, 6]. The suspension gap sensor is a constant frequency amplitude-type inductive sensor, due to the working environment and the special structure, the coil resistance becomes the main link for the sensor temperature drift [7].

2.1 Analysis of Temperature Model

The eddy current gap sensor is composed of a detection coil and a processing circuit, and the processing circuit includes an analog circuit and a digital circuit. Digital circuit deals with the digital signal after A/D sampling, so there is no temperature drift problem. Therefore, the main part that leads to the temperature drift is the detection coil and analog circuit, and the detection coil, oscillation signal source, and detection circuit are the main causes of sensor output drift. The oscillation circuit and the detection circuit of the sensor are in series relationship. After considering the temperature variable T , the detection output can be expressed as

$$\begin{aligned}
 V_0(\delta, T) &= \cos \theta \cdot V_j \\
 &= \cos \theta \cdot V(\delta, T) - \cos \theta \cdot V_c(T) \\
 &= \cos \theta \cdot V(Z(\delta, T)) - \cos \theta \cdot V_c(T),
 \end{aligned} \tag{1}$$

where $\theta = (3\pi \cdot R_d/R_L)^{1/3}$ is the angle of the detection circuit, V_j is detection output, $V_c(T)$ is detection diode voltage drop, $V(\cdot)$ is the output voltage peak of oscillation circuit, and $Z(\delta, T)$ is the impedance of detection coil which is related to gap and temperature.

According to Eq. (1), the sensor output is mainly related to $V(\cdot)$ and $V_c(T)$. In the FM modulation circuit, $V(\cdot)$ is mainly related to the detection coil impedance and other oscillator circuit parameters. While in the constant frequency amplitude modulation circuit, $V(\cdot)$ is related to the detection of the coil impedance and the stability of the oscillation signal source. In poor ventilation or static suspension state, suspension electromagnet produces large heat and makes temperature rise fast, which results the detection coil to be in a relatively poor temperature environment for a long time. Therefore, in this condition, the detection coil becomes the main link of compensation for the sensor temperature drift.

2.2 Detecting Coil Impedance

Coil equivalent impedance is [8]

$$|Z(x, T)| = \sqrt{R^2(T) + [\omega L(\delta, T)]^2}, \quad (2)$$

where R is the detection coil resistance, $L(\delta, T)$ is the detection coil equivalent inductance, and ω is vibration angular frequency.

First, we analyze the relationship between inductance and gap and temperature. The detection coil of the gap sensor is a three-dimensional coil, and the long stator is a silicon steel-laminated tooth groove structure. In order to separate the temperature drift problem, it may be appropriate to suppose the coil is equivalent to a single plane coil and the coil wire is equivalent to a circular section. To simplify the relation between detection coil and the electromagnetic surface, we suppose that the conductor is measured silicon steel plane and the detection coil is parallel to the conductor. The excitation current is $i(t) = I \cdot e^{j\omega t}$. Gap sensor measures the gap between itself and the long stator, the static magnetism plays a leading role. According to the previous derivation results, inductance under the effect of silicon steel plane is

$$L = \frac{N}{I} \iint_S \vec{B}(x, y, 0) \cdot \vec{n} ds + \frac{N}{I} \iint_S \vec{B}(x, y, 2\delta) \cdot \vec{n} ds. \quad (3)$$

In the high-frequency excitation signal, the detection coil resistance is mainly AC resistance, and the detection coil resistance can be expressed as

$$R = \frac{N\bar{l}\rho_0(1 + \alpha T)}{\sqrt{2\pi r_0\delta}} \cdot \frac{\text{ber}(kr_0)\text{bei}'(kr_0) - \text{bei}(kr_0)\text{ber}'(kr_0)}{[\text{ber}'(kr_0)]^2 + [\text{bei}'(kr_0)]^2}, \tag{4}$$

where $\text{ber}(p) + j\text{bei}(p) = J_0(pj\sqrt{j})$ is the Kelvin Function, $J_0(\cdot)$ is zero-order Bessel function, r_0 is the wire radius, \bar{l} is the coil average perimeter, and in $k = \sqrt{2}/h$, h is the skin depth of the coil wire.

$$h = \sqrt{\frac{2\rho_0(1 + \alpha T)}{\omega\mu_0\mu_r}} \tag{5}$$

where ω is coil excitation frequency, ρ_0 is the resistivity at 0 °C, α is resistivity temperature coefficient.

From the above analysis, it can be seen that the impedance is a nonlinear function of temperature.

2.3 Differential Structure and Temperature Correction Table

The detection coil uses PCB circuit board form [6], and the coil material is the copper foil. Under the action of high-frequency excitation current, the skin effect is more obvious, the AC resistance of coil plays a major role, and it becomes bigger as the temperature gets higher and higher, which results in output drift. Designed as shown in Fig. 3, the differential structure of the temperature compensation scheme, in this form, the compensation coil structure design is more important.

As shown in Fig. 4, the coil is in the form of a rectangular PCB circuit board, so that the upper layer of copper can be used to form the detection coil, and in the lower layer of copper to form a compensation coil, the two coils overlap in the space position in order to ensure that the compensation coil does not have the

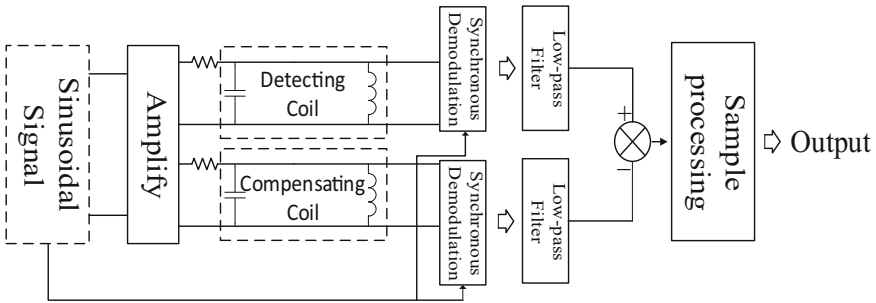
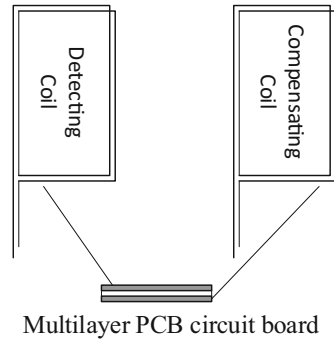


Fig. 3 Structure for temperature drifts compensation

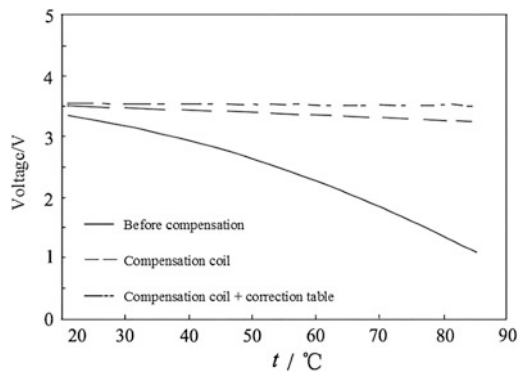
Fig. 4 Position of measurement coils and compensation coils



sensing effect to the outside conductor. Its winding is carried out as shown in Fig. 4, so that the magnetic field between each coil can cancel each other, so that the compensation coil is actually a noninductive coil, its resistance changes to the same resistance at the detection coil at the same time, when temperature changes. And the change of detection gap has no effect on it.

Although the differential structure eliminates the main effect of the temperature drift, the final output of the sensor still has some temperature drift, so it is necessary to use the appropriate temperature compensation circuit to compensate the temperature drift effectively. The circuit uses AD to acquire the temperature information of the temperature sensor PT1000, then the temperature information is digitally processed at 6-bit resolution and stored in the PROM. By judging the working temperature of the sensor and combining with the actual temperature drift, the system gives adequate compensation to the sensor. Figure 5 shows the comparison of the simulated outputs before and after temperature drift compensation. The temperature drift of the compensated sensor output in the range of 20–85 °C is within 0.05 V, and the corresponding drift is about 1% F.S.

Fig. 5 Test result compared with no compensations



3 The Processing of Gap Signal and Acceleration Signal

In the high-speed maglev train, the suspension gap sensor structure is complex. The gap sensor has a high frequency of use at a very high ambient temperature, which easily leads to sensor's breakdown [9]. The work of maintaining sensor is hard, so the sensor fault self-detection is of great importance.

The suspension gap sensor has a function with a certain fault self-detection, but this is only a self-detection for the gap measurement. The use of two independent gap detection coils is the method of fault self-test. When the error between the two gap signals is bigger than 0.15 mm, the system may determine that the gap sensor is fail. But so far there is no fault self-detection method for the acceleration sensor.

3.1 *Fundamentals and Concepts*

For the sensor self-detection, because the sensor installation space is limited, it is not enough to add new modules. This paper mainly wants to use the existing gap signal to form a kind of acceleration sensor fault self-detection.

As the acceleration sensor and the gap measurement coil in a rigid body, we think that the two are in a plane with the orbital surface as the base coordinate system. If the gap detection is normal, we can do two differential processing for the real-time vertical gap signal, then compare it with the existing acceleration signal, a simple fault self-detection method is formed.

3.2 *Design of a New Improved Tracking Differentiator*

As the gap output is nonlinear and constantly changing in real time, then the differential signal quality will be a problem, even sometimes the noise is flooded. If you can extract the "differential" signal, it will improve the controller performance and make the design of control simple.

In fact, if the known signal is discontinuous or with some random noise, then direct differential is difficult to get. In order to solve this problem, Han [10] proposed the concept of a nonlinear tracking differentiator in 1994. The tracking differentiator utilizes the fact that the numerical integration is superior to the numerical differentiation, and the differential of the given signal is transformed into the integration problem for a set of differential equations, which can track and process any signal differential.

Since the birth of the tracking differentiator, there have been many forms, and they have been widely used. Using a nonlinear form, the earliest tracking differential is based on the optimal control of the second-order system, it is with good fastness, but there is a certain flutter phenomenon. The linear tracking differential's

algorithm is simple and easy to implement, but it loses the nonlinearity of the tracer to track the rapidity of the differential. In literature [11], a high-speed tracking differentiator (hereinafter referred to as HSTD) is proposed and applied to the velocity estimation of the servo system. HSTD is simple and has good fastness, but it has some flutter phenomenon.

Aiming at the above problems, based on the fast tracer, this paper combines the advantages of linear tracking differentiator and nonlinear tracking differentiator, and proposes a new improved tracking differentiator (hereinafter referred to as NITD). When the error is large, NITD uses nonlinear links to speed up the trend toward the equilibrium point; when the error is small, the linear link is used to avoid the occurrence of flutter phenomenon. Meanwhile, the size of the linear work interval and the nonlinear interval is adjustable. The tracking differential not only has no flutter phenomenon, but also has good dynamic response and strong filtering ability, taking into the requirements of fastness and accuracy, it can achieve arbitrary signal tracking and differentiation, the algorithm is simple and easy to implement. The tracking differential is applied to the processing of the gap signal to realize the self-detection of the acceleration sensor.

In order to validate the validity of NITD, the simulation is carried out by MATLAB, and the results are compared with the discrete tracking differentiator (hereinafter referred to as TD) and HSTD. The simulation results are shown in Fig. 6.

As shown in Fig. 7, the sine wave signal is used as the gap measurement signal, and the acceleration signal obtained by NITD is basically the same as the actual acceleration signal which achieves the expected effect.

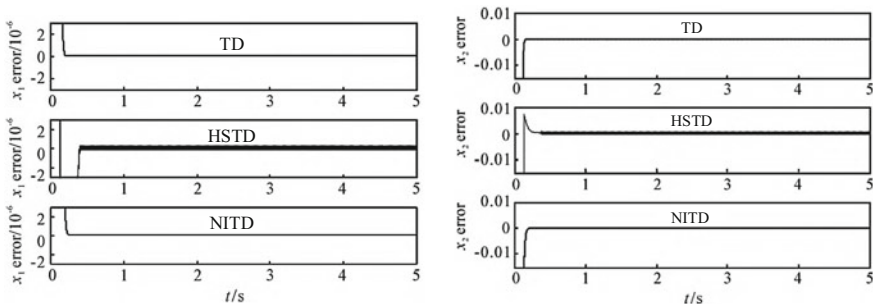
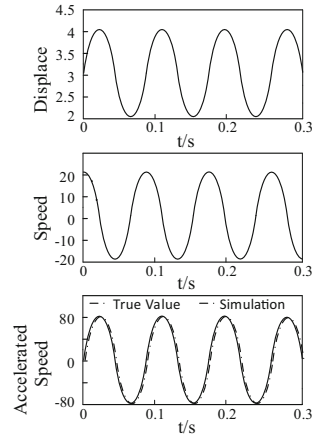


Fig. 6 Tracking error and output error under unit step input

Fig. 7 Actual simulation results



4 Conclusion

Through research, this paper has the following conclusions:

1. The use of non-inductive coil differential compensation structure and temperature correction table method can solve the sensor temperature drift problem, so that the sensor in the range of 20–85 °C temperature drift can be controlled at about 1% F. S.
2. Using NITD to deal with the change of the vertical gap signal twice, compared with the existing acceleration signal, we can get a simple fault self-test method for the acceleration sensor.

Acknowledgments National Key R&D Program of China (2016YFB1200600)

References

1. Jun W, Wenwu Z, Lu L (2009) Study on suspension gap sensor of high speed maglev train. *J Testing Technol* 05:466–470 (in Chinese)
2. Wensen C, Xiangming W, Jun W et al Gap displacement sensor for high-speed maglev train: China, ZL200410024637. 0 [P]. 2006–12 20 (in Chinese)
3. Lu L, Wenwu Z, Jun W Research on high speed maglev train levitation gap signal processing. *Chin J Sci Instrum* 29(9):2001 2004 (in Chinese)
4. Jun W, Lu L, Shujiang F et al (2004) Design of the coils for a gap sensor. *Chin J Sens Actuators* 17(3):512–515 (in Chinese)
5. Qihua F, Pei M, Xinqi T (1997) Autocompensation of temperature for eddy current sensor. *Chin J Sens Actuators* 18(2):198–201 (in Chinese)
6. Matveevich F, Mednikov S Differential eddy-current transducer. United States, US 20020121895 A1 [P]. 2002-10-07

7. Huaqing L, Zhiye L (1995) Germany maglev train. University of Electronic Science and Technology Press, Chengdu, pp 17–30 (in Chinese)
8. Lu L (2002) Design and research of high-speed maglev train gap sensor. National University of Defense Technology, Changsha (in Chinese)
9. Daqi Z (2004) Electronic equipment fault diagnosis principle and practice. Electronic Industry Press, Beijing (in Chinese)
10. Jingqing H, Wei W (1994) Nonlinear tracking differentiator. Syst Sci Math Sci 14(2):177–183 (in Chinese)
11. Xinhua W, Zengqiang C, Zhuzhi Y (2003) Nonlinear tracking differentiator with high speed in whole course. Control Theory Appl 20(6):875–878 (in Chinese)

Superiority of Unit Serial Method Used in CRTS III-Type Pretensioning Ballastless Track Slab Production

Baoqun Wang, Lei Pei, Qikai Ai, Hongbin Zheng and Wei Si

Abstract High-speed railway ballastless track two-way pretensioning prestressed track slab by the unit serial method is a new type of track slab prestressed system developed in China. The research group developed the track slab production equipment and tooling and formed a 24 h production process. This is the first time in the world to achieve large-scale production of two-way pretensioning prestressed track slab through production lines. Also, this method has been tested and put into production in Shandong Linq. Compared to the traditional matrix method of the track slab production process, this new method greatly improves the production efficiency and reduces the manufacturing cost of track slab. This can fulfill the high volume demand of track slab raised by “The Belt and Road Initiative” strategy.

Keywords CRTS III-type pretensioning ballastless · Matrix method
Production process · Unit serial method

1 Introduction

It is very common that high-speed railway uses ballastless track because of its characteristics of high smoothness, high stability, high precision and less deformation, and so on. Therefore, track slab becomes a very important part of high-speed railway track bed. The high speed of the train makes a higher demand for the smoothness of the track; the production of high-precision track slab has been a key link in high-speed railway construction technology. In 2003, China introduced the ballastless track technology from Germany and then carried out CRTS

B. Wang (✉) · W. Si
ShanDong JiaoTong University, Jinan, China
e-mail: 2009swreal@tongji.edu.cn

L. Pei · Q. Ai
Shandong Hi-Speed Rail Transit Group Co., Ltd., Jinan, China

H. Zheng
China Railway 23rd Group Co., Ltd., Beijing, China

I-type, CRTS II-type slab structure research and development. In 2011, China has formed a complete set of technology including design, manufacture, construction, repair, and maintenance of high-speed railway CRTS III-type ballastless track. In 2012, matrix method of the production process was developed and then widely used in high-speed railway.

Jinan–Qingdao High-speed Railway is an important part of the national “four vertical and four horizontal” HSR networks. The total length of this line is 307.9 km and this line is designed to travel at speeds of up to 350 km/h. It requires about 70,000 CRTS III-type track slabs. Faced with such a large number of track slab demand, the traditional matrix method of production process has been unable to fulfill the rapid development of China’s high-speed railway needs. However, if we use unit serial method, the degree of automation will be greatly improved and production efficiency will be increased by 1.5 times. Remote monitoring can be realized in production and employee demand can be reduced to two-thirds.

The implementation of unit serial method on CRTS III-type track slab will drive promotion and demonstration impact on traditional industries. The improvement of unit serial method aligns with China’s “the Belt and Road Initiative” and will enhance the international competitiveness of high-speed rail technology.

2 Advantage and Disadvantages of Traditional Matrix Method in Track Slab Production

At present, CRTS III-type track slabs are produced using matrix method universally in China, which is the traditional prestressed concrete components bench prefabrication method. Here is a brief introduction of this process: Each tensioning bench is set with 8 track slab models, arranged in a 2×4 matrix. Decide the length of prestressed tendons of track slab and cut. Do not expose the side of track slab at the end, both ends are connected to the tension bar by thread. The connectors are set between the corresponding tensioning bars of adjacent templates to achieve the continuity of the tension between the tension beam at the tension end and reaction wall at the fixed end. Then overall tension and synchronized release are made to prestressed tendons. Pour concrete and vibrate by model, followed with overall conservation to matrix unit.

Matrix method is a relatively mature technology. However, from the perspective of economic performance, matrix tension pits occupy a large area of land. Meanwhile, after the completion of the project, the reinforced concrete structure needs to be dismantled which takes a lot of manpower and material resource. From the perspective of equipment configuration, the manufacturing of track slab takes unit pit as the basis, which means many of the same equipment need to be configured. During the product steaming period, the equipment are in a state of idle which is waste. From the perspective of capacity, the manufacturing cycle of track slab is 20–22 h, which causes uncertain working period, the employees are difficult to recover mentally and physically which may increase security risk. From the perspective of administration,

the operation time of each routine is relatively short and not easy to fix personnel at a fixed position, actually one employee works for more than one position. This situation increases the difficulty for people to enhance technical professionals. Overall, matrix method has the disadvantages of low equipment utilization, large area demand of land, the difficulty of staffing assessment, higher security risks, high depreciation cost of fixed assets, and other shortcomings.

Taken together, the traditional matrix method of the production process will not be able to fulfill the demand of the rapid development of China's high-speed railway, we need to further improve the production process and improve the level of industrial production.

3 The Superiority of Unit Serial Method

HSRs such as Jinan–Qingdao High-speed Railway have huge demands of track slabs. If all CRTS III-type track slabs are produced using matrix method, 24 tension pits need to be built which occupy more than 40,000 ft², the production period will be about 2 years and at least 400 construction workers need to be organized. And after the completion of the project, the construction facilities such as tension pits and factory buildings are in the scrapped state.

Regarding this situation, our group proposed to research unit serial method in pretensioning prestressed track slab production. The main purpose is to improve the mechanization and automation of manufacturing, improve production efficiency and reduce the cost of track slab production. The way is to let the track slab model bear the tensile stress and flow in the orbit and meanwhile keep each station in the process fixed. In November 2016, the world's first high-speed railway CRTS III-type pretensioning prestressed concrete track slab production line was born in Shandong Linq. The principle of its technology has been verified, and the stability of related equipment was further improved.

Using unit serial method, the degree of automation will be greatly improved and production efficiency will be increased by 1.5 times. Remote monitoring can be realized in production and employee demand can be reduced to two-thirds. The production workshop occupies more than 30,000 ft². This project played the role of promotion and demonstration to the traditional industries. Also, unit serial method is an advanced technology which can enhance the international competitiveness of high-speed rail technology. It aligns with national "The Belt and Road Initiative" strategy.

3.1 Innovation of "Unit Serial Method" Manufacturing Technique

After analyzing previous production process of prestressed concrete sleeper and track slab, combined with the structure characteristics of two-way pretensioning

prestressed track slab, “unit serial method” production process was innovated, which is “put the CRTS III type track slab prefabricated concrete model on the wheel-rail transport flatbed (multiple sets) and then orderly enter different work areas along the track, with the completion of model shaping, steel bar and prestressed tendon installation, tension and anchorage of prestressed tendons, concrete pouring and vibrating, component maintenance, relax and stripping of prestressed tendons, quality inspection of finished product and other processes.”

The production process of CRTS III-type track slab production line, as shown in Fig. 1, mainly includes shaping, steel bar and prestressed tendon installation, tension and anchorage of prestressed tendons, concrete pouring and vibrating, component maintenance, prestressed relaxation, component stripping and migration, and other processes. The operation mode of this production process is that the models are lined up on the operation line, and flow in accordance with the time beat driven by the traction equipment, 12 different working units operate at the same time. A closed loop is formed between both the front and back ends with the steam line.

3.2 Introduction to Key Production Process of Track Slab

3.2.1 Model Design Technology of CRTS III-Type Track Slab

The track slab model is subjected to tensile stress. The slabs flow on the track and keep each process station fixed. This can improve the mechanization and automation of manufacturing, which also improves production efficiency and reduces manufacturing costs. Correspondingly, there are higher requirements for the deformation control of the model.

The track slab model is subjected to a 192t transverse tensile prestress, and a 128t longitudinal tensile prestress. The deformation of the model bottom should be kept $\leq \pm 0.5$ mm. Considering the lift capacity of operation vehicle and plant construction costs, the total weight of the model should be ≤ 12 t. Considering that the model needs to weld with bending beam and work under the vibration frequency of 80–120 Hz, selection of high-strength wear-resistant steel plate for thermal processing becomes one of the key technologies.

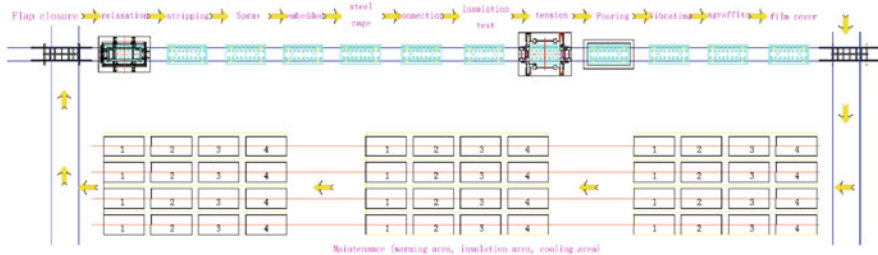


Fig. 1 CRTS III-type track slab assembly line

3.2.2 Prestressed Tendons Tension

There are 24 transverse and 16 longitudinal prestressed tendons in one CRTS III-type track slab. The production standard requires the tensile force of each prestressed tendon to be 80 kN. Tension work is divided into two steps.

The pretensioning of the prestressed tendons is carried out in the first step, and the pretension value is 30% of the control value. Check the position of prestressed tendons and anchor plates after pretensioning. The vertical position deviation of prestressed tendons shall not exceed 2 mm.

The second step is to carry out the final tension of the prestressed tendons. The tensioned beam should be used to ensure the synchronization of tension and uniformity of tension force which can avoid warping of the track slab caused by uneven tension.

Double control mode of force and displacement is used to prestressed tendons, which is, take tension measurement as the main control mode and take or monitor the elongation of prestressed tendons as an auxiliary control. The measured total tension value shall not deviate by more than 3% of the design value. The measured tension value of single prestressed tendon shall not deviate by more than 10% of the design value. This can avoid the quality problems of the track slab caused by insufficient tension.

The tensioning equipment no longer forces the foundation pit and the tensioning base but bear the tension force by itself. Lock the tension value when the tension reaches the right place and the mold will withstand tension.

3.2.3 Prestressed Tendons Release

The compressive strength of the test specimen of the track slab is not less than 48 MPa and the elastic modulus is not less than 3.4×10^4 MPa. The prestressed tendons release can be carried out when the temperature difference between the track slab surface and the environment is more than 15 °C. During the release, the steamer cover cannot be opened. The bench should be kept covered. First, loosen the jack lock nut, then the operator issues a release command through the computer in the central control room. The release begins and the jack cylinder retracts to the bottom of the cylinder.

The overall release process is the end of the track slab production process, but also is the stage with the highest technical requirements. In this process, the working mode of the two-way relaxation system is a reasonable natural relaxation of the tensile elongation. If the release of stress cannot be synchronized, it will produce uneven stress inside the track slab. This stress can cause the internal friction and residual stress of the plate concrete, which is not conducive to the durability of the track slab. Prestressed tendons should use the overall synchronized release method and super-tension is strictly prohibited. In the process of relaxation, synchronized movements of the vertical and horizontal 8 jacks must be ensured to reduce the impact of alternating stress on concrete. The rate of release should not

exceed 8 kN/s. Also, the “lubrication” mechanical lock is designed. The bearing surface of the lock is contacted by ball bearing, which greatly reduces the external force required to unlock, and therefore can completely avoid the “super-tension”.

4 Engineering Application and Conclusion

CRTS III-type ballastless track has become the preferred structure type in high-speed railway construction in China, such as Zheng Xu high-speed railway, Ji Qing high-speed railway, Lu Nan high-speed railway, and Xu Lian high-speed railway, and other projects that have been completed construction or under construction. The total demand for the next 5 years will reach about 400 thousand pieces and estimated cost can reach about 2.8 billion. The domestic production market of this product prospects, with China’s “the Belt and Road Initiative”, CRTS III-type ballastless track structure will have the strong international competitiveness and the international production market of this product will be broad.

References

1. TB 10002.3-2005 Code for Design of Reinforced Concrete and Prestressed Concrete Structure of Railway Bridges and Culverts (in Chinese)
2. China Academy of Railway Sciences (2013) Design of CRTS W slab type ballastless track slab for newly built Zhengzhou-Xuzhou railway passenger dedicated line. China Academy of Railway Sciences, Beijing (in Chinese)
3. Shuguang Z (2010) Study on technology system and system integration method of China high-speed railway. In: Proceeding of the ASME joint rail conference, 2010, JRC, pp 501–506
4. Jian She T (2005) No. 160, interim standard for constructional quality acceptance s of track engineering of passenger dedicated railway line (in Chinese)
5. An GD (2009) Technical standards and quality control of ballastless track for high speed railway. China Railway Publishing Press, Beijing (in Chinese)
6. China Academy of Railway Sciences (2010) Design of CRTS slab type ballastless track slab for newly built Panjin—Yingkou railway passenger dedicated line. Beijing: China Academy of Railway Sciences, 2010 (in Chinese)
7. TJ/GW 118-2013 Interim Technical Conditions for Pre-Tensioning Prestressed Concrete Track Plate of CRTS Slab Type Ballastless Track for High Speed Railway (in Chinese)
8. TB/T3193-2008 Technical Conditions for Clip Type Anchor, Fixtures and Connectors Used for Prestressed Tendon of Railway Engineering (in Chinese)

The Test Plan Design of Corrosion Fault Repeat and Material Selection for a Type of PCB

Xu Wang, Shaohua Du and Yingying Yuan

Abstract The failure analysis (FA) of the PCB indicates that the connecting line of relay pin and inner structure is broken for Cl corrosion. To verify the above conclusion and find the introducing path of Cl, the accelerated corrosion test profile is designed in this paper. The failure mode is repeated successfully after six cycles and Cl is found on the connecting line of relay obviously. Besides, the different design versions of PCBs with different placed orientations are compared in this test. And the better version of PCB and better placed orientation are confirmed consequently. The following works like material replacement and PCB design refinement can benefit from the results of this paper.

Keywords Corrosion · Failure recurrence · Material replacement
Test profile design

1 Introduction

A type of printed circuit board (PCB) exposed many circuit break failures after a period of running time. The FA work indicates that Cl corrosion occurs in the relay pin and inner structure. Besides, the FA work also conducted for different process stages of PCBs. The conclusion of above work shows that Cl is introduced by wave soldering process and is being aggravated in the stage of warehousing and usage. The connecting lines are broken and the PCBs fail ultimately [1].

Thus, in this paper, an accelerated corrosion test profile is designed to verify the conclusion of FA and to support the material replacement work of the relay.

X. Wang · S. Du (✉)
CRRC ZIC Research Institute of Electrical Technology
and Material Engineering, Zhuzhou 412001, Hunan, China
e-mail: dush@csrzc.com

Y. Yuan
Hunan CRRC Times Electric Vehicle Co., Ltd., Zhuzhou 412007, Hunan, China

2 The Test Plan Design

2.1 Stress Type

According to the conclusion of FA, the PCBs of different process stages are selected randomly and Cl is introduced in the relay pin and inner structure. Cl will react with metal [2]. Besides, when the salts come into the inner material, the battery system of “low potential, electrolyte solution and high potential” will appear. With the anodic dissolution, the corrosive compound is formed [3].

The key stresses which affect the react speed of the Cl and the metal are temperature, humidity, and Cl element concentrations [4]. In this paper, the Cl concentrations are not controlled to ensure the test samples be kept consistent with the usage site. Thus, the stress variables are temperature and humidity.

2.2 The Sample Size and Orientation

In this test, the chief objective is to repeat the corrosion failure of relay pin. Besides, to identify which is better between the improved version of B and the improved version of C. Finally, to conduct the mean life prediction of A-version of PCB. 10 PCBs of A-version, 4 PCBs of B-version, and 4 PCBs of C-version are chosen randomly from the storehouse. Half of each type is placed in normal orientation and others are reversed.

2.3 Design for Test Profile

To improve the efficiency of the accelerated corrosive test, we need both factors such as high temperature and high humidity. Temperature humidity bias (or 85/85) life test is used to assess the PCBs' intensity against stresses. We can continue to improve the temperature value to increase the test efficiency which is the thought of the HAST [5].

From the above, the preparing test that uses enhancement stress is conducted to verify that the 85/85 test will not destroy the PCBs [6]. Based on the 85/85 test, the temperature is increased to 90 °C, if the PCBs is normally operated, the final value of stresses is confirmed.

The designed preparing test profile is shown in Fig. 1 and accelerated corrosive profile is shown in Fig. 2.

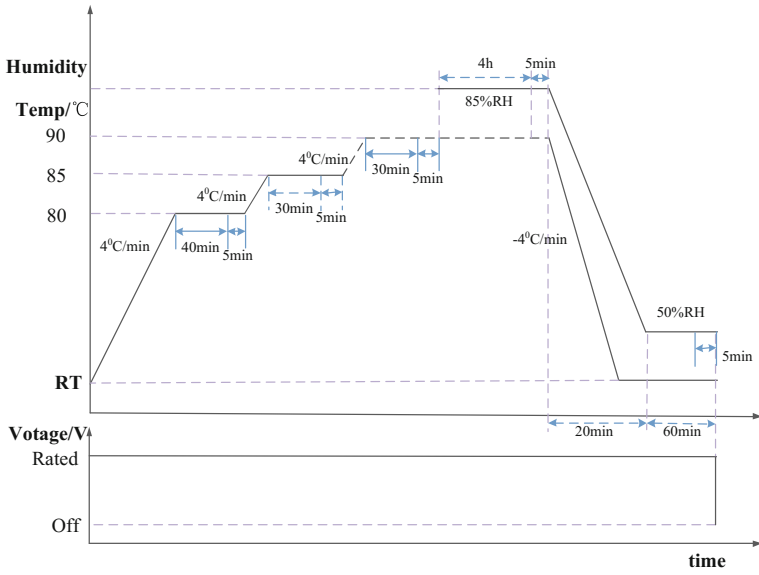


Fig. 1 The preparing test profile

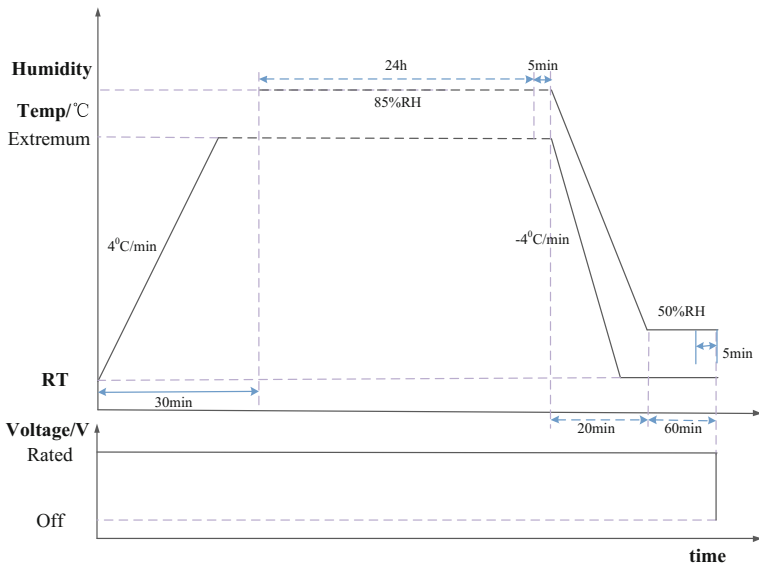


Fig. 2 The accelerated corrosive profile (one cycle)

2.4 The Mean Life Prediction

The methods of life prediction are based on the Eyring equation because two stresses are involved [7]. The acceleration factor is given by Eq. 1, developed by Peck [8]. Where RH_u and RH_s are the relative humidities at use conditions and test conditions, and T_u and T_s are the use and test temperatures. The other values, E_a and n , are constants derived from testing while k is Boltzmann’s constant.

$$AF = (RH_u/RH_s)^{-n} e^{(E_a)/k(1/T_u-1/T_s)} \tag{1}$$

According to the failure times of PCBs observed in the test, the MTTF of PCBs can be derived as Eq. 2. Where t_i is failure time of PCBs, respectively, and n is the sample size.

$$MTTF = \frac{\sum_{i=1}^n AF * t_i}{n} \tag{2}$$

3 Test Result and Analysis

3.1 The Test Result

The A-version PCB is a failure when the preparing test is continued until 90 °C, 85% RH. The failure reason is that short-circuit happened in the chip. Thus, the accelerated corrosive values in the test profile can be confirmed as 85/85.

The accelerated corrosive test is conducted for 6 cycles. The failure information is listed as shown in Table 1.

A failure of non-voltage output happens at the fifth cycle of B-version PCB and the reason is that short-circuit happened in the chip. C-version PCBs can still

Table 1 The failure information in accelerated test

	Failure number of A-version PCB (orientation)	Failure number of B-version PCB (orientation)	Failure number of C-version PCB (orientation)
First cycle	0	0	0
Second cycle	1 (reversed)	0	0
Third cycle	0 (reversed)	0	0
Fourth cycle	1 (reversed)	0	0
Fifth cycle	1 (reversed)	1 (normal)	0
Sixth cycle	2 (1 reversed and 1 normal)	0	0
Sum	5	1	0

operate normally. Four reversed A-version PCBs cannot work normally which account for 80% of total number. One normally replaced A-PCB cannot work normally which accounts for 20% of total number. All the failed PCBs are non-voltage output. Removing the relays from the failed A-version PCBs, we can see that the connecting lines of relay pins and inner structure are broken for Cl corrosion (see Fig. 3). The connecting lines of normal placed PCBs are cleaner relatively (see Fig. 4). Thus, the failure model in the use spot is repeated successfully and the effectiveness of B-version and C-version is verified. B-version PCB is not failed for the break of connecting line, but the pin of the relay is corrosive seriously. As for C-version PCB, the connecting line is hidden in the glue and is not corrosive correspondingly (see Fig. 5).

Fig. 3 The corrosive appearance of the connecting line of reversed A-version PCB

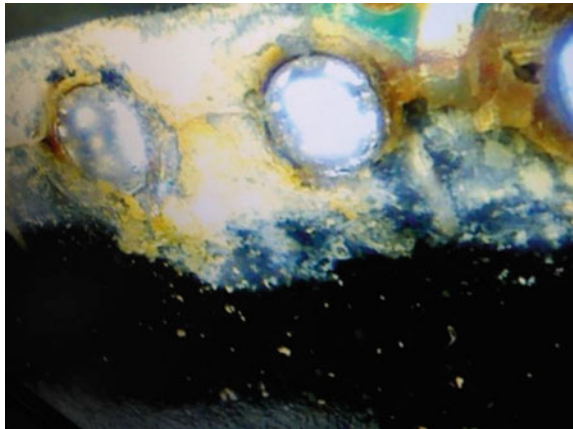


Fig. 4 The corrosive appearance of the connecting line of normal replaced A-version PCB

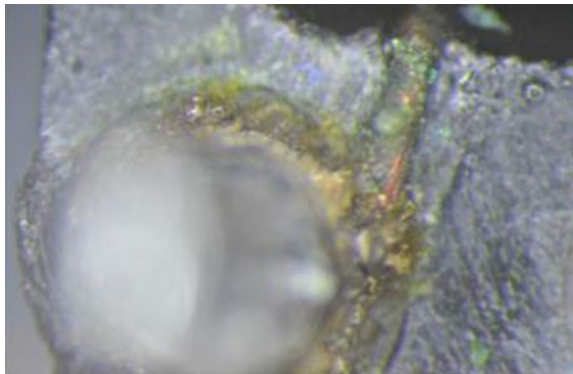
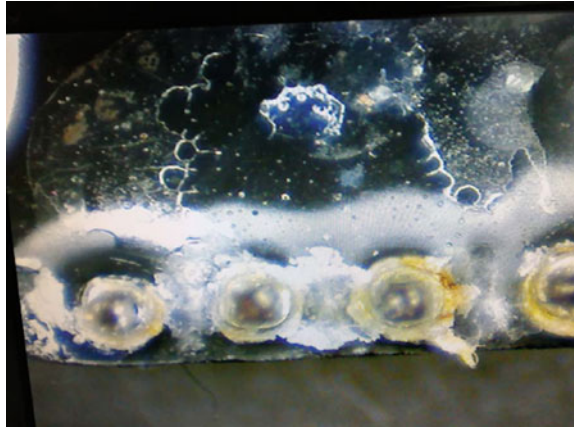


Fig. 5 The corrosive appearance of the connecting line of C-version PCB



3.2 The Test Result Analysis

- (1) The conclusion of FA is verified, that is, under the temperature and humidity, the connecting line of the relay is broke because of the Cl corrosion. Accordingly to the failure data in test and Peck model in Sect. 2.4, we can compute the mean life of A-version PCB is about 6.78 years.
- (2) Due to reversed placed, the cavity of the relay is filled with the moisture of A-version PCB. The speed of electrochemical corrosion is accelerated when there is a balance of four factors, that is, humidity, oxygen concentration, Cl, and temperature.
- (3) The diameter of connecting line is enlarged and a better coating process is used for B-version PCB. The speed of electrochemical corrosion is slowed down, however, there is still corrosive severely.
- (4) The connecting line is hidden in the glue and is not corrosive correspondingly of C-version PCBs.

4 Conclusion

According to the conclusion of FA, an accelerated corrosion test plan is designed. The conclusion of FA is verified successfully. The route of transmission of Cl is affirmed and the failure cause is found according to the placed orientation of PCBs. Besides, the effectiveness to resist corrosion between B-version and C-version is verified. The following works like material replacement and design refinement can benefit from the results of this paper.

References

1. Yuan Y (2017) The failure analysis method and applications on a type of PCB. Special Issue Electr Veh Technol 15(6):32–36
2. Zhao P, Pecht M (2005) Mixed flowing gas studies of creep corrosion on plastic encapsulated microcircuit packages with noble metal pre-plated lead frames. IEEE Trans Device Mater Reliab 5(2):268–276
3. Chen P (2014) Overview of salt spray test technology. Electron Prod Reliab Environ Test 11:24–29
4. Du Y, Guan G (2005) The research of corrosion resistance for ICs. Electron Prod Reliab Environ Test 4:30–33
5. Huang W, Bai L (2014) The influence of HAST for plastic encapsulated microelectronics devices. Environ Adaptability Reliab 4:26–29
6. Yang S (2006) Evaluation of non-operating reliability for electronic components. Guangzhou Faculty of Light and Chemical, Guangdong University of Technology
7. Galloway JE, Miles BM (1997) Moisture absorption and desorption predictions for plastic ball grid array packages. IEEE Compon Packag Manufact Technol 20(3):274–279
8. Peck DS (1986) Comprehensive model for humidity testing correlation. Reliability Physics Symposium, IEEE. 44–50

Multi-hop Communication Protocol Optimization for the Linear Wireless Monitoring Network

Xiaoping Ma, Honghui Dong, Limin Jia, Yong Qin and Ruhao Zhao

Abstract The lifetime and latency of the railway infrastructure wireless monitoring system are vital to guarantee the stability of the monitoring system and the safety of the railway operation. In this paper, we proposed the multi-hop protocols to optimize the utility of the railway infrastructure wireless monitoring system. The lifetime maximum, latency minimum, and utility maximum models are all designed in this paper. The simulation results show that the complex objective model to maximize the utility of the system is superior to the two single-objective models. The performance in lifetime is almost to the lifetime maximum model and in latency is slightly longer than the total energy consumption minimum model. In the railway system, the two factors are all very important, the utility maximum protocol is more suitable to be used in this field.

Keywords Railway monitoring system · Multi-hop · Wireless network
Utility maximum

1 Introduction

Wireless monitoring network is applied widely in the railway infrastructure inspection system [1]. The stability and efficiency of the linear wireless monitoring network are most related to the safety and operation of the railway system. For the railway infrastructure monitoring system, the stability of the wireless network should be guaranteed to make sure the inspection data could be transmitted to the data centre successfully on one hand. Due to the high-speed of the trains, the real time of the information transmission should be taken into consideration at the same time. The stability of the wireless network is related to the lifetime of the sensor and sink nodes, the sink nodes in charge of receiving the information inspected by the

X. Ma · H. Dong · L. Jia (✉) · Y. Qin · R. Zhao
State Key Lab of Rail Traffic Control and Safety,
Beijing Jiaotong University, 100044 Beijing, China
e-mail: jialm@vip.sina.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_94

929

sensors and then transmitted them to the base station. However, the energy storage, information processing and transmitting ability of the infrastructure condition sensing and communicating units are all limited, an energy-efficient protocol is needed to prolong the lifetime of the system. Meanwhile, the real-time demand of the wireless network is the other important factor in the design of the protocol.

In railway scene, the deployment of the railway infrastructure is linear along the rail, the sensor nodes are all installed on the infrastructure to inspect the service condition. The sink nodes are deployed along the rail to transmit the inspected information to the base station (BS). For the linear communication structure between the sink nodes and the base station, the distances from the sink nodes to the base station and the data packets need to be transmitted are all different, it will influence the energy consumption and even the lifetime of the sink nodes. In order to solve the problem, there are a lot of researchers aimed to propose an effective strategy to balance the energy consumption among the nodes [2, 3].

The energy consumption of the sensor/sink nodes is related to the distances from the source nodes to the destination node and the volume of the data packets. In order to save the energy consumption, Yildiz [4] proposed a protocol at the node level to trade off the communication and computation energy consumption to improve the lifetime about 20% more than the network-level protocol. Incebacak [5] proposed a strategy to save the energy by compressing/decompressing the data dynamically according to the adjustment of the transmission power. The two above methods are all based on the multi-hop routing method as shown in [6]. The method in [7] aimed to maximize the lifetime of the system while [8] working on minimizing the total energy consumption of all sensor/sink nodes. The simulation results show that the strategy in [7] is superior to in [8] in the energy efficiency while poorer performance in real time. In this paper, we proposed a modified multi-hop protocol to maximize the lifetime of the system and minimize the latency at the same time.

The remaining of the paper is organized as follows: The structure of the monitoring system and energy consumption model is introduced in Sect. 2. Section 3 describes the modified multi-hop model proposed in this paper. The simulation results are shown in Sect. 4. We conclude our work in Sect. 5.

2 The Structure and Energy Consumption Model of the Monitoring System

To improve the utility of the railway wireless monitoring system, we need to design an energy-efficient protocol to prolong the lifetime and minimize the latency based on the structure of the monitoring system. In this section, we build the structure of the monitoring system and then design the energy consumption model which will be used to optimize the system utility in the following section.

2.1 The Structure of the Monitoring System

In the railway infrastructure monitoring systems, the inspected information of the infrastructure is very important to estimate and predict the service condition of the railway system. It is vital to guarantee the stability and real time of the wireless network to ensure the operation safety of the railway. As shown in Fig. 1, the monitoring system is composed of three layers, railway infrastructure-sensing network, inspection information communication network and condition estimation and prediction centre. In the first layer, the information of the infrastructure (rail condition, slope condition, catenary condition, foreign invasion condition, *e.t.t.*) are collected by the sensors installed in the corresponding position. Then, the data packets are transmitted to the base station through the sink nodes. Finally, the information proceeds in the data centre and used to estimate and predict the service condition of the infrastructure.

2.2 The Energy Consumption Model of the Monitoring System

In the communication layer, the sink nodes and the base station are deployed in line and each sink node in charge of transmitting the information in the special region. There are several sink nodes and one base station in the communication layer.

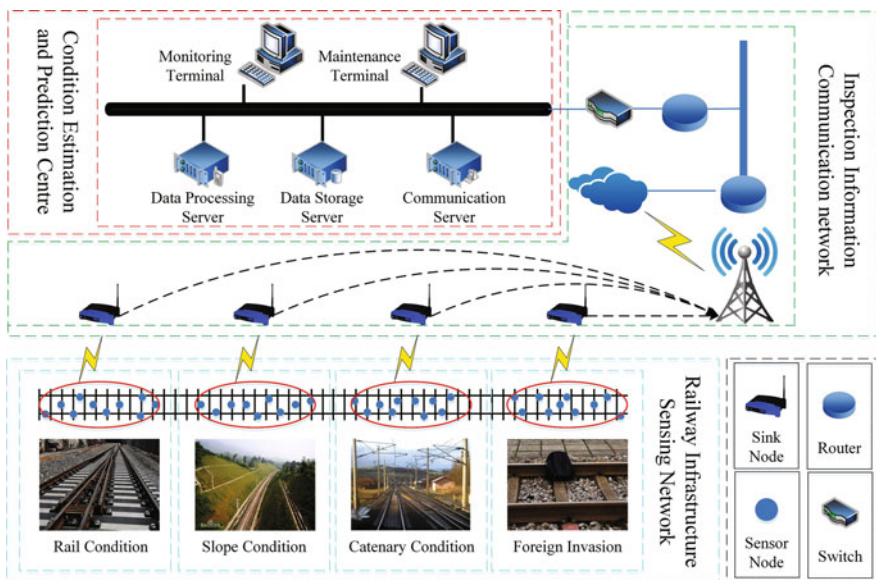


Fig. 1 The structure of the railway infrastructure monitoring system

Due to the linear deployment of the communication units, the distances from the sink nodes to the base station are different. Furthermore, the sizes of data packets to be transmitted by each sink nodes are all different because of different monitoring region and tasks. The above two factors will influence the energy consumption of the sink nodes. With the single-hop strategy that the sink nodes communicate with the base station directly will lead to the unbalanced of the energy consumption and even cause the instability of the communication network. Thus, the multi-hop strategies are introduced in the paper.

In Fig. 2, the communication units are the sink nodes, $S_i(t)$ represents the data packets received from the sensor layer, f_{ij} represents the data packets transmitted among the sink nodes and the base station. As shown in Fig. 2, in the communication structure of the multi-hop strategy, all the sink nodes receive the data packets from the sensing layer and then transmit them to the base station. In order to maximize the lifetime of the system, the data packets will be divided into several small packets and then be transmitted forward to other sink nodes, who in charge of forwarding them to the base station.

In this paper, the raw information is transmitted among the sink nodes and then to the base station in the end. We assume that there are N sink nodes and 1 base station in the communication network. The energy consumption of the i th sink node consists of two parts $E_{ri}(t)$ and $E_{ti}(t)$.

$$E_{coi}(t) = E_{ri}(t) + E_{ti}(t), \quad i = 1, 2, \dots, N \tag{1}$$

The $E_{ri}(t)$ represents the energy consumed to receive the data from the sensor layer and other sink nodes for the i th sink node at t th round is calculated by

$$E_{ri}(t) = (S_i(t) + \sum_j f_{ji}(t)) * E_{ele}, \quad i \in [1, N], j \in [2, N] \tag{2}$$

S.T. $i < j$
 $f_{ji}(t) \geq 0,$

where E_{ele} is the energy consumption of the circuit to process 1 – bit data. $S_i(t)$ represents the data packets received from the sensor layer at t th round. $f_{ji}(t)$ represents the data packets received from the j th sink node at t th round. The first

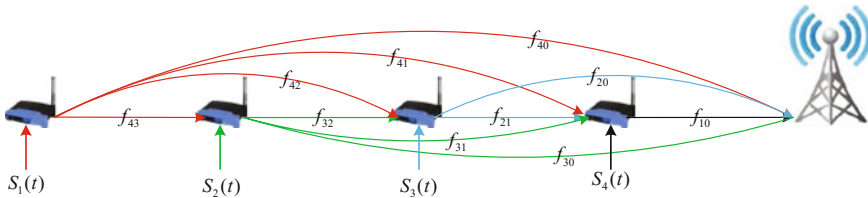


Fig. 2 The communication structure of the multi-hop strategy

constraint implies that the data packets are transmitted forward toward the base station, the second constraint says that the size of the small packets is all non-negative.

The $E_{ii}(t)$ represents the energy consumed to transmit the data packets forward to the i th sink node at t th round is calculated by

$$E_{ii}(t) = \begin{cases} \sum_j f_{ij}(t) * E_{ele} + \xi_{fs} * \sum_j f_{ij}(t) * d_{ij}^2, d_{ij} < d_0 \\ \sum_j f_{ij}(t) * E_{ele} + \xi_{mp} * \sum_j f_{ij}(t) * d_{ij}^A, d_{ij} < d_0, \end{cases} \quad i \in [1, N], j \in [0, N] \quad (3)$$

where ξ_{fs} and ξ_{mp} are the energy coefficient related to the distance between the source and the destination. $f_{ij}(t)$ represents the data packets transmitted to the j th sink node at t th round. The first constraint implies that the data packets are transmitted forward toward the base station, the second constraint says that the size of the small packets is all non-negative.

For each sink node, the input size of the data packets is equal to the output ones, it means that there are no data packets generated or disappeared during the communication process, as shown in

$$\sum_j f_{ij}(t) = \sum_j f_{ji}(t) + S_i(t), \quad i \in [1, N]. \quad (4)$$

Based on the energy consumption model of the multi-hop strategy, we design the lifetime maximizing and latency minimizing protocol for the monitoring system so as to improve its utility.

3 The Multi-hop Optimal Protocols

As shown in the last section, the energy consumption of the sink nodes is mostly related to the distances between the base station and the size of the data packets to be transmitted. In order to improve the utility of the monitoring system, we need to maximize the lifetime of the sink nodes on one hand and minimize the latency at the same time. The multi-hop strategy is an effective way to allocate the transmission task among all sink nodes to catch up on the unbalance energy consumption caused by the differences in the distances.

3.1 The Lifetime Maximum Optimization Model

The structure of the multi-hop communication among the sink nodes are introduced in the above section, the energy consumption models are all designed. In the

railway infrastructure wireless monitoring system, each sink node is in charge of the information transmission in the special monitoring region. Of course, the lifetime of all the sink nodes is co-equal and important. In this paper, the lifetime of the system is defined as the minimum communication rounds of all sink nodes until its energy exhausted.

Figure 3 formulates the model to maximize the minimum residual energy of all the sink nodes in each communication round for the objective to maximize the lifetime t .

The second constraint implies that the residual energy of all the sink nodes should be more than zero so as to complete the transmission task. The sixth constraint says that the total residual energy in t th communication round should be more than the minimum energy consumption demands in t th communication round, if not the transmission process will be shut down and the lifetime is t .

3.2 The Latency Minimize Optimization Model

In the railway infrastructure wireless monitoring system, the inspected information should be transmitted to the data centre as soon as possible and then used to estimate and predict the condition of the infrastructure so as to avoid the accident caused by the failure of this equipment. The latency as the other important factor should be taken into consideration. In the multi-hop strategy, the TDMA is used to guide the communication process, and the number of hops is mostly related to the latency of the system. In this paper, the number of hops is used to represent the system latency.

The simulation results show that minimize the total energy consumption of all sink nodes will decrease the number of hops mostly. Maximize the total residual energy of all sink nodes in each communication round to minimize the system latency as shown in Fig. 4.

$$\begin{aligned} & \text{Maximize } \min(E_{rei}(t), i \in [1, N]) \\ & \text{Subject to: } f_{ij}(t) \geq 0 \text{ for } i \in [1, N] \text{ and } j \in [0, N] \\ & \quad E_{rei}(t) \geq 0 \text{ for } i \in [1, N] \\ & \quad E_{rei}(t) = E_{rei}(t-1) - E_{coi}(t) \text{ for } i \in [1, N] \\ & \quad E_{coi}(t) = E_{ri}(t) + E_{ii}(t) \text{ for } i \in [1, N] \\ & \quad \sum_j f_{ji}(t) + S_i(t) = \sum_j f_{ij}(t) \text{ for } i \in [1, N] \\ & \quad \sum_{i=1}^N E_{rei}(t) \geq \min \sum_{i=1}^N E_{coi}(t+1) \end{aligned}$$

Fig. 3 The linear programming model to maximize the minimum residual energy of all sink nodes

$$\begin{aligned}
 & \text{Maximize } \sum_{i=1}^N E_{rei}(t) \\
 & \text{Subject to: } f_{ij}(t) \geq 0 \text{ for } i \in [1, N] \text{ and } j \in [0, N] \\
 & \quad E_{rei}(t) \geq 0 \text{ for } i \in [1, N] \\
 & \quad E_{rei}(t) = E_{rei}(t-1) - E_{coi}(t) \text{ for } i \in [1, N] \\
 & \quad E_{coi}(t) = E_{ri}(t) + E_{ii}(t) \text{ for } i \in [1, N] \\
 & \quad \sum_j f_{ji}(t) + S_i(t) = \sum_j f_{ij}(t) \text{ for } i \in [1, N] \\
 & \quad \sum_{i=1}^N E_{rei}(t) \geq \min \sum_{i=1}^N E_{coi}(t+1)
 \end{aligned}$$

Fig. 4 The linear programming model to maximize the residual energy of all sink nodes

3.3 The Utility Maximum Optimization Model

In the railway infrastructure wireless monitoring system, in order to maximize the utility of the system, we need to prolong the lifetime of the system and decrease the latency at the same time. In the paper, we combine the above two models together and proposed a new model to optimize the system utility as shown in Fig. 5.

In this model, α, β are the coefficients to balance the importance and influence of the lifetime and latency to the system. They will be adjusted based on the demands of the monitoring system. In the railway wireless monitoring system, the two factors are all very important, so we set $\alpha = \beta = 0.5$ in this paper.

$$\begin{aligned}
 & \text{Maximize } (\alpha * \sum_{i=1}^N E_{rei}(t) / N + \beta * \min E_{rei}(t)) \\
 & \text{Subject to: } f_{ij}(t) \geq 0 \text{ for } i \in [1, N] \text{ and } j \in [0, N] \\
 & \quad E_{rei}(t) \geq 0 \text{ for } i \in [1, N] \\
 & \quad E_{rei}(t) = E_{rei}(t-1) - E_{coi}(t) \text{ for } i \in [1, N] \\
 & \quad E_{coi}(t) = E_{ri}(t) + E_{ii}(t) \text{ for } i \in [1, N] \\
 & \quad \sum_j f_{ji}(t) + S_i(t) = \sum_j f_{ij}(t) \text{ for } i \in [1, N] \\
 & \quad \sum_{i=1}^N E_{rei}(t) \geq \min \sum_{i=1}^N E_{coi}(t+1) \\
 & \quad 1 \geq \alpha, \beta \geq 0 \\
 & \quad \alpha + \beta = 1
 \end{aligned}$$

Fig. 5 The linear programming model to maximize the utility of the system

4 Performance Evaluation

In this section, the performances of the three communication protocols proposed in this paper are tested and compared via computer simulation. In the simulation, we assume that there are 4 sink nodes with the index [1, 2, 3, 4] and 1 base station with index 0. In each communication round, the sizes of the data packets of all sink nodes received from the sensor layer are equal $S_i = 200$ bit, $i = 1, 2, 3, 4$. The simulation results are shown as follows. The initial energy of each sink node is set as $E_i = 200$ mJ, $i = 1, 2, 3, 4$, the distances between the sink nodes and the base station are the same as $d = 100$ m. The simulation results from Figs. 6, 7, 8, 9, 10, 11, and 12 are all obtained from the simulation results by the Lingo optimization software based on the optimization model Figs. 3, 4, and 5.

In Figs. 6, 7, and 8 we display the optimal routing based on the three different models (Max T, Mix ET, Min E) in the first round.

Fig. 6 The optimal routers based on lifetime maximize protocol

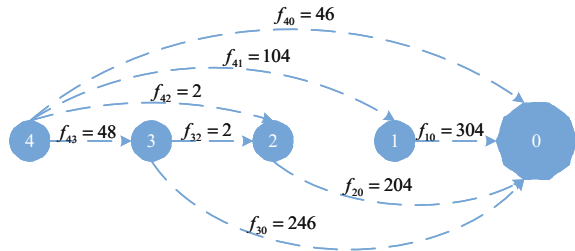


Fig. 7 The optimal routers based on latency minimize protocol

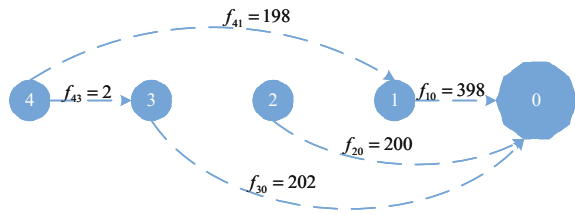
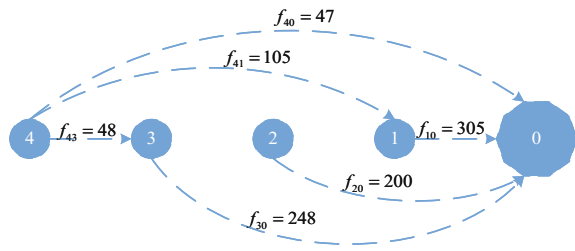


Fig. 8 The optimal routers based on utility maximize protocol



Using Figs. 9, 10, 11 and 12, it is easy to find that the communication rounds which represent the lifetime time of the system in Max T Model and Mix ET Model are longer than Min E Model. In Fig. 9, the minimum residual energy of all sink nodes with Mix ET Model is similar to Mat T while much more than Min E Model. As shown in Fig. 10, the total residual energy of all sink nodes is similar to each other with different models while the variance of the Min E Model is much bigger than the other two models. From Fig. 11, the variance of the residual energy in Mix ET is similar to Mat T but much smaller than Min E, it means that the energy consumption in the two methods is more balanced and it is benefited to improve the system lifetime. In Fig. 12, the number of hops in Mix ET Model is much less than Max T Model while slightly more than Min E Model in most of the communication rounds, which means that in the performance of latency, the Mix ET Model is close to the Min E Model which with the single-objective to minimize the latency.

Fig. 9 Comparison of the minimum residual energy of all sink nodes

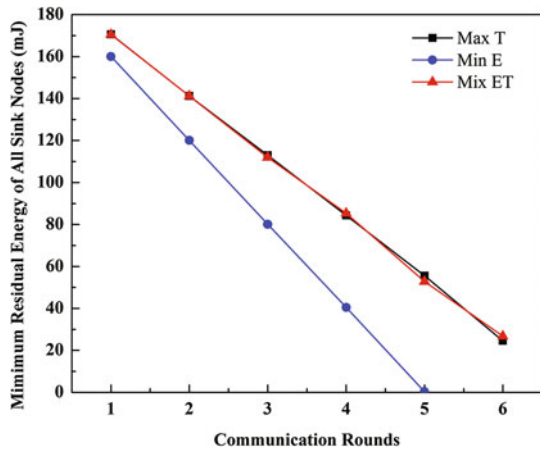


Fig. 10 Comparison of the total residual energy of all sink nodes

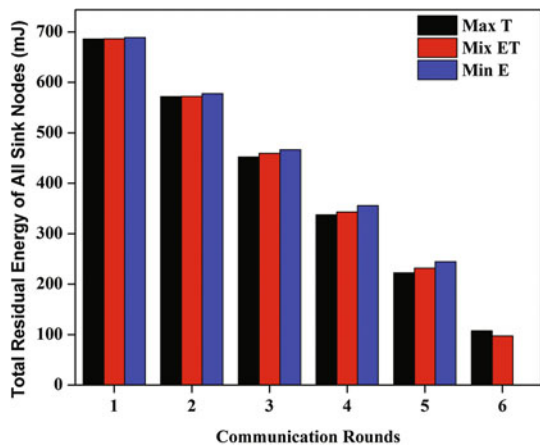


Fig. 11 Comparison of the residual energy variance of all sink nodes

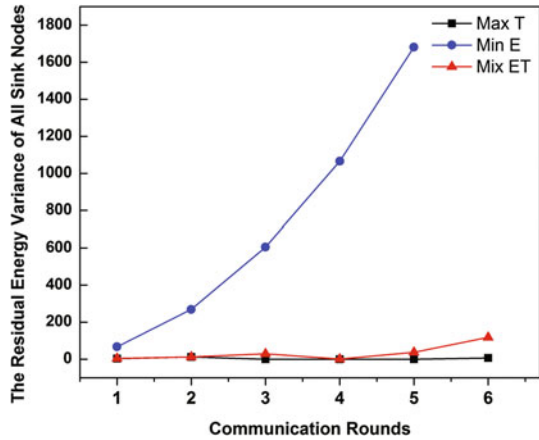
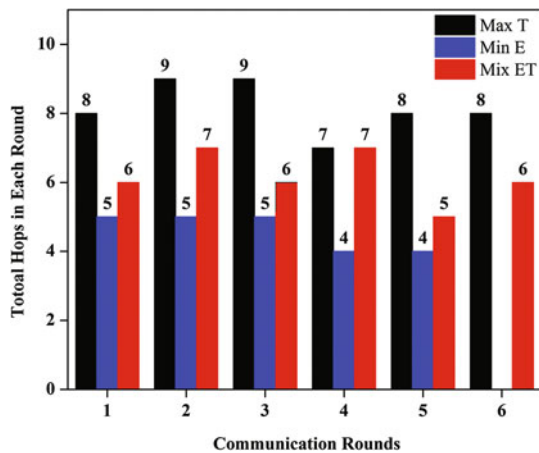


Fig. 12 Comparison of the number of hops for all sink nodes



In general, the utility of the Mix ET Model is superior to the other two models, whose lifetime is same to Max T Model but more than Min E Model and the latency is close to Min E Model while less than Max T Model.

5 Conclusions

The lifetime maximum, latency minimum, and utility maximum models based on the multi-hop communication protocol are proposed in this paper. The multi-hop communication protocol aims to divide the transmitted data packets into several small packets and allocated the transmission tasks among all the sink nodes to balance the energy consumption. The first two single-objective models aim to maximize the lifetime and minimize the latency of the system respectively.

The third model aims to improve the utility of the system, which takes the two factors into consideration at the same time. The simulation results show that the performance of the utility model is superior to the other two models and is more suitable to be used in the railway infrastructure wireless monitoring system.

Acknowledgement This work was supported by the National Science and Technology Major Project Grant No. 2016YFB1200100; State Key Laboratory under Grant No. RCS2016ZT018.

References

1. Shafiullah GM, Azad SA, Shawkat Ali ABM (2013) Energy-efficient wireless MAC protocols for railway monitoring applications. *IEEE Trans Intell Transp Syst* 14(2):649–659
2. Pantazis NA, Nikolidakis SA, Vergados DD (2013) Energy-efficient routing protocols in wireless sensor networks: a survey. *IEEE Commun Surv Tutor* 15(2):551–591
3. Conti M, Willemsen J, Crispo B (2013) Providing source location privacy in wireless sensor networks: a survey. *IEEE Commun Surv Tutor* 15(3):1238–1280
4. Yildiz HU et al (2016) Maximizing Wireless Sensor Network lifetime by communication/computation energy optimization of non-repudiation security service: node level versus network level strategies. *Ad Hoc Netw* 37:301–323
5. Incebacak D et al (2015) Optimal data compression for lifetime maximization in wireless sensor networks operating in stealth mode. *Ad Hoc Netw* 24:134–147
6. Ergen SC, Varaiya P (2005) On multi-hop routing for energy efficiency. *IEEE Commun Lett* 9(10):880–881
7. Chang JH, Tassiulas L (2004) Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Trans Networking* 12(4):609–619
8. Bicakci K, Tavli B (2010) Prolonging network lifetime with multi-domain cooperation strategies in wireless sensor networks. *Ad Hoc Netw* 8(6):582–596

Application of High Accelerated Life Testing for an Integrated Control Board

Yingying Yuan, Xuhui Zhang, Shaohua Du, Zaiwu Peng,
Xiaochun Xiao and Jiani Zuo

Abstract As a core part of the electric vehicle motor controller, the integrated control board plays an important role in reliability and safety of the controller running. In order to ensure the stability of the performance of the new controller, this paper carried out a high accelerated life test for the relatively weak integrated control board, so as to quickly stimulate the weakness and defects of the PCBA in design and production process. Based on the discovered faults, we can improve and optimize the design and process as early as possible. As a result, we can reduce the failure rate of the product and finally improve its inherent reliability level. In addition, this test may also provide parameters for the following reliability tests such as high accelerated stress screening test, so the plan design proposed in this paper is of great value and significance.

Keywords High accelerated life test · Test profile · HALT · Stress limit

1 Introduction

In recent years, based on the development trend of high power density of motor controller, we developed a highly integrated control board, which combines the functions of power control, motor control, and vehicle control. In order to meet the fast and changing market demand, this type of integrated control board is urgently needed to be assessed by reliability test. However, the conventional reliability test, such as routine test and environmental test which is of low stress level and long testing time, cannot quickly and completely expose the potential defects of product

Y. Yuan (✉) · X. Zhang · Z. Peng · X. Xiao · J. Zuo
Hunan CRRC Times Electric Vehicle Co., Ltd., Zhuzhou
Hunan 412007, China
e-mail: yuanyy1@csrzc.com

S. Du
CRRC ZIC Research Institute of Electrical Technology and
Material Engineering, Zhuzhou, Hunan 412001, China

efficiently. For the shortcomings of the traditional test, this article applied highly accelerated life test (HALT) [1–3] which was first proposed by Dr. Hobbs for the integrated control board. The biggest feature of HALT technology is time saving, that is, it can stimulate the potential defects of the product efficiently by using the increasing stress which is much larger than technical specifications [4–6], so it can stimulate the defects of PCBA in design and production process rapidly. After design optimization and test verification, we can ultimately improve the reliability of the product.

In this paper, we conduct HALT technology on the integrated control board. Through the optimization and improvement of the problems explored by HALT, we can improve the inherent reliability of the PCBA [7, 8]. What is more important is that the consumption of human and material costs of the problems founded in the design and development stage by HALT is far less than the investment cost of post-failure processing, thus saving huge operating costs.

2 Test Procedure of HALT

Based on the experience of practical engineering application, this paper draws out the HALT test procedure, which is divided into four steps and is demonstrated in Fig. 1.

First, according to verification needs in design and development stage, and based on the historical fault data, we should clear the objects of HALT, such as the machine, key parts, the product which is of the high frequency of occurrence.

Second, we should clear the test environment and carry out the preparatory work required for the HALT of PCBA, which includes the selection of test equipment, the number of test samples and the test items, the fixture production, the test standards, the failure criteria, the system connection of test, and the test operation precautions.

Third, design the test profile. At the beginning, we should find the stress limitation of temperature and vibration of the PCBA such as the working limit where the failure can be restored and the damage limit where the failure cannot be restored, which can provide the parameter input for the temperature cycle and comprehensive tests. After that, based on the stubborn failure exposed in temperature cycle and comprehensive test, we can provide effective fault information for the following design optimization and improvement of the product.

Fourth, according to the fault inspired by HALT, the designers could discuss and analyze the fault type, and conduct the fault diagnosis and fault location, and propose feasible measures, and then verify the feasibility and effectiveness of measures. Finally, we output the test report.

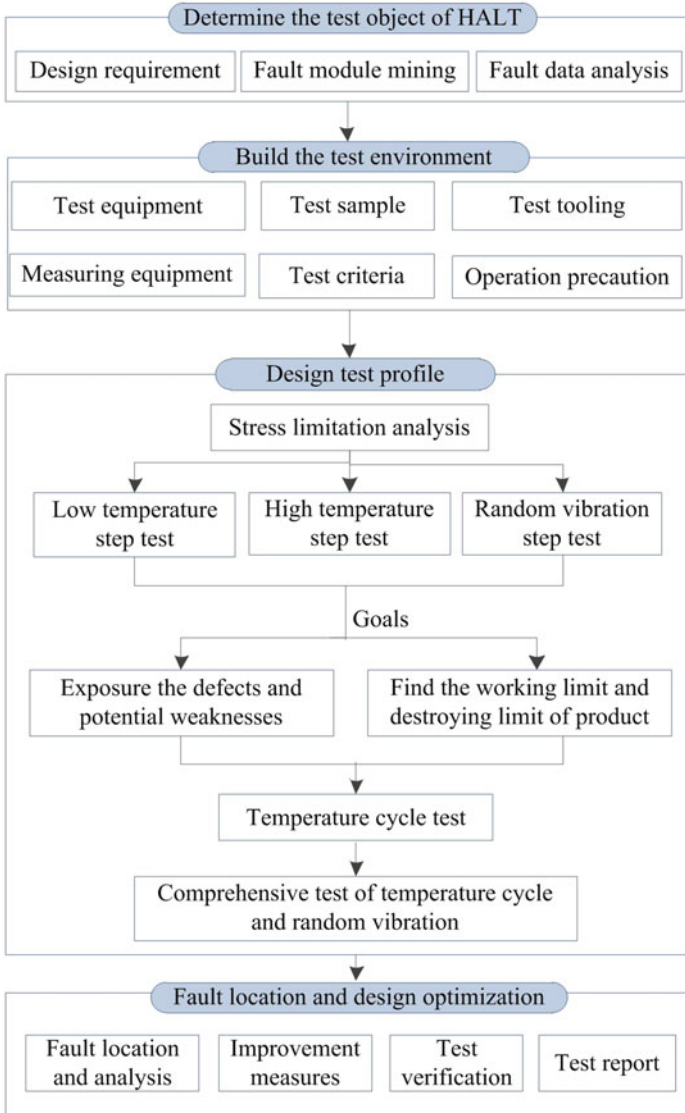


Fig. 1 Test procedure of HALT

3 Test Plan Design

3.1 Test Object

In this paper, the object of HALT is the integrated control board, and the number of samples is eight. The specific distribution of the samples in HALT is shown in Table 1.

3.2 Test Environment

The basic performance parameters of the HALT test equipment is shown in Table 2.

The design of test tool of the integrated control board should consider the following six factors, such as the test efficiency, the size and weight of the samples and the tools, the installation location of wires, and the vibration level of the HALT equipment. In this HALT, we adopt the aluminum design, where the plate can place two PCBAs at the same time, which saved half of the test time.

The system connection diagram of the integrated control panel is shown in Fig. 2. And the test equipment includes low-voltage power supply, precision power supply, signal generator, oscilloscope, etc., and the performance test of PCBA is referred to the debugging outline.

Table 1 The distribution of the integrated control board

No	Test item	Test samples
1	Low-temperature step test	1#, 2#
2	High-temperature step test	3#, 4#
3	Random vibration step test	5#, 6#
4	Temperature cycle with vibrating step test	Good samples selected from the above six samples (spare: 7#, 8#)

Table 2 The performance parameters of the HALT test equipment

No	Item	Performance parameters
1	Temperature range	-100 °C to +200 °C (liquid nitrogen cooling, and the temperature deviation is ±3 °C)
2	Temperature variability	Not less than 60 °C/min
3	Vibration type	Random vibration (the bandwidth of vibration energy is 2-10,000 Hz, and the output of vibration level is at least 80 g in case of no load, and the deviation of vibration is ±3 g)

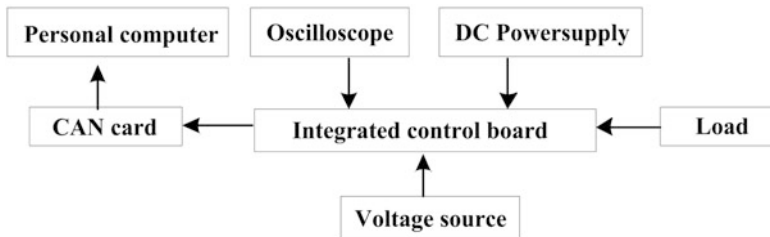


Fig. 2 The connection of test system

3.3 Test Profile Design

The HALT test profile of the integrated control board includes four stages, such as low-temperature step test, high-temperature step test, vibration step test, and integrated stress test.

The test samples of low temperature step are 1#, 2#. The test conditions are that the temperature variability is 60 °C/min, and the time of cold soaking and performance test is 30 and 5 min. Before the routine performance test, we should carry out five power off tests. The beginning step value is -10 °C. When the temperature is close to the working limit, the step value is -5 °C. When the temperature reaches the low-temperature working limit, the test is ended. The specific low-temperature test profile is shown in Fig. 3.

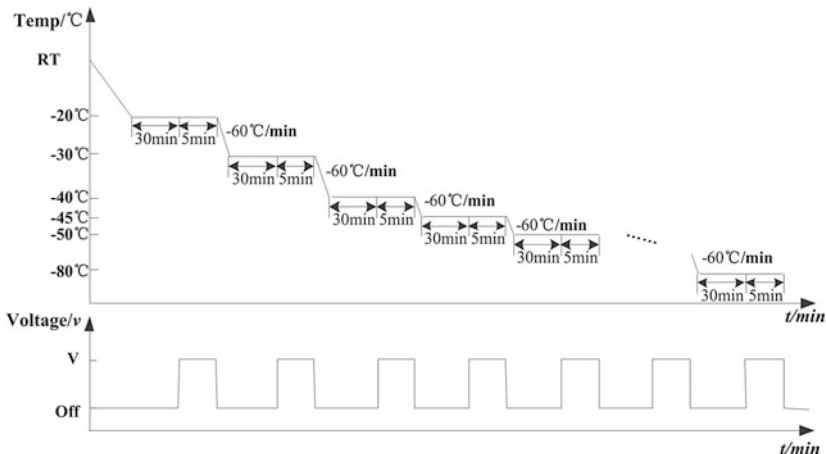


Fig. 3 The low-temperature step test profile of integrated control board

When the temperature is $-80\text{ }^{\circ}\text{C}$, the performance test of the integrated control board is still normal. According to the stop criterion, we end the test, and make $-80\text{ }^{\circ}\text{C}$ as the low temperature working limit of the PCBA.

The test samples of high temperature step are 3#, 4#. The test conditions are that the temperature variability is $+60\text{ }^{\circ}\text{C}/\text{min}$, and the time of heat soaking and performance test is 30 and 5 min. The whole test process is of electricity. The temperature rises from the room temperature to $65\text{ }^{\circ}\text{C}$ at first, and the step value is $+10\text{ }^{\circ}\text{C}$. When the temperature is close to the high-temperature limit, we reduce the step value to $+5\text{ }^{\circ}\text{C}$ until it reaches the high temperature working limit of the PCBA. The specific high-temperature test profile is shown in Fig. 4.

Through the high-temperature step test, we find the high-temperature working limit ($135\text{ }^{\circ}\text{C}$) and the high-temperature damaging limit ($137\text{ }^{\circ}\text{C}$) of the integrated control board.

The test samples of vibration step are 5#, 6#. The acceleration rate of vibration is $30\text{ g}/\text{min}$, and the random vibration begins from 5 g to the value that we find the working limit of the integrated control board. Each vibration step includes two stages, and the time of stabilization stage and performance test stage is 10 min and 5 min. Before each performance test stage, we should conduct 3 min of weak vibration which the value is 5 g , so as to find the failure of the PCBA efficiently. The specific vibration test profile is shown in Fig. 5.

Through the vibration step, we find the vibration working limit (40 g) of the integrated control board.

The samples of the integrated stress test are 1#, 2#, 7#, 8#. The test includes three kinds of stress—the fast temperature cycling, the random vibration, and the electric stress. The temperature cycle includes two parts: cold/hot penetration and performance test, and each of the time is one hour and five minutes. Based on the working limit of the low/high-temperature step test and the vibration step test, we can determine the parameters of the integrated stress test profile. Normally we set

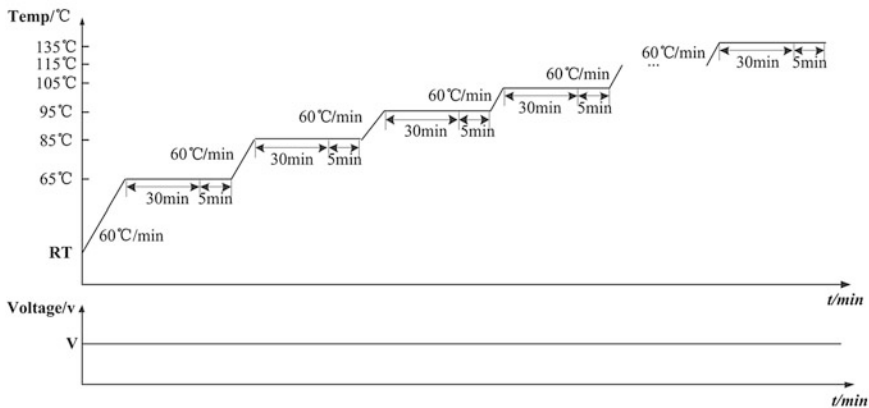


Fig. 4 The high-temperature step test profile of integrated control board

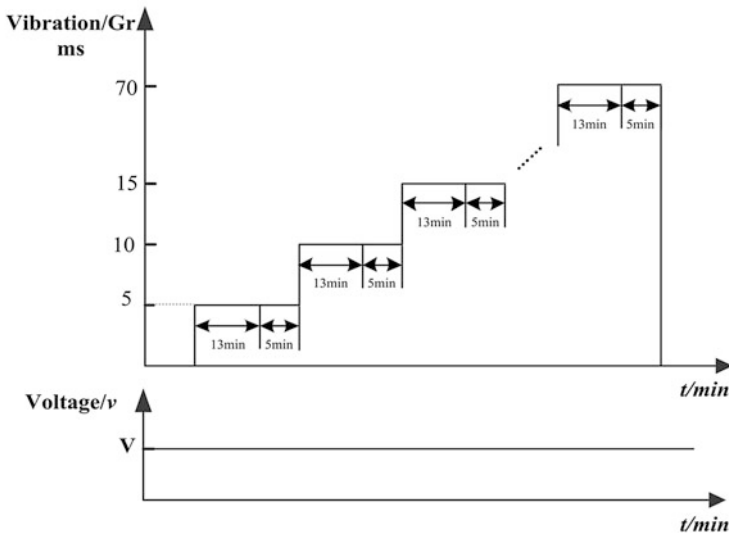


Fig. 5 The random vibration step test profile of integrated control board

up five cycles, however, if the PCBA is still of a good performance after five cycles, we could improve the test stress by increasing the number of cycles or increasing the stress level of the temperature or vibration for stimulating the potential fault of the product. The specific comprehensive test profile is shown in Fig. 6.

According to the low/high temperature and vibration step test, we find the upper and lower limits of the comprehensive stress test, which are 130 and -75°C , and set the temperature change rate to $60^{\circ}\text{C}/\text{min}$. The vibration step will be started at 20 g.

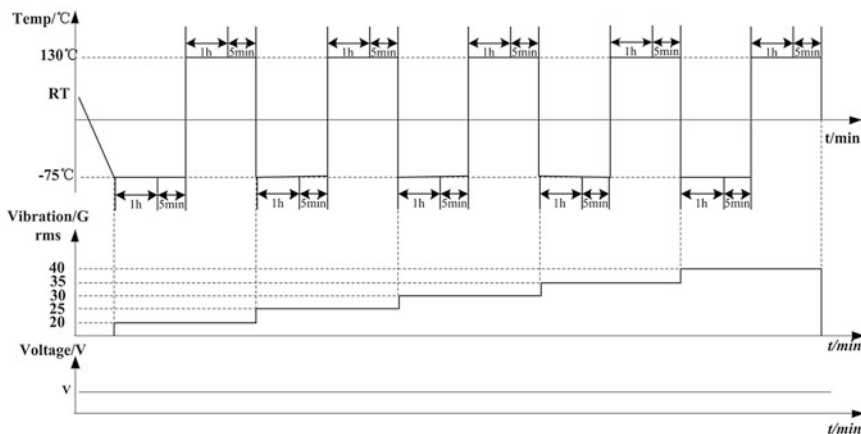


Fig. 6 The comprehensive stress test profile of integrated control board

4 Test Result

Through the test profile design and implementation of HALT, we finally stimulate 13 potential faults, and find the working limitation of low/high temperature and the random vibration which are $-80, 135\text{ }^{\circ}\text{C}$, and 40 g , and the damage limit of high temperature is $137\text{ }^{\circ}\text{C}$.

The weakness of the integrated control board which is sensitive to the stress of low/high temperature, random vibration, and integrated test we found is shown as follows. The power module D6 and diodes such as V23, V24, V31, V32, V33 are failure in $105\text{ }^{\circ}\text{C}$. Power modules such as D3 and D2 are failure at 135 and $137\text{ }^{\circ}\text{C}$. The power module like D6 is failure at the vibration level of 40 Grms . The power modules (D3, D6), and the capacitance (C12), and the inductance (L5) are failure at the stage of integrated stress.

5 Fault Location and Design Optimization

Through the HALT test of the integrated control board, we found five non-recoverable faults. The specific fault phenomena and the corresponding fault analysis and suggestions for improvement and optimization of the PCBA are shown in Table 3.

It can be proved that the failures stimulated by HALT can be eliminated after adopting the suggestions proposed in Table 3, which once again proves that the measures carried out are feasible and effective. According to the conclusions of this article and the information we obtained from other environmental experiment and white box test, the designers can revise and optimize the design of PCBA, which can avoid the potential risk of failure thoroughly.

Table 3 Fault location and analysis

No	Fault phenomenon	Cause analysis	Suggestion
1	D2 is damaged at the moment of power-up in high temperature	The voltage difference between the 26th and 28th pins of D2 is about 10 V , which is greater than the voltage threshold of 8.5 V	The output time of D2 is extended from 4.2 to 10 ms
2	Under the condition of high temperature and high vibration, the pins of C12, L5, D3, and D6 are loose and disconnected	These components are not placed very well	Strengthen the dispensing of the components and similar devices
3	D3 and D6 are damaged at the temperature of $135\text{ }^{\circ}\text{C}$	The value has exceeded the temperature specified by the technical specification	Looking for the alternative models

(continued)

Table 3 (continued)

No	Fault phenomenon	Cause analysis	Suggestion
4	V23, V24, V31, V32, and V33 are affected by high temperature, resulting in abnormal high/low-voltage sampling	Diodes are affected by high temperature, resulting in increased reverse leakage current and reverse drift	Remove the relevant diodes or redesign the location of the diodes to eliminate the impact of its temperature characteristics
5	The signal transmission of plugs (X7, X8) are abnormal, such as alarming module failure	Under the condition of high temperature and high vibration, the plugs of X7, X8 are contacted abnormally	Strengthen the material control and wiring harness production process

6 Conclusion

In this paper, we established a complete set of process system of HALT technology, which includes the test samples selection, the bench building, the test profile designing, and the fault location and redesign optimization. It could provide a reference for other reliability experiments in future. We conducted the HALT by taking integrated control board as an example, and from the test, we stimulated 13 failures. Based on the fault location and analysis of failures, we proposed improvement measures, and once again through the HALT test to verify its effectiveness, and finally closed the cycle of fault problems. What needs to be pointed out is that this HALT test successfully reproduced the failure of power chip of D2, which behaves abnormally in the actual electric drive system, and this demonstrates the high efficiency of HALT test profile design. Additionally, through the low/ high-temperature limitation found in this article, we can provide basis parameters input for HASS and dynamic aging test.

As an effective mean of positive reliability design, the effective use of HALT test at the design and development stage can quickly stimulate the defects of PCBA. Only when we have the ability to stimulate fault, correct fault, and verify the effectiveness of the methods proposed for HALT technology, we can reduce the failure rate of PCBA and achieve the growth of reliability genuinely.

References

1. Jiang T (2000) Reliability enhancement test. *Environ Technol* 1:35–42 (in Chinese)
2. Jiang T, Jin M (2000) Reliability enhancement test and the application in aviation engineering. *Aero Eng Maintenance* 1:258–266 (in Chinese)
3. Luo M, Jiang T (2010) Reliability Enhancement Test Technology for Accelerated Range of Product Safety. *J Ordnance* 1:24–29 (in Chinese)
4. Yao J, Jiang I (2006) Calculation and control of profile parameters based on electric vibration table. *Environ Technol* 3:30–35

5. Wang X, Zhang Z, Zhong Y (2016) Research on reliability enhancement test technology and application of electronic products. *Electron Prod Reliab Environ Test* 5:26–31 (in Chinese)
6. Chen T, Wu C, Wang H (2015) Application of reliability enhancement test technology in the development of air-to-air missile. *Aviat Weapons* 4:11–17 (in Chinese)
7. Li C, He Y, Xu W, Guo L (2015) Thoughts on development of reliability test technology for space launch vehicle. *Qual Reliab* 2:159–163 (in Chinese)
8. Liu W, Liu H, Liu H (2013) Application of reliability enhancement testing in the development of aero-engine accessories. *Aviat Maintenance Eng* 3:212–217 (in Chinese)

Remote Monitoring System of Electrical Equipment Based on GPRS

Yu Wang and Xuedong Jiang

Abstract Based on GPRS, the remote monitoring system of electrical equipment realizes the data sharing between the electrical equipment of each train and the remote terminals. This system consists of three parts: vehicle acquisition module, mobile network equipment and remote terminal. The CAN controller of the vehicle acquisition module transmits the collected power parameter data to the protocol converter, and the main controller periodically transmits the data to the remote terminal through the GPRS module, then the remote terminal stores and displays the data. The system can greatly improve the efficiency of the accident handling and improve the train running safety factor.

Keywords Remote monitoring system · GPRS · Data sharing

1 Introduction

With the vigorous development of railway construction, the high-speed railway has become the inevitable trend of the modern railway. Then the question of how to improve the safety factor of the railway becomes particularly important. Remote monitoring system realizes data sharing, so the safety problem has been solved [1–3]. As the whole power source of the train, the electrical equipment has a direct impact on the safe operation of the train. Through GPRS [4–6], the power parameters of electrical equipment will be shared. Based on these information, the dispatch centre can make the right decision to let the past vehicle change the line or keep running, then sends trailer and maintenance personnel to deal with the problem. This can effectively shorten the accident processing time and the normal operation of the railway line is guaranteed—to the maximum extent [7, 8]. In addition, compared to the previous system, this system uses advanced technology module, so its performance, stability and power consumption have been greatly improved.

Y. Wang · X. Jiang (✉)

School of Electrical Engineering, Beijing Jiaotong University, Beijing, China
e-mail: 15121482@bjtu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_96

951

Fig. 1 Structure of remote monitoring system



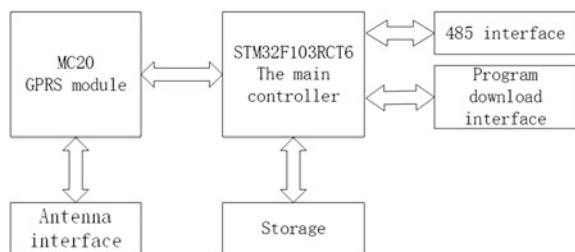
The remote monitoring system is simple in structure, the vehicle acquisition module is installed on the train, the remote terminal is located in the traffic control centre, and the mobile network equipment is taken as the middle link. The concrete structure is shown in Fig. 1.

2 Vehicle Acquisition Module

The vehicle acquisition module mainly includes the main controller, CAN controller, communication protocol conversion module, and storage. The CAN controller collects and transfers the data to the protocol conversion module, and convert the data that conforms to the CAN protocol to the data in the LKJ-93 format that conforms to the 485 communication protocol. The main controller transmits the real-time data to the GPRS module at regular intervals. Hardware of the vehicle acquisition module is shown in Fig. 2.

Through the remote monitoring equipment based on GPRS, the power parameters on the CAN bus can be collected in real time. The application of the equipment can effectively shorten the processing time of the accident.

Fig. 2 Hardware of the vehicle acquisition module



2.1 The Main Controller

The main controller receives the data sent by the protocol converter, caches it and sends the data to the GPRS module at regular intervals for long-distance transmission.

The main controller uses 32-bit RISC processor STM32F103RCT6. The STM32F103 [9] family uses a high-performance, low-cost, low-power ARM Cortex-M3 kernel. The maximum operating frequency of the main controller is 72 MHz, which has 256 K bytes flash, 64 K bytes of SRAM, 4–16 MHz crystal oscillator, 1.25 DMIPS/MHz, and three low-power modes: sleep, downtime, and standby. The module has 64 pins, 51 general-purpose I/O port, communications ports that supports multiple protocols and 3 general 16-bit timer. The working temperature range is $-40-85^{\circ}\text{C}$.

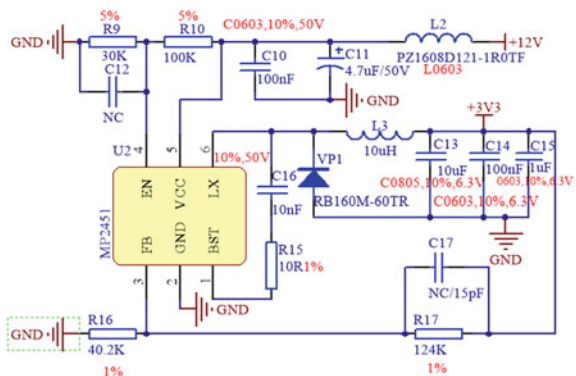
The main controller’s supply voltage is 2.0–3.6 V. In this paper, through the 2541 chip and filter circuit, the battery with 12 V voltage is converted to a more pure 3.3 V. The power supply part of the schematic diagram is as Fig. 3.

2.2 The CAN Controller

CAN controller is a serial data communication bus developed to solve the data exchange between various measuring and controlling components. In this paper, SJA1000 is used as an alternative to the PHILIPS semiconductor PCA82C200 CAN controller with a bit rate of up to 1 Mbps.

The CAN controller controls the transmission and reception of data frames according to the CAN bus protocol. The controller can store a full standard or extended message and pass the required data to the main controller via the protocol converter.

Fig. 3 Schematic of power supply circuit



2.3 The Communication Protocol Conversion Module

Turn the CAN bus protocol into the RS485 protocol. This communication method transmits data through differential signals. The interface signal level is reduced by 232, so the service life of the interface chip can be extended. The interface uses a combination of balanced drives and differential receivers, which enhances the anti-common mode interference capability. The converter supports up to 32 nodes at the same time.

This article uses the products of Beijing An Rong Hu Tong Technology Co., Ltd. The product is equipped with high-speed photoelectric isolation opt coupler and DC/DC isolated power module, which ensures that the converter has a strong anti-interference ability, while improving the reliability of the system in harsh environments.

3 Mobile Network Equipment

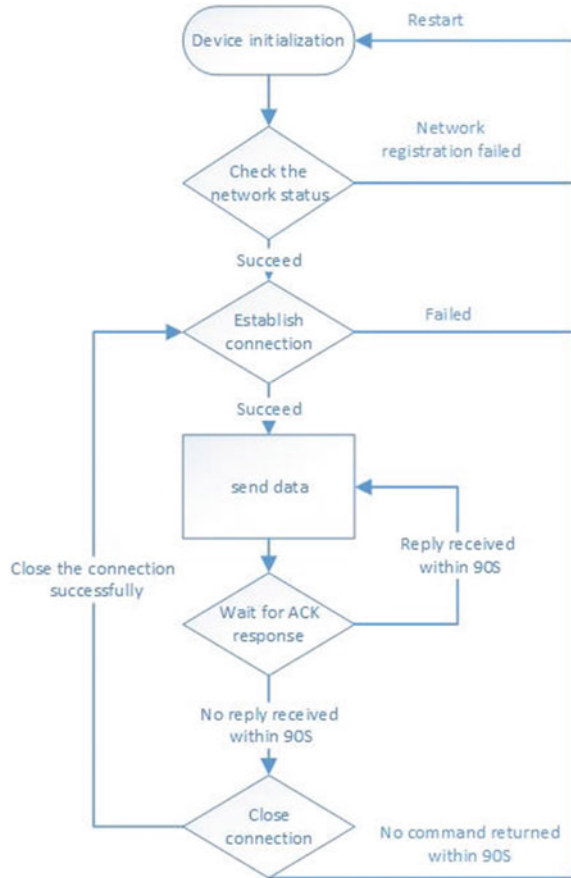
We use the Quectel MC20 module as mobile network equipment, through the asynchronous serial port module connected to the main controller. The module uses MediaTek's latest multi-function communication positioning chip, which has ultra-small size, low-power consumption, embedded rich network protocols (such as TCP, UDP, PPP, FTP, HTTP and SSL), integrated multi-constellation satellite system (such as the Big Dipper, GPS, QZSS), which can provide wireless mobile communications and accurate positioning function. Its receive sensitivity can be up to -149 dBm. The power consumption in low-power mode (GLP Mode) is only 40% of the normal operating mode.

Data exchange between module and server uses AT command, which is easy to understand. The data transmission speed can be controlled by adjusting the baud rate. The data transfer process is as shown in Fig. 4.

4 The Remote Terminal

The remote terminal is located in the dispatch centre, which mainly includes access equipment, database and host compute. GPRS terminal (access device) and PC (data centre) have established a TCP connection to the directory server, and the link maintenance of the host computer and the directory server is done automatically by the TCP/IP protocol. Link maintenance operations still need to be performed between the GPRS terminal and the directory server. We use the Microsoft SQL

Fig. 4 The data transfer process



Server database as the database management software, which manages all the locomotive power parameter information, data analysis, report generation, access permissions and user terminal management.

The remote terminal receives data sent by the mobile network device and stores, processes and displays the data in the database. When the number of abnormal data occurs, the dispatcher can do the corresponding processing according to the data. The topology interface of the remote terminal is shown in the following Fig. 5.

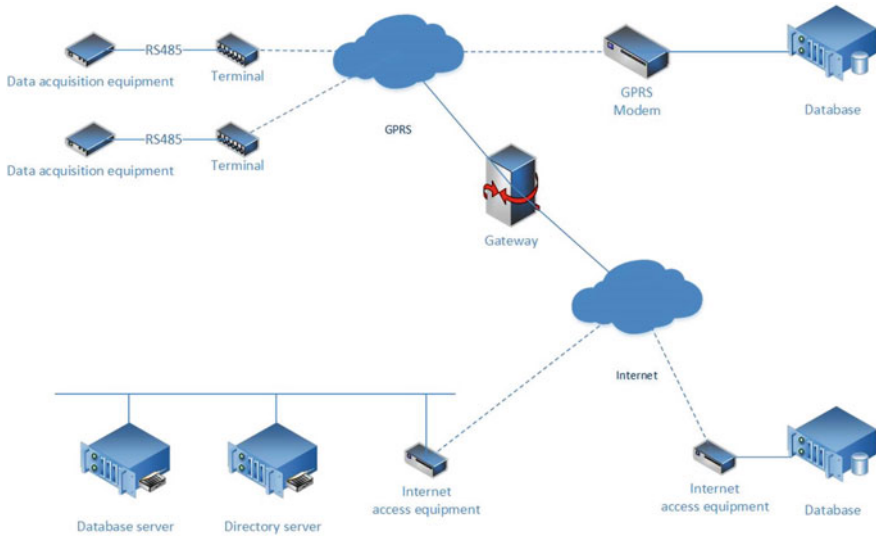


Fig. 5 Topology of the remote terminal

5 Conclusions

Through the remote monitoring equipment based on GPRS, the power parameters of the power equipment can be shown in real time for us to make the right choice. The application of the equipment can effectively shorten the processing time of the accident.

References

1. Fan Z (2010) Design of the railway vehicle remote monitoring system. Jilin University (in Chinese)
2. Shaokai L (2015) Research on remote diagnosis system for rail transit vehicles. Beijing Jiaotong University (in Chinese)
3. Yuhong C (2007) The research and development of the host computer of the train running monitor and record device. Zhejiang University (in Chinese)
4. Pinglong M, Puxuan D (2007) Remote real time monitoring system of locomotive signal based on GPRS. *Railway Signal Commun* 01:53–55 (in Chinese)
5. Shiyang D (2011) Remote monitoring system of pressure vessel based on GPRS intelligent information technology application association. In: Proceedings of the 2011 international conference on mechanical engineering and technology (ICMET 2011) intelligent information technology application association, 4
6. Yingjie L, Yuguang F (2015) Research on mineral yield in remote monitoring system based on GPRS Northeastern University IEEE Singapore section. In: The 27th Chinese control and decision conference, vols 2, 4, pp 5760–5763

7. Zhichao Y (2004) Design and application of the GPRS wireless data transmission terminal. Xi'an University of Technology (in Chinese)
8. Jian P, Feng Y, Jingze X (2011) Function design and application research of automobile remote monitoring and diagnosis system Shanghai Auto 3:24–29 (in Chinese)
9. Shuying S, Zhijia C, Chao K (2010) The developing and application of the new generation of embedded MPU-STM32F103 12:59–63 (in Chinese)

YQuery: A Novel Privacy- and Integrity-Preserving Range Queries in Two-Tiered Sensor Networks

Lvyou Yu, Jingwen Yu, Lei Wang and Linai Kuang

Abstract A novel protocol named YQuery was proposed in this paper to prevent attackers from gaining the information processed by or stored in storage nodes, which can not only achieve the goals of data privacy and integrity preserving, but also ensure the secure range query without incurring false positive, and allow sink node to detect the compromised storage nodes when they misbehave. To preserve privacy, YQuery uses an order encryption mechanism by adopting stream cipher to encrypt/decrypt both sensed data and queries. To preserve integrity, YQuery uses a new data structure called S2L, which allows sink node to verify the integrity of retrieved data via the so-called verification information. Analysis and experimental results demonstrate the effectiveness of YQuery.

Keywords Two-tiered wireless sensor network · Secure range query
Data privacy · Integrity preserving

1 Introduction

Wireless sensor networks (WSN) are composed of spatially distributed autonomous sensors to monitor physical or environmental conditions and to cooperatively pass their data through the network to the main location [1].

A range query is a fundamental operation in WSNs. Recently, some efficient range query protocols with privacy and integrity protection have been proposed for two-tiered WSNs already. For example, Sheng et al. [2] proposed a verifiable privacy-preserving range query scheme in a two-tiered WSN, in which, the bucketing

L. Yu · J. Yu · L. Wang (✉) · L. Kuang
College of Information Engineering, Xiangtan University, Yuhu District,
Xiangtan 411105, Hunan, People's Republic of China
e-mail: 364715358@qq.com

L. Yu · J. Yu · L. Wang · L. Kuang
Key Laboratory of Intelligent Computing and Information Processing,
Xiangtan University, Yuhu District, Xiangtan 411105, Hunan, People's Republic of China

technique was utilized to mix the data in a certain range, and a verifiable query protocol was designed by employing encoding numbers to enable sink to validate the query reply. Later, Chen et al. [3] proposed an efficient secure range query protocol named SafeQ based on the prefix membership verification [4], in which, neighborhood chain was constructed to guarantee the completeness of query result, and the Bloom-Filter [5] was adopted to reduce the cost of communication. Tsou et al. [6] proposed a privacy- and integrity-preserving rang query method named EQ, in which, an order encryption mechanism based on stream cipher was used to achieve privacy protection, and a new data structure named XOR-linked list was constructed to achieve the integrity of query result.

In this paper, a secure range query protocol with privacy and integrity protection was proposed for two-tiered WSNs, in which, for data privacy protection, an order encryption mechanism was used by adopting stream cipher to encrypt/decrypt both sensors sensed data and sink-issued queries such that storage nodes can only process sink-issued queries over stored data in the encryption domain, and for data integrity preserving, a new data structure called S2L was constructed to generate the integrity verification information so that inquirers can verify the integrity of query result. Comparing with existing methods, simulation results show that our newly proposed method has better performance of privacy and integrity protection, smaller communication overheads, and required less energy and storage space for range queries.

The rest of this paper is organized as follows. Section 1 summarizes related work. Network model and attacker model are described in Sect. 2. Section 3 introduces the proposed scheme. System analysis is elaborated in Sect. 4. Section 5 presents the performance evaluation. And Sect. 6 concludes this paper with some future works.

2 Network and Attack Models of Two-tiered WSNs

Network Model: We consider two-tiered sensor networks as shown in Fig. 1. A two-tiered WSN can be partitioned into multiple cells, and each cell contains a storage node and some sensor nodes. Hence, in the two-tiered WSN, there are three kinds of nodes such as sink, storage node, and sensor. Among them, sensors are inexpensive sensing devices with limited storage, computation, and power, and every sensor will collect the environmental data in a fixed rate and periodically submit them to the storage node which is in charge of the cell that the sensor lies in [7]. The storage nodes have the strong computing power and a large storage capacity relative to the sensors. Each sensor periodically transmits the data collected by itself to the nearest storage node. When the sink receives a query request from a user, it will rewrite the query request and sends it to the storage node, and then, the storage node will return the data that satisfies the condition to sink according to the query request.



Fig. 1 Architecture of two-tiered sensor networks

In the two-tiered WSN, we state the problem of range queries in the following way: a range query that aims to find all the data items collected by the sensor node SN_i at the current time slot t in the range $[low, high]$ is denoted as $Qt = (SN_i, t, [low, high])$.

Attack Model: For two-tiered sensor networks, we assume that sensors and sink are trusted, but the storage nodes are not credible. Obviously, if a storage node is compromised, then the data collected by the sensors in the same cell of the storage node and the query sent from sink might be known by the attacker. In addition, the compromised storage node may be manipulated to return the wrong or incomplete reply in response to the sink-issued query. Therefore, in this paper, we mainly focus on the security of storage nodes.

3 The Proposed Scheme

Data Encryption: Once a storage node is attacked, the goals of data privacy and integrity preserving cannot be achieved. In order to address the above problems, We use the order encryption mechanism (OEM) proposed by Tsou et al. [6] to encrypt/decrypt both sensors sensed data and sink-issued queries. And the OEM function will be set up for the sink and sensor nodes at system initialization. To perform OEM, the domain of values of sensed data will be mapped into γ equivalent and consecutive nonoverlapping intervals such as $P_1, P_2, \dots, P_\gamma$ according to the network owner first, where these γ intervals are taken together shall cover the whole sensed data domain (d_{min}, d_{max}) , and each interval P_ρ is associated with a serial number ρ and a key $k_\rho (1 \leq \rho \leq \gamma)$ that is decided by the network owner and would be used as parameters of the OEM function.

In this following, we will describe the behaviors of sensor nodes, the storage node, and sink, in detail, respectively. Without loss of generality, we will focus on one cluster consisting of N sensor nodes such as $\{SN_1, SN_2, \dots, SN_N\}$ and a storage node S_0 .

Behavior of the Sensor Node: First, each sensor node will sort all sensed data with each of n bits collected by itself in the current time slot. For any given sensor node SN_i , let $\{d_1, d_2, \dots, d_m\}$ be all of the m sensed data values (plaintexts) collected by it in the current time slot, where there is $d_{min} < d_1 \leq d_2 \leq \dots \leq d_m < d_{max}$. Let the interval size used in OEM be μ , then SN_i will map each sensed data $d_i (1 \leq i \leq m)$ to a corresponding interval $p_{\lceil \frac{d_i}{\mu} \rceil}$, where $\lceil \frac{d_i}{\mu} \rceil$ means to round $\frac{d_i}{\mu}$ up to

an integer. And furthermore, for the integrity protection of query results, we proposed a new data structure S2L. Each sensed data $d_i(1 \leq i \leq m)$ will be concatenated with a $\lceil \log_2(m+1) \rceil$ -bits Tag_i and expressed as $\text{Tag}_i|d_i$, where $\text{Tag}_i = i \oplus d_i$, “|” represents a concatenation operation and \oplus represents XOR operation.

Next, we will expand the set $\{(1|d_1), (2|d_2), \dots, (m|d_m)\}$ obtained above to a new set $\{(0|d_{\min}), (1|d_1), (2|d_2), \dots, (m|d_m), (m+1|d_{\max})\}$. And then, $(i|d_i)$ instead of the single $d_i(1 \leq i \leq m)$ will be encrypted for confidentiality. Let E_k represent the OEM function and satisfy $E_k(x_1|x_2|\dots|x_m) = E_{k_{\lceil \frac{m}{\mu} \rceil}}(x_1) \dots | E_{k_{\lceil \frac{m}{\mu} \rceil}}(x_m)$. Then SN_i will upload the packet $\langle i, t, E_k\{(0|d_{\min})|(1|d_1)|\dots|(m|d_m)|(m+1|d_{\max})\} \rangle$ to its closest storage node during the current time slot t .

In order to illustrate the behavior of sensor nodes more intuitively, we give an example as follows: Let $D = \{d_i\}_{i=1}^{10} = \{2, 270, \dots, 800, 1000\}$ be the sensed data collected and sorted by the sensor node SN_1 at the time slot 3, and then, as shown in Fig. 2, first, SN_1 will map each sensed data in D to a corresponding interval (Step 1 in Fig. 2). Next, SN_1 will concatenate each sensed data with a Tag and obtain a set (Step 2 in Fig. 2). And later, SN_1 will expand the set to a new set such as $\{(0|d_{\min}), (1|2), \dots, (10|1000), (11|d_{\max})\}$ (Step 3 in Fig. 2). Finally, SN_1 will upload the packet $\langle 1, 3, E_k\{(0|d_{\min})|(1|2)|\dots|(10|1000)|(11|d_{\max})\} \rangle$ to its closest storage node (Step 4 in Fig. 2).

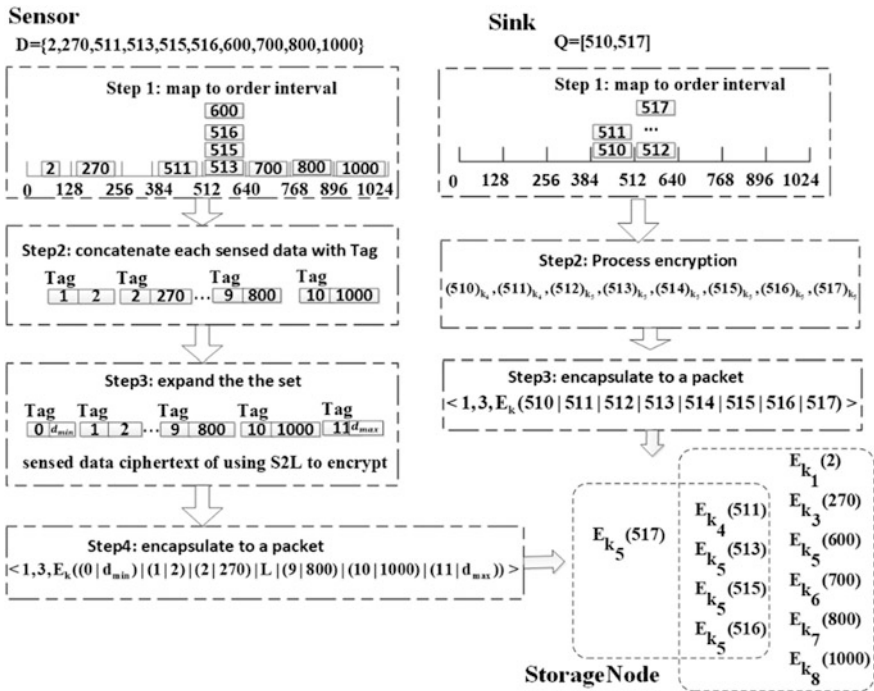


Fig. 2 The flowchart for data retrieval

Behavior of the Storage Node: When a packet arrives, it is obvious that the storage node can easily retrieve each sensed data value by slicing the ciphertext in the packet according to the structure of S2L. When the sink wants to query the sensor node SN_i for the data within the range [low,high] and sensed during the time slot t , it will construct a query packet $\langle i, t, E_k(d_{t1}|d_{t2}|\dots|d_{tk}) \rangle$ first, where there is $d_{t1} \leq \dots \leq d_{tm}$ and low and high are the lower and upper bound of $\{d_{t1}, d_{t2}, \dots, d_{tk}\}$, respectively, and then it will send the query packet to the storage node. After receiving the query packet, the storage node will search for the requested data stored in its database. Let $DR = \langle E_k(d_w|\dots|d_p) \rangle (1 \leq w \leq p \leq m)$ be the set of all encrypted sensed data stored in the database of the storage node and satisfying the sink-issued query request, where $d_w = \min(d_w|\dots|d_p)$ and $d_w = \max(d_w|\dots|d_p)$ are the min and max of the set DR . Specifically, in storage node's database, there are $d_{\min} \leq d_{w-1} < d_w \leq \dots < d_p < d_{p+1} \leq d_{\max}$. d_{w-1} is the left neighborhood of d_w and d_{p+1} is the right neighborhood of d_p in the database of the storage node. d_{w-1} and d_{p+1} satisfy the following: $(d_{w-1}, d_w) \cap \{d_1, d_2, \dots, d_m\} = \emptyset$, $(d_p, d_{p+1}) \cap \{d_1, d_2, \dots, d_m\} = \emptyset$.

If there is $DR \neq \emptyset$, the storage node shall include $(\text{Tag}_{d_{w-1}}|d_{w-1})$ and $(\text{Tag}_{d_{p+1}}|d_{p+1})$ in the query result(QR) because $(\text{Tag}_{d_{w-1}}|d_{w-1})$ and $(\text{Tag}_{d_{p+1}}|d_{p+1})$ ensure that the query result does include all encrypted sensed data that satisfy the query as the query result is bounded by them, which is expressed as $\langle E_k\{(\text{Tag}_{d_{n_1-1}}|d_{n_1-1})|(\text{Tag}_{d_{n_1}}|d_{n_1})|\dots|(\text{Tag}_{d_{n_w}}|d_{n_w})|(\text{Tag}_{d_{n_w+1}}|d_{n_w+1})\} \rangle$. Then, storage node sends the QR to sink as result; If $DR = \emptyset$, storage nodes do not send any message to sink as result. For example, when the sink wants to query the storage node for the data in the range [510, 517] sensed by SN_1 in time slot $t = 3$, it builds the query packet $\langle 1, 3, E_k(510|511|512|\dots|516|517) \rangle$ and sends to the storage node. When storage node receives the query packet, the storage node searches for the requested data stored in its database. In addition, the encrypted sensed data stored in the storage node's database intersected to the query range is the QR for the range query, such as $E_{k_4}(511), E_{k_5}(513), E_{k_5}(515)$ and $E_{k_5}(516)$, as shown in the storage node of Fig. 2. Then, while the network is reliable, storage node will send $QR = \{E_{k_4}(2|270)|E_{k_4}(3|511)|\dots|E_{k_5}(6|516)|E_{k_5}(7|600)\}$ to sink as result.

Behavior of the Sink: When the sink wants to query the storage node for the data in the range [low, high] sensed by SN_i during the current time slot t , first, it will map the query range into the corresponding interval. Then using the OEM to encrypt the query range, such as $E_k(d_s|\dots|d_e)$, where $\text{low} \leq d_s \leq \dots \leq d_e \leq \text{high}$. Next, encapsulating a query packet expressed as $\langle i, t, E_k(d_s|\dots|d_e) \rangle$ and then sending to the storage node SN_i . For example, when the sink wants to query the storage node for the data in the range [510, 517] sensed by SN_1 in time slot 3, first, it maps the query range into the corresponding interval (Step 1 in the storage node of Fig. 2). Then encrypting the query range, such as $E_k(510|511|\dots|516|517)$ (Step 2 in the storage node of Fig. 2). Finally, building a query packet and sending to the storage node, such as $\langle 1, 3, E_k(510|\dots|517) \rangle$ (Step 3 in the storage node of Fig. 2).

After a sink receives the QR obtained via Eq. (1).

$$E_k \left\{ (\text{Tag}_{n_1-1} | d_{n_1-1}) | (\text{Tag}_{n_1} | d_{n_1}) | \dots | (\text{Tag}_{n_w} | d_{n_w}) | (\text{Tag}_{n_w+1} | d_{n_w+1}) \right\} \quad (1)$$

where $d_{n_1-1} \leq d_{n_1} \leq \dots \leq d_{n_w} \leq d_{n_w+1}$. Then let the actual query results be DR. For each encrypted item $E_{k \left[\frac{d_j}{P} \right]} (\text{Tag}_j | d_j)$, it can be decrypted to get $(\text{Tag}_j | d_j)$ using the key $k \left[\frac{d_j}{P} \right]$ decryption, where $d_{n_1-1} \leq d_j \leq d_{n_w+1}$. The following is a description of the rules for integrity validation.

1. If there exists $n_1 < j < n_w$, but $(\text{Tag}_j | d_j) \notin \text{QR}$, the sink can know that there is missing sensed data by calculating $\text{Tag}_{j+1} - \text{Tag}_{j-1} > 1$. We assume the correct QR is $\{(2|270)|(3|511)|(4|513)|(5|515)|(6|516)|(7|600)\}$. If (4|513) is missing, then sink will receive a wrong QR. Sink can know that there is missing sensed data between (3|511) and (5|515) by calculating $\text{Tag}_5 - \text{Tag}_3 = 2 > 1$.
2. If $(\text{Tag}_{n_1} | d_{n_1}) \notin \text{QR}$, then sink can also detect that there is missing sensed data by calculating $\text{Tag}_{n_1+1} - \text{Tag}_{n_1-1} > 1$. For example, if (3|511) is missing, the sink can detect that there is missing sensed data by calculating $\text{Tag}_4 - \text{Tag}_2 > 1$.
3. If $(\text{Tag}_{n_w} | d_{n_w}) \notin \text{QR}$, sink can detect that there is missing sensed data $\text{Tag}_{n_w+1} - \text{Tag}_{n_w-1} > 1$, for example, if (6|516) is missing, then sink can detect that there is missing sensed data by calculating $\text{Tag}_7 - \text{Tag}_5 > 1$.

4 System Analysis

Privacy Analysis: In the phase of receiving and processing the packet, the sensed data in storage node is encrypted. Therefore, it is impossible to calculate the actual value of the sensed data without knowing the corresponding key; What is more, in the phase of processing sink-issued queries and sending query results, storage node only receives the encrypted range query. Without knowing the key used in OEM, it is computationally infeasible to compute of sink-issued queries. Through paper [6], we can get that our method satisfies the property of a negligible function which is used in security analysis to characterize the probability that a computationally bounded adversary successfully breaks a computationally secure encryption scheme.

Integrity Analysis: The goal of the integrity verification is to detect whether there is missing data. The data structure S2L satisfies this purpose. The integrity of QR can be verified firmly as described in the Sink Behavior of Sect. 3. And its computational complexity is $O(m)$.

5 Performance Evaluation

In this section, the performance of our method is evaluated and compared with the schemes of SafeQ-Bloom [3], and EQ [6]. The performance comparison is analyzed in these aspects: (1) Energy consumption of transmitting encrypted and transmitting encrypted query result. (2) Storage node storage space overhead.

We use MATLAB to implement SafeQ-Bloom, EQ, and our methods. In the following, we will introduce the experimental model and the experimental results.

Experimental model: We randomly distributed 100 sensors in a field of 100-meter long and 100-meter wide, and then placed a storage node in the center of the field. For sending L bit data over a distance d , the total energy consumed by a node (E_t) is as follows [8]:

$$E_t = \begin{cases} LE_{\text{elect}} + L \in_{fs} d^2 & d < d_0 \\ LE_{\text{elect}} + L \in_{mp} d^4 & d \geq d_0, \end{cases} \tag{2}$$

where d_0 denotes the shortest crossover distance, E_{elect} represents the energy required to activate electronic circuits, E_{elect} is the energy required to activate electronic circuits and \in_{fs}, \in_{mp} denotes the parameters related to receiver’s sensitivity and noise shape. Each sensor periodically collects 10,128-bit sensed data and sends encrypted sensed data to its nearby storage node. In particular, each encrypted sensed data in our method consists of a 5-bit tag and 128-bit sensed data value.

Evaluation Results: The experimental results show that the YQuery protocol is superior to SafeQ-Bloom and EQ scheme in terms of energy consumption and storage space cost.

Energy Consumption of Transmitting Encrypted Data: As shown in Fig. 3a, we can observe that our method consumed less energy than other methods. When compared with YQuery, EQ consumes about 1.5 times more energy and SafeQ-Bloom consumes about 5.7 times more energy.

Energy Consumption of Transmitting Query Results: From the results shown in Fig. 3b, we can observe that our method consumes 37% less energy than EQ. Compared with our method, SafeQ-Bloom consumes 8.1 times more energy.

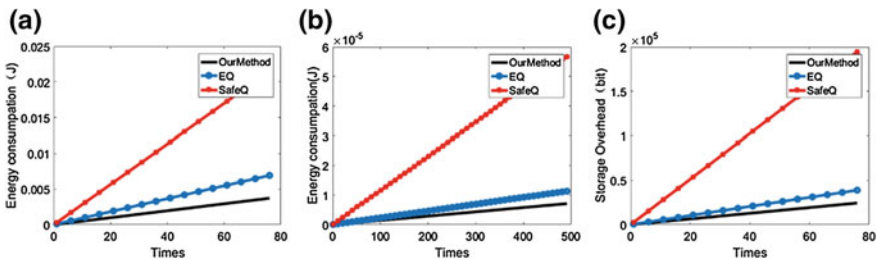


Fig. 3 The evaluation results. **a** Energy consumption of transmitting encrypted data. **b** Energy consumption of transmitting query results. **c** Space overhead in the storage node

Space Overhead in Storage Node: As shown in Fig. 3c, we can obtain that our method is much better than SafeQ-Bloom in terms of space overhead. What is more, EQ consumes 1.6 times more space overhead than our method.

6 Conclusion

We propose an efficient protocol for handling range queries in two-tiered sensor networks in a privacy- and integrity-preserving fashion. The protocol allows the sink to correctly handle the query, it also protects the privacy of the data and the integrity of the query results effectively. We use the stream cipher and S2L structure to achieve the protection of data privacy and data integrity. In this paper, the simulation test is carried out and the experimental results show that our method has lower communication energy and storage space cost than SafeQ-Bloom and EQ.

Acknowledgments The authors thank the anonymous referees for suggestions that helped to improve the paper substantially. And the project is partly sponsored by the Construct Program of the Key Discipline in Hunan province, the National Natural Science Foundation of China (No. 61640210, No. 61672447), the CERNET Next Generation Internet Technology Innovation Project (No. NGII20160305), and the Upgrading Project of Industry-University—Research of Xiangtan University (No. 11KZ|KZ03051).

References

1. Zanjireh MM, Larijani H (2015) A survey on centralised and distributed clustering routing algorithms for WSNs. In: Vehicular technology conference, IEEE, pp 1–6
2. Sheng B, Li Q (2011) Verifiable privacy-preserving sensor network storage for range query. *IEEE Trans Mob Comput* 10(9):1312–1326
3. Chen F, Liu AX (2010) SafeQ: secure and efficient query processing in sensor networks. In: Conference on information communications, IEEE Press, pp 2642–2650
4. Cheng J, Yang H, Wong S HY et al (2007) Design and implementation of cross-domain cooperative firewall
5. Bloom BH (1970) Space/time trade-offs in hash coding with allowable errors. *Commun ACM* 13(7):422–426
6. Tsou YT, Lu CS, Kuo SY (2013) Privacy- and integrity-preserving range query in wireless sensor networks. In: Global communications conference, IEEE, pp 328–334
7. Ma L, Liu J, Yao Y (2012) Privacy-preserving top-K query in two-tiered wireless sensor networks. *Int J Advancements Comput Technol*
8. Delavar AG, Baradaran AA (2012) CRCWSN: presenting a routing algorithm by using re-clustering to reduce energy consumption in WSN. *Int J Comput Commun Control* 8(1):61

Simulation Study on High-Speed Milling Performance of Rail U71Mn Material

Chao Pan, Jianqiang Zhang, Yuelei He, Haitao Xia
and Baosheng Wang

Abstract The high-speed milling of rail is one of the important directions in the development of rail maintenance technology. In order to study the high-speed milling performance of rail material, the U71Mn material milling simulation model was established using metal cutting simulation software AdvantEdge to analyze the changing rule of cutting force and cutting temperature during milling. Results show that the cutting force decreased with the increase of the milling speed of the rail, and the cutting force decreased with the cutting speed of 1200 r/min. The cutting heat increased significantly with the cutting speed, and the heat is mainly taken away by the chips, and the temperature rise is small. The research results to achieve high-speed rail milling have important guiding significance.

Keywords U71Mn · High-speed milling · Milling force · Milling temperature

1 Introduction

The mobile maintenance of rail milling is an important means for ensuring the status of rail. With the development of technology, the method of rail maintaining will be developed from grinding to multiple methods of milling, from low speed to

C. Pan (✉) · J. Zhang · H. Xia
Shanghai Railway Administration Science and Technology Institute,
No. 251 Tianmudong Road, Jing an District, Shanghai, China
e-mail: apan626@126.com

Y. He
Collaborative Innovation Center of Detection and Assessment for Operation
Safety of Railway Transit, Shanghai University of Engineering Science,
No. 333 Long Teng Road, Song Jiang District, Shanghai, China
e-mail: hyl doc@163.com

B. Wang
Research Department of Intelligent Manufacturing Equipment, Nanjing Institute
of Technology, No. 1 Hongjing Road, Jiangning District Nanjing, China
e-mail: clxwbs@njit.edu.cn

high speed, and the maintenance efficiency and quality will be greatly enhanced. The high-speed mobile milling of rail will be one of the important directions in rail maintenance technologies.

Based on the theory of high-speed milling, it can greatly improve the material removal rate and raise the surface processing quality of the material. Since rail is a long and huge workpiece, with the continuous working of the equipment, the high-speed milling technology is suitable for online maintenance of rail. Scholars in China and abroad have done a lot of studies on the high-speed properties of metal materials, especially on the physical performances such as milling process and cutting force, cutting heat and made a series of achievements. Pang et al. conducted an experiment on high-speed milling 45 Steel, and studied the influence of cutting speed, workpiece hardness, material properties on cutting force, and made prediction studies on the surface roughness of 45 Steel in milling [1–3], Su Fa et al. studied the high-speed milling performances of Cr12-quenched steel and analyzed the influence of milling speed, back engagement feeding speed on milling force, Wang et al. studied the wearing status and wearing mechanism of hard alloy tool with different surface coatings on quenched steel; Cui et al. studied the influence of CBN tool on the form of chips and roughness on the surface of workpiece in the milling process of quenched steel [4–8]. There are many varieties and types of steel, however, there are few reports on the studies of the high-speed milling performance of rail.

This paper adopted the simulation software AdvantEdge of metal cutting, established a high-speed milling simulation model on the large radius hard alloy tool with TiAlN surface coating, studied and analyzed the changing rules of cutting force, cutting temperature distribution under different milling speeds, and laid a solid theoretical foundation for manufacturing of a new generation of mobile online high-speed rail milling equipment.

2 Finite Element Modeling

2.1 Constitutive Model

The constitutive relationship of material describes the influence of strain, strain rate, and temperature on flowing stress; it plays an important role in the simulation and processing of finite element modeling and is directly related to the precision in simulation. This paper adopted the power law material model, which has isotropic elastoplastic performances and stress hardening, the stress function of the model is

$$\sigma(\varepsilon^p, \dot{\varepsilon}, T) = g(\varepsilon^p) \times \Gamma(\dot{\varepsilon}) \times \Theta(T) \quad (1)$$

where $g(\varepsilon^p)$ is the working hardening function, $\Gamma(\dot{\varepsilon})$ is the stress velocity effect function, and $\Theta(T)$ is the thermal softening function.

The stress rate in the power law can be expressed as follows:

$$1 + \frac{\dot{\varepsilon}^p}{\dot{\varepsilon}_0^p} = \left(\frac{\bar{\sigma}}{\sigma_f(\varepsilon^p)} \right)^{m_1}, \quad (2)$$

where $\bar{\sigma}$ is the von Mises stress, ε^p is the accumulated plastic strain, $\dot{\varepsilon}_0^p$ is the plastic strain rate for reference, σ_f is the flow stress, and $\dot{\varepsilon}^p$ is the plastic strain rate.

The flow stress in the power law can be expressed as follows:

$$\sigma_f = \sigma_0 \Theta(T) \left(1 + \frac{\varepsilon^p}{\varepsilon_0^p} \right)^{1/n} \quad (3)$$

where σ_0 is the initial yield stress at the temperature of T_0 , $\Theta(T)$ is the thermal softening function, ε_0^p is the plastic deformation for reference, and n is hardening index.

2.2 Rail U71Mn and Parameters of Milling Tool

The steel of the rail is a special steel, and normally the high-carbon and micro-alloyed steel after hot-rolling and quenching. At present, the U71Mn is used on 60 kg/m rail in China, and its physical and mechanical properties are shown in Table 1.

The mobile rail milling is peripheral milling. The cutter is fixed with blades of different shapes and sizes. In order to save calculation time and ensure the reliability of the results, the model was simplified, and the cutting edge was simulated in a rectilinear motion. The milling tool material of the rail milling train is hard alloy, with the advantages of high hardness, good abrasive resistance, high temperature resistance, and corrosion resistance. The surface coating on the tool was TiAlN. The surface coating material as the chemical and thermal shield reduced the diffusion and chemical reaction between the tool and workpiece. TiAlN is a kind of ideal surface coating material, as its hardness has good chemical stability with a small affinity with metal, low friction coefficient, good lubricating properties, and can greatly reduce friction and wearing of the tool and prevent caking and cold welding. The cutter-head of rail milling train is large and its diameter was set as 600 mm as simulation parameter. In order to minimize simulation calculation, the teeth number of the tool was set as 2. The parameters of the cutters are shown in Table 2.

Table 1 Physical and mechanical properties of U71Mn

Tensile strength σ_b	Yield strength σ_s /(GPa)	Elasticity modulus E/(GPa)	Ductility $\delta_{0.5}$ (%)	Brinell hardness HBS
0.98–1.03	0.77–0.82	211	50–80	170–220

Table 2 Parameters of cutters

Material	Hard alloy
Surface coating	TiAlN
Thickness of surface coating μ	5
Diameter (mm)	600
Teeth number	2
Front angle ($^{\circ}$)	11
Rear angle ($^{\circ}$)	1
Corner radius (mm)	0.02

2.3 Setting of Milling Parameters

Based on the actual working status of the rail milling train, the simulated parameters were set. During working, the disk milling cutters above the rail rotated around the horizontal shaft perpendicular to the rail and realized the climb milling of the rail, as is shown in Fig. 1. In order to avoid the adverse effect of cutting fluid on the friction coefficient and environmental pollution, dry milling was adopted. The simulation parameters can be found in Table 3, and the spindle speed increased gradually from 200 r/min.

3 Milling Process Simulation and Result Analysis

3.1 Influence of Milling Speed on Cutting Force

During the process of rail milling, cutting force determines the power consumed in milling and also directly influences the wearing of rail milling tool, quality of surface processing, precision and stability of milling system of the rail milling train. The simulation results of the influence of milling velocity on cutting force are shown in Fig. 2. Since climb milling was adopted in rail milling, at an early period

Fig. 1 Mobile rail milling

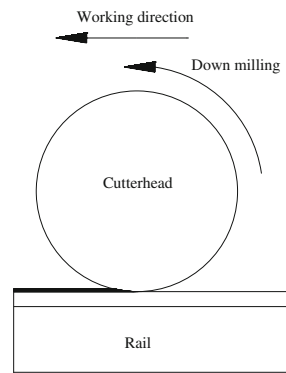
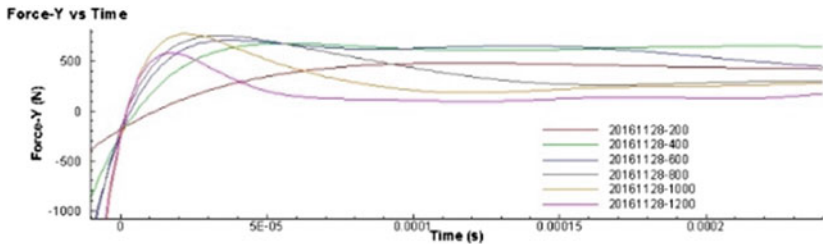
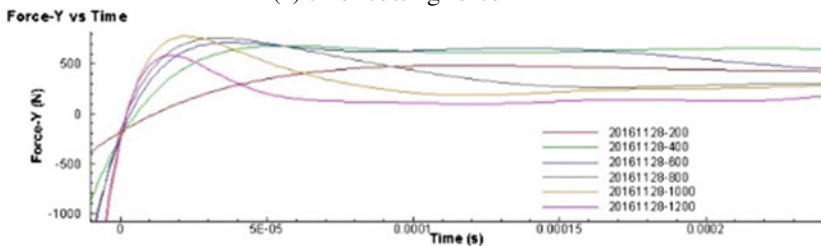


Table 3 Parameters of milling process

Milling method	Milling width (mm)	Feed per tooth (mm)	Milling depth (mm)	Initial temperature (°C)
Down	2	2	3	20



(1) chief cutting force



(2) Feeding resistance

Fig. 2 The change of cutting force with milling speed

of milling, the thickness of cutting materials by cutter tooth was the maximum, and with the rotation of the cutter, the milling thickness decreased, so the cutting force first reached a maximum value and then gradually decreased. With the increase in milling velocity, the cutting force also increased. When spindle speed reached to 1200 r/min, the chief cutting force and feeding resistance both decreased significantly. The decrease of cutting force renders great advantage to the high-speed milling of rail milling train, and can reduce the power consumption of the milling train, improve the working efficiency and velocity of the rail milling train.

3.2 Influence of Milling Speed on Cutting Temperature

During rail milling, the cutting heat directly influences the rail processing quality and wearing speed of the tool, the cutting temperature distribution of U71Mn steel is shown in Fig. 3. Figure 3 shows that the cutting heat is mainly within the contact area between the front tool face and chips. This was caused by the fact that a large

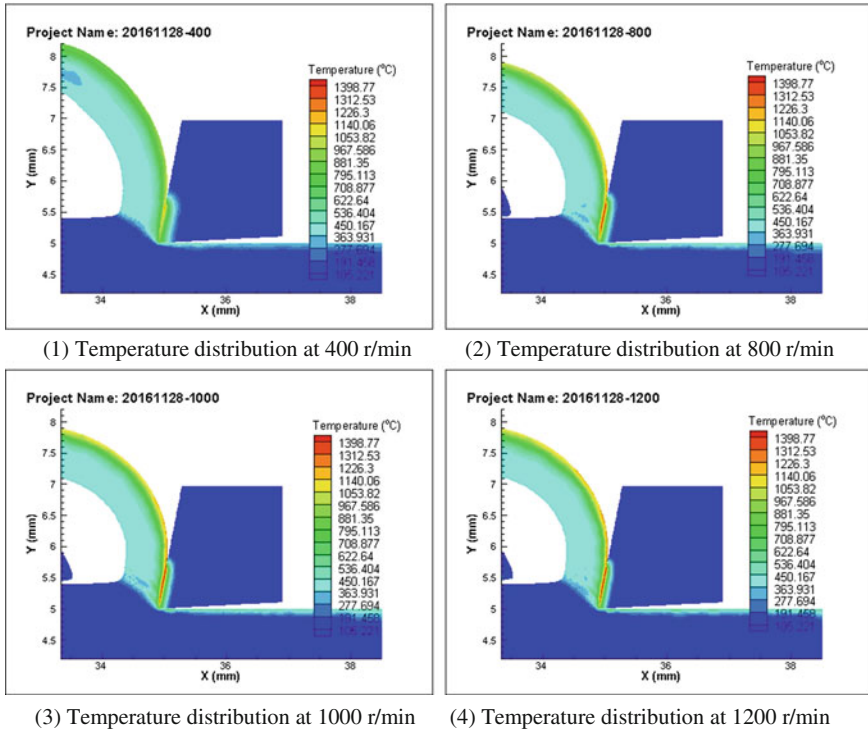


Fig. 3 Distribution of cutting temperature

amount of heat produced in friction could not be dissipated and the heat in friction gradually increased along the contact area. Figure 2 shows that, with the increase in milling speed, the high-temperature area was continuously expanded between chips and front tool face, and the area influenced by the heat of the chips also became larger. It shows that the heat produced in milling was removed by chips with the increase of milling speed. The heat affected zone in the front tool face gradually reduced, with increasing temperature gradient, besides, the maximum temperature area was not at the blade, but some distance from the blade. The total heat produced in the milling area increased rapidly, however, most of the heat was brought away by the chips, there was little heat transferred to rail and the tool, besides, with the increase of milling speed, the more heat was brought away by chips and less heat was transferred to rail, which would lower the temperature rise on rail surface and temperature stress [9].

Figure 4 shows the changing rules of peak temperatures under different milling velocities, and with the increase in milling speed, the peak temperature also increases gradually. The high cutting temperature will soften tool material and reduce the strength and milling performance of the tool, accelerate tool wear, and affect the working precision of rail and working quality on the surface.

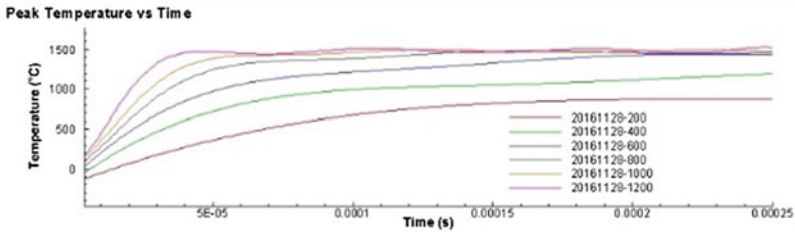


Fig. 4 The change of peak cutting temperature with milling speed

4 Conclusions

- (1) High-speed mobile milling of rail will be one of the important directions in rail maintenance technologies, and high-speed milling of rail can greatly improve the material removal rate and raise the working quality of rail surface.
- (2) In the process of milling, the cutting force also increased with the increase of milling speed, when spindle speed reached to 1200 r/min, the cutting force significantly decreased.
- (3) In the process of milling, the heat produced in the milling area and peak cutting temperature also increased with the increase of milling speed, but most of the heat was removed along with chips, and it was good for lowering temperature rise on rail surface and temperature stress.

Acknowledgments This work was supported by National Natural Science Foundation of China under Grant number 51405210.

References

1. Pang JZ, Wang M, Duan C et al (2007) Chip morphology and cutting forces in high speed end milling of hardened P20 steel. *China Mech Eng* 18(21):2543–2546 (in Chinese)
2. Duan C, Hao Q (2015) Surface roughness prediction in high speed milling of 45 steel. *J Harbin Eng Univ* 36(9):1229–1233 (in Chinese)
3. Hao Q, Kong W, Hou H et al (2016) Experimental research on the influence of workpiece hardness on cutting force and surface roughness while milling 45 steel at high speed. *Manufact Technol Mach Tool* 8:83–86 (in Chinese)
4. Wang Y, Han Z, Ye G et al (2015) Surface integrity and fatigue life in high-speed milling of die steel SKD61. *Comput Integr Manufact Syst* 21(11):2995–3000 (in Chinese)
5. Lu J, Hu Z, Song Y et al (2010) Experimental study on high speed milling force of quench steel cr12. *Mach Des Manufact* 2:170–171 (in Chinese)
6. Su F, Liu Y, Lian G (2016) Experimental study of high speed milling on PAC5000 die steel. *Mach Tool Hydraulics* 44(15):105–111 (in Chinese)

7. Cui X, Guo J, Zhao J et al (2015) Chip temperature and its effects on chip morphology, cutting forces, and surface roughness in high-speed face milling of hardened steel. *Int J Adv Manuf Technol* 77:2209–2219 (in Chinese)
8. Wang CY, Xie YX, Qin Z et al (2015) Wear and breakage of TiAlN- and TiSiN-coated carbide tools during high-speed milling of hardened steel. *Wear* 336–337:29–42
9. Pang L, Hosseini A, Hussein HM et al (2015) Application of a new thick zone model to the cutting mechanics during end-milling. *Int J Mech Sci* 96–97:91–100

Adaptive Control with Asymptotic Stability Guarantees for High-Speed Train Systems with Uncertain Input Nonlinearities

Dongyue Yang and Lin Niu

Abstract The running process of high-speed train is a complex, multivariable, and nonlinear dynamic system with many uncertain factors. Nonlinear, time-varying and other factors make it difficult to achieve the desired performance with the linear time-invariant feedback controller. A class of adaptive learning control method is proposed to solve the problem of unknown time varying in high-speed train operation. The parametric system is used to describe the high-speed train dynamics model, and the nonlinear characteristics of the system are studied. A parameter-adaptive learning control method is given. Based on the Lyapunov functional analysis method, it is proved that the adaptive learning control method and its improved form can guarantee the train to converge to the desired speed trajectory every point. Finally, numerical simulations are conducted to verify the effectiveness of the proposed method.

Keywords Adaptive learning control • High-speed train • Traction and braking Uncertain system

1 Introduction

With the development of high-speed trains (HST), automatic driving has become an essential means of modern train control which is needed to accurately control train traction and braking system [1].

The theory and application of advanced control strategy has made great progress in recent years, and some achievements have been made in the research of train

D. Yang (✉)

Overseas Business Division, CRRC Zhuzhou Locomotive Co., Ltd.,
Zhuzhou, Hunan 412000, China
e-mail: 362100729@qq.com

L. Niu (✉)

Honghe University, Mengzi, Yunnan 661100, China
e-mail: 1220360837@qq.com

safety operation control. However with the increase of train running speed over 350 km/h, the traditional train operation safety mode cannot satisfy the development demand of high-speed train. Compared with the traditional low-speed train, braking and tracking control faces great challenge in high-speed train, such as system parameter uncertainty, model uncertainty and resistance uncertainty, hierarchical traction and braking input nonlinearities and input saturation index, the brake device circuit structure, non-affine characteristics of train. Therefore, the adjustment of train traction/braking control faces the following problems:

- (1) Aerodynamic effect: The traditional medium and low-speed trains have little influence on the stability and safety of the train because of its low running speed. Therefore, in most of the existing train control technologies, the influence of aerodynamic effects on operational safety has not been studied deeply. For high-speed trains with speed of over 350 km/h, the aerodynamic effect is very obvious, which is enough to endanger the safety of train operation. Therefore, it must be taken as an important factor to consider and study the control strategy and method of high-speed train safety operation satisfying the aerodynamic law [2, 3].
- (2) Kinetic model: In the traditional research on low-speed train control, the connection of cars of a train can be considered as rigid connection, thus operation safety of traction/braking control can be studied with the train as a whole for dynamic modeling, ignoring the interaction characteristics between cars and the operation safety influence. However, for the actual high-speed train, each car is connected by spring and damping mechanism, and the car has different microscopic forms in speed, displacement, posture, and vibration characteristics. When the train is running, these microscopic effects will evolve into important factors that affect the safety of the operation [4–6].
- (3) Inherent uncertainty, nonlinear and non-affine characteristics of the system: The operation of the train can be changed by many uncertain factors, such as weather conditions, line curvature, and slope, the weight of the car, which makes it very difficult to accurately measure the drag coefficient, high-speed train additional resistance, and other environmental disturbances. At the same time, the passenger numbers and train freight volumes are different in operation stages, so that the quality of high-speed train changes in the process of train operation, the time-varying characteristics makes it difficult to establish the precise model and accurate information to the measurement system parameters of the train system [7, 8].
- (4) Traction/braking system failure: It is difficult to avoid the fault of traction/braking system in long time running of high-speed train. This requires that train control system has the fault tolerance ability to deal with the fault automatically. At present, the fault detection and diagnosis and redundancy technology have their advantages, but also have shortcomings. Therefore, it is necessary to establish more effective fault response system [9].

A novel adaptive control method is proposed to solve the above problems.

2 Dynamic Model of High-Speed Train

Consider a multi-mass point model of trains. Each carriage of the train is regarded as a mass point, and each point is connected by a link system. Model form is taken as follows:

$$m_i \ddot{x}_i = \ddot{F}_i - f_{di}, \tag{1}$$

where $m_i (i = 1, \dots, n)$ for the i carriage mass, $F_i (i = 1, \dots, n)$ is traction force or braking force for corresponding driving, $f_{di} (i = 1, \dots, n)$ indicates the resistance used for the i carriage. It is necessary to point out that the exact value of m_i is not easy to be determined due to the different passenger number and at each carriage.

3 Adaptive Control

Rewriting formula (1) forms following adaptive control format:

$$\begin{aligned} \dot{x}_i &= \sum_{j=1}^R \eta_j(\varpi) (\kappa_i^j(X_i) + \vartheta_i^j(X_i)x_{i+1}) \\ \dot{x}_n &= \sum_{j=1}^R \eta_j(\varpi) (\kappa_n^j(X_n) + \vartheta_n^j(X_n)u), \end{aligned} \tag{2}$$

where $X_n \in R^n$ is state vector, $u \in R$ is the input signal, $i = 1, 2, \dots, n - 1$, $X_i = [x_1, x_2, \dots, x_i]^T$ and $\eta_j (j = 1, 2, \dots, R)$ are interpolating functions, $\varpi \in R^q$ is an exogenous scheduling variable.

Denote for convenience

$$\begin{aligned} \kappa_i^c(X_i, \sigma) &= \sum_{j=1}^R \eta_j(\sigma) \kappa_i^j(X_i) \\ \vartheta_i^c(X_i, \sigma) &= \sum_{j=1}^R \eta_j(\sigma) \vartheta_i^j(X_i). \end{aligned} \tag{3}$$

Assume that κ_i^c and ϑ_i^c are sufficiently smooth in their arguments, and for all $X_i \in R^i$ and $\sigma \in R^i$, $i = 1, \dots, n$ we have

$$\begin{aligned} \kappa_i^c(0, \sigma) &= 0 \\ \vartheta_i^c(X_i, \sigma) &\neq 0, \end{aligned} \tag{4}$$

where the interpolation functions η_j are given, the functions κ_i^c and ϑ_i^c are uncertain, and a stabilizing control law can be built based on the approximations to the plant dynamics.

For each vector X_i , we assume the existence of a compact set $S_i \subset R^i$ such that $S_i \subset S_{i+1}, i = 1, \dots, n - 1$. We will also assume

$$0 < \bar{\vartheta}(X_i\sigma) \leq \vartheta_i^c(X_i, \sigma), \quad i = 1, \dots, n - 1. \tag{5}$$

For the system dynamics, we let

$$\kappa_i^j(X_i) = \theta_{\kappa_i^j}^{*T} \zeta_{\kappa_i^j}(X_i) + \delta_{\kappa_i^j}(X_i) \quad \text{for all } X_i \in S_i, \tag{6}$$

$$\vartheta_i^j(X_i) = \theta_{\vartheta_i^j}^{*T} \zeta_{\vartheta_i^j}(X_i) + \delta_{\vartheta_i^j}(X_i) \quad \text{for all } X_i \in S_i. \tag{7}$$

To obtain optimal parameter vector $\theta_{\kappa_i^j}^*$ and $\theta_{\vartheta_i^j}^*$, we let

$$\begin{aligned} \theta_{\kappa_i^j}^* &= \arg \min_{\theta_{\kappa_i^j} \in \Omega_{\kappa_i^j}} \left[\sup_{X_i \in S_i} \left| \kappa_i^j - \theta_{\kappa_i^j}^T \zeta_{\kappa_i^j}(X_i) \right| \right] \\ \theta_{\vartheta_i^j}^* &= \arg \min_{\theta_{\vartheta_i^j} \in \Omega_{\vartheta_i^j}} \left[\sup_{X_i \in S_i} \left| \vartheta_i^j - \theta_{\vartheta_i^j}^T \zeta_{\vartheta_i^j}(X_i) \right| \right]. \end{aligned} \tag{8}$$

Choose the function approximators as

$$\begin{aligned} \widehat{\kappa}_i^c(X_i, \sigma) &= \sum_{j=1}^R \eta_j(\sigma) \widehat{\theta}_{\kappa_i^j}^T \zeta_{\kappa_i^j}(X_i), \\ \widehat{\vartheta}_i^c(X_i, \sigma) &= \sum_{j=1}^R \eta_j(\sigma) \widehat{\theta}_{\vartheta_i^j}^T \zeta_{\vartheta_i^j}(X_i). \end{aligned} \tag{9}$$

The parameter errors for the j th approximator in the i th state are $\Phi_{\kappa_i^j} = \theta_{\kappa_i^j}^* - \widehat{\theta}_{\kappa_i^j}$ and $\Phi_{\vartheta_i^j} = \theta_{\vartheta_i^j}^* - \widehat{\theta}_{\vartheta_i^j}, i = 1, \dots, n - 1, j = 1, 2, \dots, R$.

Assume $\zeta_{\kappa_i^j}$ and $\zeta_{\vartheta_i^j}$ to be at least $n - i$ times continuously differentiable for $i = 1, \dots, n - 1, j = 1, 2, \dots, R$

$$\left| \frac{\partial^{n-1} \zeta_{\kappa_i^j}}{\partial X_i^{n-i}} \right| < \infty, \quad \left| \frac{\partial^{n-1} \zeta_{\vartheta_i^j}}{\partial X_i^{n-i}} \right| < \infty. \tag{10}$$

4 Adaptive Control Theorem

In case of the scheduling vector ϖ_{n-1} bounded and available for measurement, assume that $X_i(0)$ lies within a sufficiently small subset of $S_i \subset R^i$ $i = 1, \dots, n-1$, where S_i are compact sets specified a priori. Consider the following diffeomorphism:

$$\begin{aligned} y_1 &= x_1 \\ y_i &= x_i - \alpha_{i-1}^s - \alpha_i - 1, \quad i = 2, \dots, n, \end{aligned} \quad (11)$$

where $\alpha_1 = \frac{1}{\vartheta_1^c} (-\hat{\kappa}_1^c - c_1 y_1)$, $\alpha_1^s = -\frac{y_1}{\vartheta_1} (\beta_{11} + \beta_{12} \alpha_1^2)$.

For $m = 2, \dots, n$

$$\begin{aligned} \alpha_m^s &= -\frac{y_m}{\vartheta_m} \left(\beta_{m1} + \beta_{m2} \alpha_m^2 + \sum_{i=1}^{m-1} \beta_{i1} \left(\frac{\partial \alpha_{m-1}}{\partial x_i} + \frac{\partial \alpha_{m-1}^s}{\partial x_i} \right)^2 + \sum_{i=1}^{m-2} \beta_{i2} x_{i+1}^2 \cdot \right. \\ &\quad \left. \left(\frac{\partial \alpha_{m-1}}{\partial x_i} + \frac{\partial \alpha_{m-1}^s}{\partial x_i} \right)^2 + 2\beta_{(m-1)2} x_m^2 \left(\frac{\partial \alpha_{m-1}}{\partial x_{m-1}} + \frac{\partial \alpha_{m-1}^s}{\partial x_{m-1}} \right)^2 + 2\beta_{(m-1)2} y_{m-1}^2 \right), \end{aligned} \quad (12)$$

where $c_i > 0$, $\beta_{i1} > 0$, $\beta_{i2} > 0$, $i = 1, 2, \dots, n$ are design constants, $\gamma_{\kappa_i^j} > 0$, $\gamma_{\vartheta_i^j} > 0$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, R$ are adaptation constants.

Define auxiliary functions ($m = 2, \dots, n$)

$$\tau_{\kappa_{i1}^j} = \eta_j \zeta_{\kappa_i^j} y_i - \frac{b_{\kappa_i^j}}{\gamma_{\kappa_i^j}} \hat{\theta}_{\kappa_i^j}, \quad i = 1, 2, \dots, n \quad (13)$$

$$\tau_{\vartheta_{i1}^j} = \eta_j \zeta_{\vartheta_i^j} y_i \alpha_i - \frac{b_{\vartheta_i^j}}{\gamma_{\vartheta_i^j}} \hat{\theta}_{\vartheta_i^j} \quad i = 1, 2, \dots, n \quad (14)$$

$$\tau_{\vartheta_{(m-1)2}^j} = \tau_{\vartheta_{(m-1)1}^j} - \eta_j \zeta_{\vartheta_{(m-1)1}^j} \times \left(\left(\frac{\partial \alpha_{m-1}}{\partial x_{m-1}} + \frac{\partial \alpha_{m-1}^s}{\partial x_{m-1}} \right) y_m x_m - y_{m-1} y_m \right) \quad (15)$$

$$\tau_{\kappa_{i(m+1-i)}^j} = \tau_{\kappa_{i(m-i)}^j} - \eta_j \zeta_{\kappa_i^j} \times \left(\frac{\partial \alpha_{m-1}}{\partial x_i} + \frac{\partial \alpha_{m-1}^s}{\partial x_i} \right) y_m, \quad i = 1, 2, \dots, m-1 \quad (16)$$

$$\tau_{\vartheta_{i(m+1-i)}^j} = \tau_{\vartheta_{i(m-i)}^j} - \eta_j \zeta_{\vartheta_i^j} \left(\frac{\partial \alpha_{m-1}}{\partial x_i} + \frac{\partial \alpha_{m-1}^s}{\partial x_i} \right) y_m x_{i+1}, \quad i = 1, 2, \dots, m-2 \quad (17)$$

Then $b_{\kappa_i^j} > 0$, $b_{\vartheta_i^j} > 0$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, R$

When $m = 3, \dots, n$ (Let $g_2 = 0$)

$$\begin{aligned}
 g_m = & - \sum_{i=1}^{m-2} \sum_{j=1}^R \left[\left(\sum_{l=2}^{m-1} y_l \left(\frac{\partial \alpha_{l-1}}{\partial \hat{\theta}_{\kappa_i^j}} + \frac{\partial \alpha_{l-1}^s}{\partial \hat{\theta}_{\kappa_i^j}} \right)^T \right) \gamma_{\kappa_i^j} \zeta_{\vartheta_i^j} \times \left(\frac{\partial \alpha_{m-1}}{\partial \hat{x}_i} + \frac{\partial \alpha_{m-1}^s}{\partial \hat{x}_i} \right) \right. \\
 & \left. + \left(\sum_{l=2}^{m-1} y_l \left(\frac{\partial \alpha_{l-1}}{\partial \hat{\theta}_{\vartheta_i^j}} + \frac{\partial \alpha_{l-1}^s}{\partial \hat{\theta}_{\vartheta_i^j}} \right)^T \right) \gamma_{\vartheta_i^j} \zeta_{\kappa_i^j} \times \left(\frac{\partial \alpha_{m-1}}{\partial \hat{x}_i} + \frac{\partial \alpha_{m-1}^s}{\partial \hat{x}_i} \right) x_{i+1} \right].
 \end{aligned}
 \tag{18}$$

Assume that the function $\zeta_{\kappa_i^j}(X_i)$ and $\zeta_{\vartheta_i^j}(X_i)$ satisfies the condition (10). Consider the adaptive laws for the parameter vectors

$$\hat{\theta}_{\kappa_i^j} \in R^{N_{\kappa_i^j}} \quad \text{and} \quad \hat{\theta}_{\vartheta_i^j} \in R^{N_{\vartheta_i^j}}$$

where $N_{\kappa_k^j} \in N$, $N_{\vartheta_k^j} \in N$, $i = 1, 2, \dots, n, j = 1, 2, \dots, R$.

Let $y_{n+1} = 0, x_{n+1} = 0$.

Then, the control law

$$u = \alpha_n + \alpha_n^s.
 \tag{19}$$

5 Simulations

Consider an example:

$$\begin{aligned}
 x_1(k+1) &= x_2(k) \\
 x_2(k+1) &= (1 - e^{-x_1(k)}) / (1 + e^{-x_1(k)}) - 0.2x_1(k) \sin(t) + u(k) + d(k) \\
 y(k) &= x_1(k),
 \end{aligned}
 \tag{20}$$

where $d(k) = 0.2 \cos(0.3t)$ is the external disturbance.

$y_d(t) = \sin(0.2t) + \cos(0.4t)$ is the reference trajectory.

Where we choose the internal noise to obey Gaussian distribution, and train the data sample with this distribution data.

5.1 Case 1—Free of Internal Noise

Figures 1 and 2 gives control signal $u(k)$, where the tracking performance of the output trajectory is $y(k)$ and the reference trajectory is $y_d(k)$.

Fig. 1 The output trajectories of $y(k)$ and $y_d(k)$

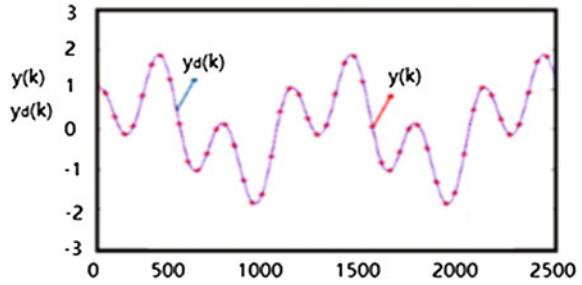
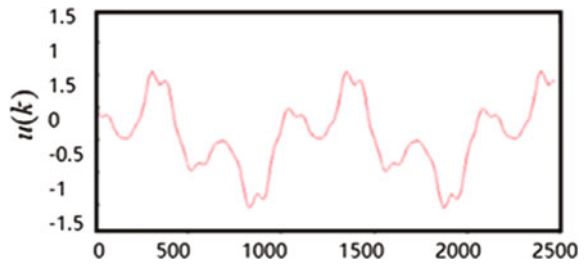


Fig. 2 Control input $u(k)$

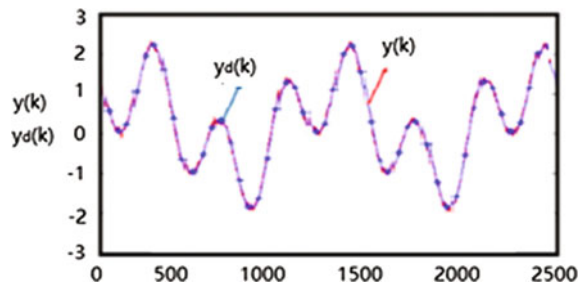


5.2 Case 2—Training Data Corrupted

The tracking performance of the output trajectory and the reference trajectory are shown in Figs. 3 and 4 respectively. Figure 5 gives mean square of the tracking error in Case 1 and Case 2 separately.

The experimental results show that the tracking performance is satisfactory.

Fig. 3 The output trajectories of $y(k)$ and $y_d(k)$



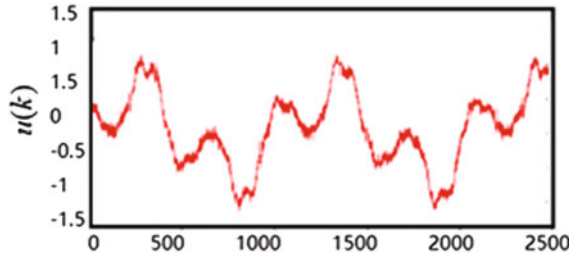


Fig. 4 Control input $u(k)$

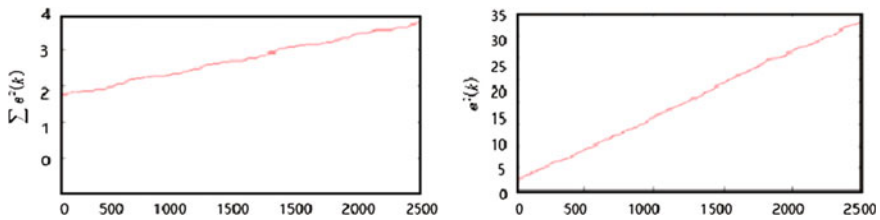


Fig. 5 Mean square of the tracking error of Case 1(left) and Case 2(right)

6 Conclusion

Aiming at the complex nonlinear system of high-speed train, this paper discussed systematically the train system model uncertainty, time-varying parameters, external disturbance and uncertain actuator/sensor fault traction, and braking tracking control problem, the train model is established based on nonlinear dynamics principle. Analysis of Lyapunov theory and related model based on the establishment of the traction and the brake-adaptive control method is reliable, stable tracking of the ideal velocity displacement trajectory makes the system uncertain and nonlinear in a variety of circumstances, to ensure that the train can run safely in normal conditions and possible fault conditions. The high-speed train tracking control problem are systematically analyzed and discussed, but there are still many problems to be further explored, such as the analysis of the relationship between the control performance and sensitivity of designed parameters.

Acknowledgements The project was supported by the National Natural Science Foundation of China (Grant No. 61463013).

References

1. Masoud M, Kent G, Kozan E (2016) A new multi-objective model to optimize rail transport scheduler. *J Transp Technol* 06(2):86–98

2. Gao R, Wang Y, Lai J (2016) Neuro-adaptive fault-tolerant control of high speed trains under traction-braking failures using self-structuring neural networks. *Inf Sci* 368:449–462
3. Wang Y, Song Y, Gao H (2016) Distributed fault-tolerant control of virtually and physically interconnected systems with application to high-speed trains under traction/braking failures. *IEEE Trans Intell Transp Syst* 17(2):535–545
4. Cai W, Liao W, Li D (2014) Observer based traction/Braking control design for high speed trains considering adhesion nonlinearity. *Abstr Appl Anal* 1:1–10
5. Zhou R, Zolotas A, Goodall R (2014) Robust system state estimation for active suspension control in high-speed tilting trains. *Veh Syst Dyn* 52(sup1):355–369
6. Shaltout R, Ulianov C, Baeza L (2015) Development of a simulation tool for the dynamic analysis of railway vehicle-track interaction. *Transp Probl* 10(SE):47–58
7. Liu D, Lechner B, Freudenstein S (2016) Evaluation of high-speed track quality using dynamic simulation of vehicle-track interaction. *J Transp Technol* 06(1):9–14
8. Liu Y, Chen C, Guoxing G (2011) Adaptive neural output feedback tracking control for a class of uncertain discrete-time nonlinear systems. *IEEE Trans Neural Networks* 22:1162–1167
9. Song Q (2014) Robust adaptive and fault-tolerant control of high speed trains [Ph.D]. Beijing Jiaotong University, Beijing (in Chinese)

Research on Key Technology of Platform Screen Door Control System for High-Speed Railway

Zhifei Wang

Abstract Platform screen door (PSD) control system is a complicated distribution parameter control system. It integrates construction, machinery, electronics control, and other disciplines at an organic whole. Aiming at the problem of low localization of the original PSD products, the paper realizes the autonomy of two core technologies, door machine and door control. Control system can be divided into signal system level, local control level, and the manual operation of three control levels from high to low order. The key technology realizing the platform screen door system can be divided into the central control panel design, gating unit design, and software design. At last, the model machine of control system has been tested in platform screen doors controlling experiment. The test results show that the system has an excellent dynamic performance and a high reliability, and thus the validity and the feasibility of the proposed approach are proved. The system has the advantages of high national productivity, reducing project cost, and owning independent intellectual property right.

Keywords High-speed railway · Platform screen door · Control system

1 System Architecture

PSD is composed of the machine and electrical part; the electric part comprises a power supply system and control system. PSD system is shown in Fig. 1. In the PSD system of Rail transit, the central Control Panel (PSC), monitoring system and power supply system are set in the heart of the equipment, in which the monitoring system is composed of the instruction processing module and the monitoring module [1–3]. The monitoring module via the CAN bus with each of the DCU door controller connected to the output end of the monitoring module LSL each display

Z. Wang (✉)

Institute of Computing Technology, China Academy of Railway Sciences,
No. 2 Daliushu Road, Haidian District, Beijing, China
e-mail: wfbuaa@126.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_100

985

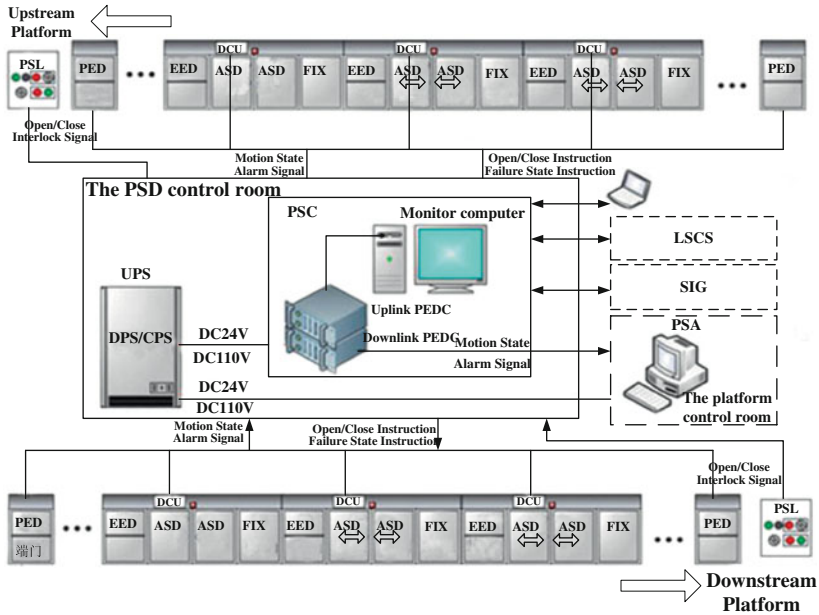


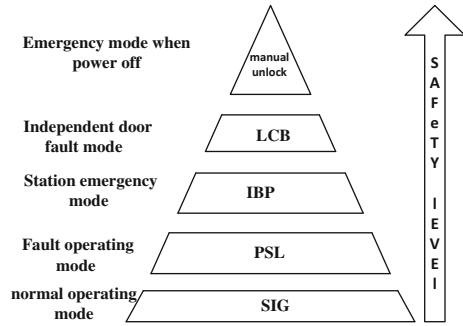
Fig. 1 The system diagram

connected to the input of the monitoring module with the JC safety detector connected via local bus and monitoring module PSA monitoring machine connected to a PC monitor module and portable two-way connection, the monitoring module, respectively, and the station control system signaling devices, electromechanical equipment monitoring device is connected; said PSL site controller output door controller with DCU input terminal; said platform security gate power supply input terminal connected to the station, providing the system components and the power circuit.

2 Control System

Control system can be divided into signal system level, local control level, and the manual operation of three control levels from high to low order. The Door Control unit (DCU) consists of a logic control unit, a motor drive control unit, an interface unit, and related software. As a security door control system at the core of the system, door control unit receives the control from the central controller command, the logic control unit deals with all kinds of signals, and then through the drive control unit control brushless DC motor has to perform open–close action; and through the MODBUS communication status information feedback of the door, the real time to effectively hold exit at work of all kinds of data is uploaded to the

Fig. 2 Priority sketch map of the screen door control system

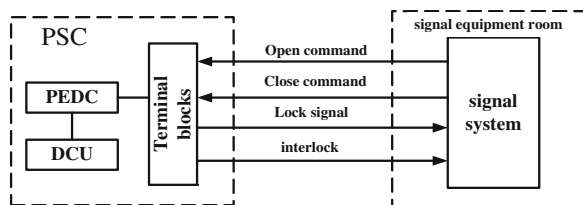


central processing ark, to guarantee the normal work of PSD. The screen door control system ranks from low to high as its priority includes signal system level, local control level, and manual control level. In the three levels, signal system level shares the lowest priority and the manual control shares the highest priority. The lower priority operations can be done only when the high priority level is done. Sketch map of the priority appeared is shown in Fig. 2.

3 Central Control Design

PSC includes cabinet, LCU, monitoring host and display terminal, and the signal system, port instruments of the ISCS, meters and the related cases, and condition lights of PSC board. The central control board includes two mutually independent PEDCs, each PEDC includes a control unit, controlling the opposite side's screen door at the stage, and the redundancy protection that is realized by the safe relay. The PSC connects the DCU, LCP, and IBP board through rigid line, to realize the key control and signal tickling. There is a monitoring host PLC at each side, it can realize the real-time communication with the very side's stage via the trunk function of the redundancy site, and to transfer the fault, operating condition to PLC through trunk. PSC also includes the rigid line port with signal system, port with ISCS system (RJ45 port), the internet port with display terminal of PSC board, related buttons and switches on cabinet boards and terminal block, indicating light, repair outlets, and so on. The wiring diagram of screen door system and signal system is shown in Fig. 3.

Fig. 3 The wiring diagram of screen door system and signal system



System on each side uses the delay logistic circuit to realize the redundancy protection of controlling function. When the PLC working normally, the relay switches to the PLC control mode, and all controlling input signals get into PLC. When the PLC comes with faults, the relay switches to relay mode and all controlling input signals get into relay circuit, unceasingly realizing the command making function of screen door.

The monitoring host completes the condition monitoring to screen door system. It realizes the system to collect and analyze on related information, realizes the information exchanging between DCU, LCP, and ISCS, and can do condition monitoring on signal system and IBP board port. The system owns the function of real-time monitoring and self-diagnosis. The system can monitor all monitoring sites of the stage, which can monitor the whole screen doors or independent door of the alarm, condition, accident information, and update to main monitoring system through live trunk network, the main monitoring system will do collection, analysis, and recording on these information, and do the information exchanging with the DCU units and external system. It can also do online altering to the operating data of DCU (including the choice on path of open/close door, clamp force when door closing, times of closing door when suffering barriers, delaying time when reopening doors, and so on) and automatically note the faults and major operating; the faults and operating notes share the functions of display, check, delete, and so on.

4 Design on Door Control System

DCU is one of the most important parts of the screen doors. The DCU receives all signals gives out by PSC control cabinet; controls switches of the sliding doors, electric lock, open/close door when suffering barriers, infrared block, and so on; realizes remote, manual, auto, and other modes; realizes what the real-time gives out of condition indicating, as well as that the specifications can pass the deploy of the upper computer. The DCU also returns signals to PSC cabinet monitoring computer to display the current condition of the screen door. DCUs of each side are connected in series, and the rigid lines are connected to the central control board at the end and head, forming the safety circuit (as shown in Fig. 9), with ring protection, to ensure any independent circuit's broke would not influence the open/close function of screen doors. The DCU fault influences only on single door. The central control board's control to DCU owns the safety relay redundancy protection. The system controls the LCB to allow the station or staffs use the LCB to change the operating modes when adjusting or maintenance; the major operating includes auto mode, manual turn on command, manual shut down command, and so on.

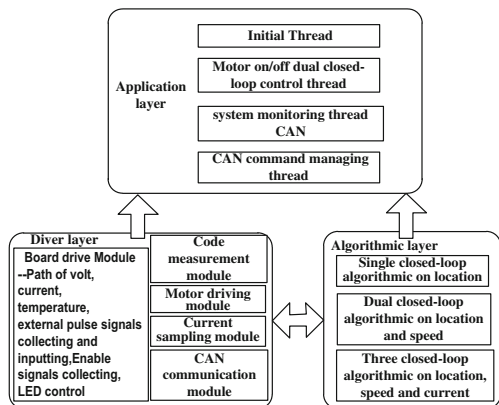
The DCU door control system consists of three parts: motor driving module, logical control module, and power sourcing module. The motor driving module mainly does the digitally control on motors based on the feedback from digital

encoder after algorithm. Logical control module achieves the demanded curvilinear motion controlling based on the returned signals after logical analysis and transfers the control order to motor driving module. The power sourcing module major concerns the actual applying space and the compatibility of strong and weak electric, and then designed the auxiliary control circuit module to realize the power supply for motor driving module, logical control module, controlling function of the 110 V YL, and relay safety circuit. The power sourcing module consists of four parts: motor driving module power, source, logical controlling module power source, safety control circuit, and YL logical control circuit (Fig. 4).

The functions completed by DCU door controlling include the following:

- (1) DCU stores the necessary velocity curve internally and sets up a series of doors clamp forces, reclosing interval time and the delay time, times, and velocity curve.
- (2) The DCU realizes the controlling on motors through the self-learned setting the velocity curve, can collect the condition information of doors, and locks the related faults information.
- (3) It shares the function of offline adjust on controlling units, through the internal encoding port to do resetting and encoding on specification.
- (4) Realize the control commands from system level and stage level.
- (5) DCU can control the condition indicating lights right including flashing, long-lighting, and long die-out.
- (6) Once the faults happen when the door opening or closing, the two processors of DCU can realize the fault from speed-down of motor and the increasing of current, so there is no need of ports and special sensor, which highly ensures the redundancy.

Fig. 4 Control processing



5 Software Design of the System

Parts of the system are high performance Interconnect ARM environment, the program written in C language. The system language uses the modular construction and time-sharing operation, to give out the control command and realize the screen door’s real-time control based on the signal system’s open/close command and sliding door condition (Fig. 5).

The high-speed railway platform screen door with the method of system control, the metro driven into station and stopped within the allowed mistake scope, the train signal system gives out the open/close command to the central control system, the control command of screen door sent to the DCU through signal system, DCU controls the sliding door in real-time, and realize the open/close operating on screen door. The processes of normally open on screen door are shown in Fig. 6.

When the high-speed train signal system is running right, the railway metro stops within the allowed mistake, and the signal system cancels the “OPEN” command and and starts the CLOSE program. The flow chart of the CLOSE control process is shown in Fig. 7.

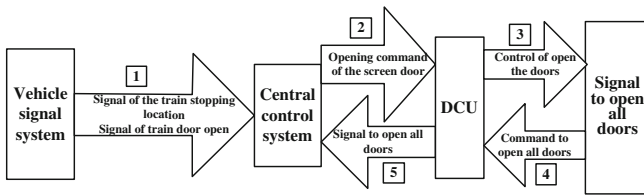


Fig. 5 The processes of normally open on screen door

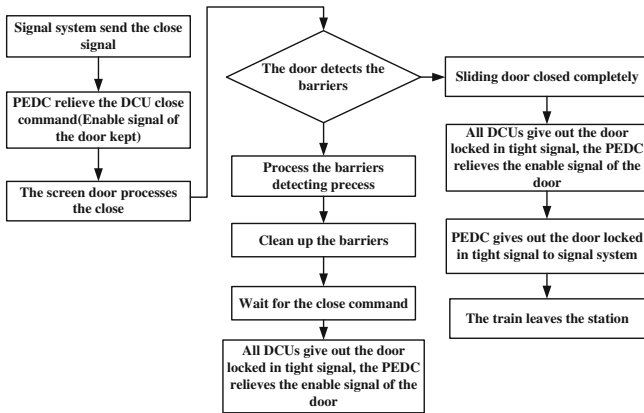


Fig. 6 The flow chart of the CLOSE control process

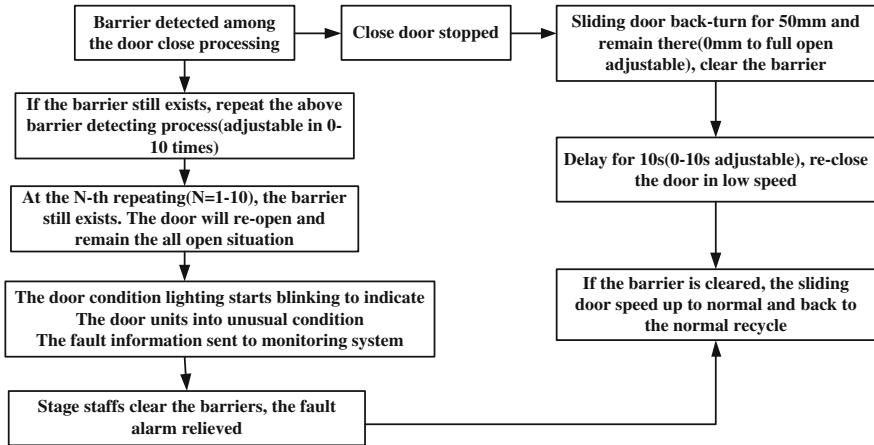


Fig. 7 Flow chart of the barriers detecting OPEN/CLOSE doors

The screen door system is running under the right condition; if the sliding door at any side suffers from the barriers, the sliding door will stop and can be open manually; and other sliding doors still under right condition. After delayed waiting (adjustable among 0–10 s), the door would try to open for three times; if the barrier still exists, the door condition lighting starts and the door is at a free condition till when the sliding door closed and locked.

When the signal system gives out the CLOSE command, the doors launching low-speed command would be shut down automatically. If the screen door suffers barriers during the CLOSE command, by receiving the CLOSE command, the screen door will turn to barriers detecting the function on CLOSE. The flow chart of the barriers detecting OPEN/CLOSE doors is shown in Fig. 7.

6 Experimental Verification

The experimental apparatus is shown in Fig. 8 consisting of platform screen door control system by China academy railway sciences. The platform screen door consists of one gated unit, a door machine system, a central control Panel, one in situ control disk, and a power supply system. Power switch uses AC380 power supply, door control unit using Dc110v power supply, central control panel using Dc110v, and \$literal power supply.

The door with door control unit to the design of this article has carried out the related test; the results are shown in Fig. 9. The first part of figure is door operation curve: door runs with a maximum shutter speed 580 mm/s, the open time of whole process is 2.5 s, door in place target speed keeps to prevent, and the passengers can restore to open the door after door in place. The second part of figure is door closing



Fig. 8 System test picture

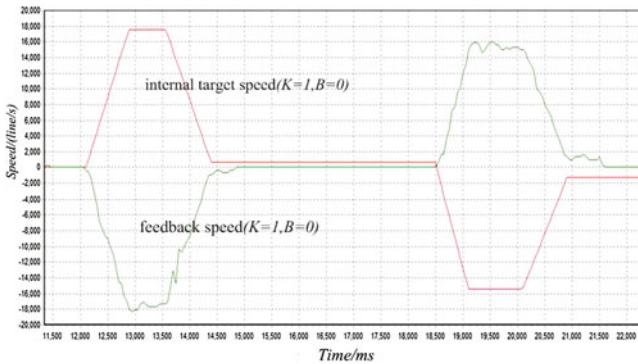


Fig. 9 The experimental speed curve

curve: door closes with a maximum shutter speed 510 mm/s and the close time of whole process is 2.5 s. Opening and closing time can be changed through the maximum opening and closing step-less adjustment of speed. Opening and closing time can be changed through step-less adjustment of the maximum opening and closing speed.

From the graph, we can see that internal target curve of motor driver and feedback speed curve of the waveform are in agreement. It explains that machine door control unit can better achieve security door controls.

7 Conclusion

The control system is the key technology of the platform screen door of high-speed railway, which will play a very important role in improving the safety and reliability of the high-speed railway platform. This paper focuses on the control system, which can be divided into signal system level, local control level, and the manual

operation of three control levels from high to low order. The key technology to realize the platform screen door system can be divided into the central control panel design and gating unit design. On this basis, the control system software design and control system algorithm are completed. The established control system is validated on the platform of the high-speed railway safety gate. The test results show that the system can complete the switch door of the platform door accurately and has the practical value of engineering.

References

1. Lao Y (2011) Metro platform screen door's control system research. South China University of Technology, Guozhou
2. Zhou C, Zhou J, Zhang H (2009) The system frame and control style of platform screen doors for bus rapid transit. *J Dev Innov Mach Electr Prod* 122(4):138–140
3. Haihui C, Yueming H, Jianming X (2002) Proposal of control system of metro platform screen door. *J S China Univ Technol (Nat Sci Ed)* 30(4):44–47

Information Platform Framework for the Integrated Management of Intercity Arterial Roads

Haiyang Wang

Abstract This chapter considers the comprehensive management of urban arterial roads, analyzes the current status and key components of integrated information management systems—in terms of road and transport related socioeconomic management issues—and describes the target and significance of the information system framework. By using system analysis combined with traffic engineering and relevant information system design, an integrated management information system framework, for an urban trunk road system, is put forward, including its main framework components and their functions.

Keywords Arterial road · Integrated management · Informatizaion System framework

1 Introduction

With rapid development of the road system in China, an ever increasing intercity arterial system, and transport-related socioeconomic activity, the traditional fragmented traffic management system used for arterial roads clearly cannot meet developmental needs. It is imperative and necessary to have an integrated management information system framework for effectively coordinating each of the components of road system management based on a sharing of information [1, 2].

By learning from the experience of developed countries in terms of road system management information systems, with their successful application of intelligent transportation system (ITS) technology in traffic monitoring; weight–load controls; roadwork-related data collection; and processing management information platform construction and information release, we can effectively improve the efficiency and capacity of the operation and management of the roads and traffic, and improve the

H. Wang (✉)

China Academy of Transportation Science, 240 Huixinli, Chaoyang District,
Beijing, China
e-mail: 273133342@qq.com

quality of the transport service available to the public in China. It is in this context that the current situation of developing an urban arterial road management information system will be considered [1].

2 Existing Problems of the Road Management Information System in China

Great effort has been made by each of the major cities in China in terms of research and development of road management information systems and their components (such as road traffic monitoring, traffic overload control, and roadwork and office management). However, there are many problems with the present mode of operation of China's management information system [3, 4].

- Each component of the road management system, in terms of its operation, is relatively independent and therefore the system lacks effective integration.
- Information hubs need to be well-organized and based on information sharing for an effective and efficient use of resources.
- There is a lack of development of hardware systems for database/information service centers.
- There is no uniform geographic information service platform to provide a practical visual display.
- There is a requirement for the development of more efficient, multifunctional, comprehensive data/information collection systems.
- A greater effort is required to develop multiterminal publications containing travel information and services.
- The existing automated systems running in network offices needs improvement.

3 Objective and Significance of the Framework Design

The framework design includes road traffic management, road facilities and management of equipment, and data/information management and services to support decision-making in both the public and private sectors. This study represents a comprehensive analysis of the key issues of road infrastructure development and the application of an intelligent transportation system. To this end, a number of frameworks are proposed for a comprehensive integrated management information system, including those required for road traffic control, transport facilities and equipment, traveler information and services, among others [5].

It is expected that such a comprehensive data/information service would be provided to administration departments at all levels in order to realize total office automation and daily management, for use in both public and private sectors,

ensuring timely, accurate, comprehensive, and adequate information services to support informed decision making and to enhance the capacity for scientific decision making in terms of routine management, emergency response, and first aid.

4 Design of the Integrated Road Management Information System

4.1 Overall Framework for the Information System Platform

On the basis of the current status of ITS research and the road management information system development in China, this chapter introduces the theory of management information-sharing platforms for the integrated development of the road management system, with an initial application of an integrated road management system framework based on the common information platform, as shown in Fig. 1. The framework consists of the following four levels [1].

1. A user interface layer through Browser/Server and Client/Server application for different users with different user application interfaces.
2. A middle application layer, including application frameworks, database engine, and a communication module responsible for transaction processing and control of database operations in order to achieve transmission and processing of data, accepting and processing client requests, and taking responsibility for the system's logic processing across a variety of applications.

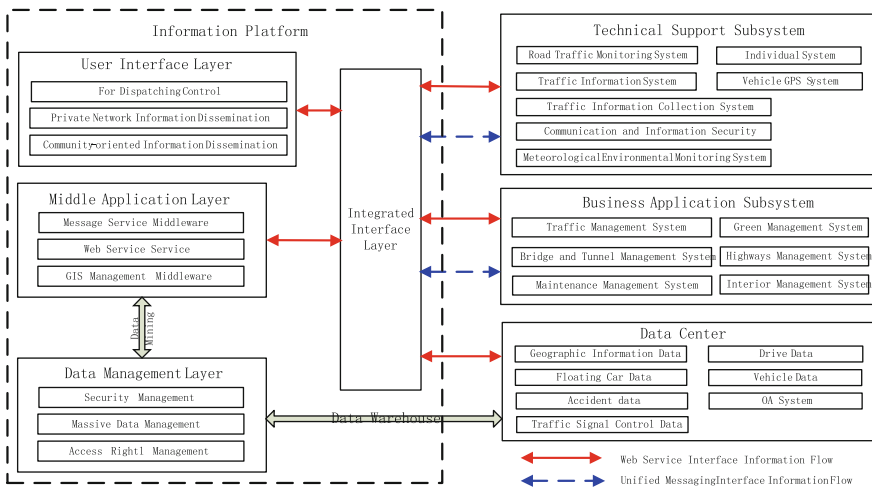


Fig. 1 Framework design for an integrated road management information system

3. A data management layer used as an underlying database for the entire system's operation, providing the necessary data services for the middle application layer, including vast amounts of spatial data and system information data.
4. An integrated interface layer responsible for information exchange and control between a management information platform and various service subsystems, through unified messaging and web service interface exchange of information via a technical subsystem and external systems, such as a traffic management data center. It should also be connected to a basic geographic information platform remotely via Spatial Database Engine.

In addition, the integrated information platform is composed of seven support subsystems and six business service subsystems, respectively. The supporting subsystems are implemented to ensure information platform functions. The business subsystems should realize data sharing and functional communication through the information platform [6].

4.2 Technical Support Subsystems

Traffic Information Collection System

Due to user demand for dynamic and static traffic data, this system is an information system incorporating dynamic road traffic operational data monitoring, through a seamless linked network, providing reliable, accurate, and real-time traffic performance data for the integrated road management information platform, permitting the release of traffic information and system services.

Road Traffic Monitoring System

Through a monitoring system, real-time images of traffic conditions, infrastructure status, and weather conditions can be compiled. The system includes timely detection of abnormal traffic events, including traffic jams, accidents, and vehicle identification (ID) monitoring, as well as the use of rest service areas, and the monitoring of abnormal real-time events, which can act as a basis for the emergency services [2].

Meteorological Environmental Monitoring System

Weather conditions represent one of the main impacts on road travel. This monitoring system is designed to monitor meteorological environmental conditions and information that may have an impact on road traffic, including visibility, weather-related surface conditions, and other relevant meteorological environment conditions.

Traffic Information Release System

The information release system relies on the fusion of traffic data/information and analysis, based on multi-sourced data/information collection, fusion, processing, analysis, and in-depth data excavation, with intelligent data calculation processing results seamlessly connected to the integrated management information platform, so

as to achieve the system functions allowing real-time data access, communication, and dynamic traffic information dissemination. The main release methods include variable message signs, traffic radio, a telephone customer service hotline, mobile apps, and internet sites [3–7].

Individual System

This system is defined as a system for on-duty officials with contingency communication equipment, operating through a 3G wireless network. This system makes it possible for the on-duty officials to send live images, location information, and record in real-time back to a command center. At the same time the system allows administrators at the command center to dispatch on-duty officials in real-time when needed for emergency responses [8]. The real-time communication equipment used in such emergency responses should be able to collect and store information and allow real-time communication.

Vehicle-Positioning System

The vehicle-positioning system is based on the functions of onboard computers (mobile electronic police) with cameras, and sound-recording functions, in addition to a 3G communication module and global positioning system (GPS) module, permitting mobile patrols, traffic law enforcement, and similar tasks. Using an information access platform, it is possible to achieve a two-way dialog between the platform and the vehicle, and ultimately to achieve an effective cross-regional uniform command for vehicle dispatch, monitoring, and management.

Communication and Information Security System

The system includes two types of information communication, one for information communication between different departments inside the command center, the other one for use between the command center and field. In the command center room are three types of internet system, i.e., the common Internet, a special network for road authorities, and a mobile Virtual special network. The communication system includes a front-end node, trunk transmission, and mobile communications. Information security is concerned with network security (including system-specific network security, a separation of the special network from the Internet, and security of information transmission) and software security (including data storage security, confidential data transmission, user password security, logging and analysis, and protection against viruses [3]).

4.3 Traffic Service Application Subsystem

Road Conditions Management System

This management system provides basic data maintenance, data query, and print functions. It not only provides a basic data platform for road maintenance, road administration, a geographical information system (GIS), and other subsystems but can also be used as a stand-alone subsystem. At the same time, this subsystem can

provide a record of all data items of historical information and data revocation functions in particular, to ensure integrity and accuracy of the underlying data.

Bridge and Tunnel Management System

This system maintains detailed records of basic information, as well as archived information, for bridges and tunnels, including records and analysis of routine management, regular inspections, maintenance, and management, with a focus on providing a variety of tools for data recording, mission planning, assessment analysis, and statistical reports required for bridge and tunnel administration. This system helps achieve a more scientific and efficient management of bridges and tunnels, as well as improving maintenance work.

Road Maintenance and Management System

As an important part of road operations, the maintenance and management system is an essential link to ensure the smooth implementation and completion of daily maintenance work. It is also an essential link to justify the necessary financial input and verify the success or failure of the maintenance work, so as to ensure a reasonable allocation of maintenance funds and facilitate informed decision making in maintenance and work planning. The main functions of this system include a statistical summary of minor repairs and monthly maintenance reports, road bulletins, road maintenance inspections, development of cost plans for minor repairs and maintenance, assessment of the technical conditions of roads, and basic data management [8].

Green Management System

This system relies on information technology being used for the management of tree planting, saplings, greening spaces, and other relevant areas. It includes rich botanical knowledge and is used to develop an early warning mechanism for pest control. The scientific data analysis and processing methodology of the system provides fast, accurate, high-quality pictures and texts for use by environmental departments.

Road Administration System

This system is in compliance with the relevant laws and regulations of the road, and makes use of computer technology and network technology as an auxiliary management tool. It establishes the scientific management and real-time monitoring mechanisms for enforcement, application reviews and approvals, road patrols, and obstacle removals. It continuously improves road administration by following the principles of responsibility, standardization, and transparency, in order to reduce duplicated work and improve efficiency and effectiveness.

Maintenance Management System

This system focuses on the scientific assessment of decisions taken for road maintenance as well as the efficiency of daily management. Its task is to achieve basic data collection, assessment of maintenance quality, evaluation of routine inspections, maintenance project management, and maintenance scheme evaluation based on full life cycle analysis of information from the management system [4].

5 Functional Design of the Information Platform

5.1 An Integrated Display Function Based on Geographical Information Systems

A GIS platform can transform data/information into digitalized traffic maps, i.e., the road network infrastructure, traffic volumes, resources for emergency response services, road conditions according to inspections, monitoring the weight of goods vehicles, traffic incidents, etc.

The system could be used to provide support for decision makers, establishing corresponding thematic maps based on decision maker demands.

5.2 Road Network Operational Monitoring and Early Warning System

This platform conducts correlation analysis of collected and shared data/information concerning the road network's environment, traffic operations, traffic incidents, road condition tests, road meteorological conditions, and geological hazards. The platform can also make road network condition assessments and dynamic forecasts, provide an analysis of developing road operational trends, provide a mid- and short-term impact forecast service, and offer the possibility of forecasting and analyzing unusual road network incidents [2].

5.3 Emergency Response Command and Dispatch Management

Based on the preliminary knowledge bank of emergency response command and management, the information service platform establishes a mode of enquiry into the resources used by the emergency services. In the event of an emergency, a set of emergency response procedures shall be followed, including initiating a preliminary plan, road network coordinated traffic control, and resource dispatch and monitoring. All of these are supplemented with emergency response information, including processing results and statistical analysis.

5.4 Information Dissemination and Public Information Services

There are two ways to achieve information release. One is an internal traffic management information release (by use of a browser, official users may logon to

the integrated management information platform set by the roads authority), the other is an external information release for the public (by use of a different website or media terminal, such as that of a vehicle management system, road department website, traffic radio broadcast, on-board terminal, road bureau WeChat platform, or road and other information broadcast apps). The information released would include: road infrastructure, service facilities, travel planning, traffic conditions, road accidents/incidents, roadworks, road environment, emergency rescue incidents, traffic administration, auxiliary service information.

5.5 The Facility's Intelligent Operation and Maintenance

This platform's function includes time corrections to the system, a showcasing of facilities and equipment, real-time monitoring of networked devices, alarm information automatic acquisition, manual prompting of malfunctions and relevant incidents, operational and maintenance issues, knowledge base, facility design, facility work management.

5.6 Global Positioning System Vehicle (Driver) Management

This platform uses standardized interfaces integrated with a GPS vehicle (driver), to achieve vehicle identification and display, fuzzy inquiry, vehicle tracking, sending instructions, track playback and etc.

5.7 Hotline Service

This platform uses an integrated interface connected to its users, to provide a hotline service, with configured functional modules, including those for dispatching communications, making digital recordings, and coordinating emergency responses, in order to achieve a thematic map display of hotline information, hotline service scheduling management, storage of digital recordings, incident management, as well as information enquiries, etc.

5.8 User Management Service

This includes the management service for unit users, individual users, defined system roles, and access rights.

6 Conclusions and Outlook

Based on analysis of an intercity urban arterial road information management system and ITS technology, this paper has established an integrated management information system framework for road system management, including a logical configuration of functional modules within the framework. The functions of seven subsystem were designed to support the information integration platform. The functions of six service functions were designed to establish an effective and efficient integrated road management information system framework.

In future studies, more emphasis should be given to ITS data-processing technology, so that data fusion, data mining, GIS, and other technologies associated with data processing can be better integrated in road transport system management, to further optimize and improve key technological applications to facilitate a continuous improvement of the road management information system.

References

1. Jiadong T (2008) Management information service. Dongbei University of Finance and Economics Press (in Chinese)
2. Guoqing W, Yantao L (2004) Function analysis and system design of a provincial highway network management information system. *J Transp Syst Eng Inf Technol* 4(1):97–101 (in Chinese)
3. Hua L (2010) The basic ideas of emergency management system of highway. *Study Times* (in Chinese)
4. Wen X, Xiangwu Y (2011) Highway road network management center system framework design. *J Highw Transp Res Dev* 1002-0268(2011)S1-0040-06, 40–45 (in Chinese)
5. Yuanhua J, Yanmin K (2002) Discussion on programs of expressway information management in China. *J Northern Jiao Tong Univ* 1000-1506(2002)02-0051-04, 51–54 (in Chinese)
6. Yingying H (2007) Highway information management system based on GIS. *J Qing Hai Transp Technol* U495:8–11 (in Chinese)
7. Zhonghua Z, Liming Z (2005) Highway management information system based on GIS. *J Chang'an Univ (Natural Science Edition)*. 1671-8879(2005)01-0069-04, 69–72. (in Chinese)
8. Yunpeng W, Hantao Z (2006) Research on the frame of expressway system integration based on common information platform. *J Highw Transp Res Dev* 1002-0268(2006)02-0128-05, 128–132 (in Chinese)

Study on a Mn–Cr–V–Ni Bainitic Forging Steel for Railway Transportation

Yunlei Lin and Qingyue Zhou

Abstract A bainitic forging steel for railway transportation was developed and produced with secondary refining, vacuum degassing, blooming casting, and billet slow cooling, with addition of silicon, manganese, chromium, nickel, and vanadium. The influence of cooling rates on the microstructure of the Mn–Cr–V–Ni alloy was studied by optical microscopy and SEM. And the influences of heat treatment on the mechanical properties and microstructure were investigated, the microstructure analysis and tensile tests were carried after 920 °C annealing–normalizing–tempering. The results show that carbide-free bainite structure was obtained by cooling from 950 °C to room temperature at less than 0.3 °C/s, martensite structure was obtained at 8.0 °C/s. A good combination of strength and toughness was obtained through the annealing–normalizing–tempering, with tensile strength of 1353 MPa, yield strength of 1057 MPa, an elongation of 17%, a reduction of 52%, impact toughness of 110 J/cm², and 42.0 HRC.

Keywords Bainitic steel · Railway transportation · Carbide-free bainite

1 Introduction

Railway transportation is one of the most primary traffic modes in China, and it plays an important role in national economy and country's public transport system. George Stephenson made the first steam locomotive with tubular boiler in 1814, then, the first passenger railway was opened between Stockton and Darlington.

Y. Lin (✉)

China Academy of Railway Sciences, No. 2 Daliushu Road, Haidian District,
Beijing, China

e-mail: carslinyunlei@163.com

Q. Zhou

Metals and Chemistry Research Institute, China Academy of Railway Sciences,
No. 2 Daliushu Road, Haidian District, Beijing, China

e-mail: zhouqingyue93480@126.com

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_102

1005

Over the years, railway has increasingly become a new vehicle, a combination of capacity, velocity, and reliability, alongside economic development and technical progress.

As an essential equipment of railroad, the function of rail is guiding train wheels to move forward, bearing and transferring load. Railway industry and relevant industry have always paid great attentions to research on rail both at home and abroad. 800–1280 MPa rails have been developed for various lines in U.S., Japan, Europe, China, etc., which offers technical guarantee to construct safe, reliable, rapid railway. However, now exist in the use and maintenance the following problems: rails in the sharp-radius curve have poor wear-resisting performance, upper rails appear side abrasion and spalling defects, lower rails appear corrugation and crack on surface. These issues influence the service life of rail and increase the maintenance workload of rail and wheel. A railroad turnout is one of the mechanical devices enabling railway trains to be guided from one track to another. As the operation condition of turnout is rigorous, the damage of turnout is fearful, as shown in Fig. 1.

In the wake of developments in high speed and heavy haul railway, more and more attentions are paid to the quality of rail, with a strong focus on strength–toughness properties in steel. In order to improve performance of available pearlitic rail steel, all countries are eagerly developing high strength and toughness bainitic steel. Bainitic steel rail has high strength, high toughness, good wear resistance, and high rolling contact fatigue (RCF) life, compared with pearlitic steel rail. Therefore, it has broad development prospect in future railway transportation technology.

The microstructure of bainitic steel usually consists of ferrite and cementite, similar to the microstructure of pearlitic steel. Pearlite has good plasticity and toughness, but it is limited to improve the strength to 1300 MPa or greater of pearlitic rail steel by decreasing the distance between rows of pearlite through alloying and heat treatment. So researchers fixed their sights on bainitic steel [1, 2]. Unlike that of typical lamellar pearlite structure, the morphology of ferrite in bainitic is mainly acicular or lath form and the carbides are discrete particles. Bainitic rail steel is superior in mechanical properties and fracture toughness to pearlitic rail steel. Yates [3] thought that it was difficult to improve the properties of pearlitic steel rail even more and he developed a low-carbon carbide-free bainitic steel. Pacyna [4],

Fig. 1 Spalling defects of turnout



a researcher of AGH University of Science, developed a new type bainitic steel RB390, the tensile strength, yield strength, elongation, and cracking resistance in dynamic conditions (KU2) of J6 were, respectively, 1347–1353 MPa, 825–832 MPa, 13.0–14.9%, and 73.1–79.5 J. It is worth mentioning that the KU2 and resistance to crack propagation were better than the same properties of traditional rails.

In China, Chen et al. [5] have studied the microstructure and mechanical properties of a kind of low-alloy bainitic steel. Huang et al. [6] have examined effects of a quenching-long partitioning (Q-LP) heat treatment on the microstructure and mechanical properties of a bainitic steel. They found that the impact toughness is increased from 83.8 J/cm² to a maximum value of 103.8 J/cm² after the Q-LP heat treatment with a partitioning time of 60 min.

In recent years, with the increase in the annual amount and axle load of heavy haul railway in China, the service conditions of rail become more and more harsh. Especially in small-radius curve, though heat treated rails with tensile strength of 1280 MPa, percent elongation of more than 8%, impact toughness of more than 30 J and harness of rail top surface of 370 HB–410 HB have been laid, upper rail side abrasion remains and shelling defects on lower rail occurs. It is urgently needed in developing high strength, good weldability, and contact fatigue resistance rail to meet the need of heavy haul railway development in China.

The 1180 MPa grade and 1280 MPa grade air-cooling bainite rail have been developed successfully. The 1180 MPa grade air-cooling bainite rail shows excellent performances in contact fatigue resistance and wear resistance after using in the Railway Administration of Beijing, Taiyuan, Chengdu, Shenyang, etc. The integrated performances are better than those of heat treated pearlite rail. But there are some problems existing in the 1280 MPa grade air-cooling bainite rail, such as serious micro-segregation, poor weldability, unstable quality, and so on.

In order to reduce cost as well as to further increase strength, toughness, weldability to ensure the safety, and prolong the service life, it is meaningful to develop a long-life rail mainly composed of lath martensite and lower bainite by using alloying treatment and online heat treatment in combination, which has tensile strength of 1300–1400 MPa and impact toughness greater than 100 J/cm².

In this paper, in order to seek a combination of high strength and toughness, a Mn–Cr–V–Ni bainitic forging steel for railway transportation is developed, and its mechanical properties and microstructures after heat treatment are studied.

2 Experimental Procedure

A Mn–Cr–V–Ni alloy was prepared as base material, as shown in Table 1. Manganese and chromium were added to lower the bainite start (Bs) temperature and increase the hardness of the steel, while vanadium was added to increase strength. The bainitic billets were produced with secondary refining, vacuum degassing, blooming casting, billet slow cooling, and so on. After the ingot was hot-forged and then hot-rolled to bars, four test pieces were cut from the bars.

Table 1 Composition (wt%) of studied alloy

C	Mn	Si	Cr	Ni	Mo	V
0.28–0.29	1.7–1.8	1.5–1.8	0.8–1.0	0.2–0.4	0.3–0.5	0.08–0.12

The tests consist of two parts: cooling test and heating treatment test. For part one, in order to study the microstructure of the test alloy after different cooling rates (referred to test 1, 2, 3), three samples were chosen to be austenitized at 10 °C/s to 950 °C for 5 min under vacuum, then the samples were cooled by chosen cooling rates (0.05, 0.3, 8.0 °C/s). Samples for microstructure analysis were mechanically ground, chemically polished, and etched in 4% Nital, and then microstructures were observed by optical microscope and transmission electron microscope.

For part two, combined with the results of above tests, the remaining one piece was subjected to a heat treatment (referred to test 4). The test piece was austenitized at 200 °C/h to 920 °C for 100 min, followed by furnace-cooling to room temperature. Then, the test piece was heated at 200 °C/h to a temperature (>850 °C) for 100 min, followed by air-cooling to room temperature. The last step was tempering.

For the piece treated by heat treatment, hardness, tensile, and Charpy-U impact tests were carried to measure the Rockwell hardness, the strength, plasticity, and toughness. The hardness was examined in a Rockwell hardness tester, the tensile testing was tested by MTS system, and the Charpy-U tests was tested by a pendulum impact testing machine. Samples for microstructure analysis were mechanically ground, chemically polished and etched in 4% Nital, and then microstructures were observed by optical microscope and scanning electron microscope.

3 Results and Discussion

3.1 Roles of Alloy Elements

Carbon is an essential part of steel, it can form carbides and solid solution. To some extent, the strength and wear resistance of steel increase with the increase of carbon content. But high carbon content will decrease toughness and influence welding quality. And carbon contributes to stabilize austenite, because of its high solubility in austenite and low solubility in ferrite.

Silicon is a non-carbide forming element and one of the strengthening elements in steel. In low-carbon bainite steel, the amount of retained austenite and its heat stability and mechanical stability will improve along the increase of silicon [7]. Silicon can suppress precipitation of cementite and bainitic ferrite. In general, the silicon content of steel is 1–2.5%.

Manganese and Nickel [8] are added to enhance the hardenability of martensite and bainite, and improve the amount of them by reducing the temperatures of martensite and bainite phase transformation. Meanwhile, manganese and nickel promote the formation of austenite.

Chromium is a crucial and functional element in steel, it can suppress ferrite formation and increase hardenability of bainite during continuous cooling [9]. Chromium is added to refine the microstructure of bainite and enhance the strength and toughness of steel. But chromium carbide and nitride are easily formed in steel.

Vanadium is a strong carbide forming element, it dissolves in ferrite or exists in carbo-nitride form in steel mostly. Vanadium plays a role of precipitation strengthen and it increases retained austenite content.

3.2 The Effects of Different Cooling Rates to the Microstructures of the Mn–Cr–V–Ni Alloy

Three samples were austenitized at 10 °C/s to 950 °C for 5 min under vacuum, then cooled by different cooling rates (0.05, 0.3, 8.0 °C/s). Microstructures of test alloy are listed in Table 2 and Figs. 2 and 3. The microstructure of the Mn–Cr–V–Ni alloy consisted of bainite and a small amount of proeutectoid ferrite after cooling at 0.05 °C/s, filminess retained austenite existed discontinuous between bainite ferrite. Because the silicon content was high, the steel had a microstructure of carbide-free bainite and a few proeutectoid ferrite due to high silicon content. And it consisted of carbide-free bainite (lath bainite ferrite and filminess retained austenite between laths) when cooled from 950 °C to room temperature at 0.3 °C/s, lath martensite and a little bainite were obtained when cooling rate increase to 8.0 °C/s. The addition of Si and Mn increased the amount of retained austenite with higher stability, and still, improved the strength and toughness. High cooling speed was the main reason to form a martensitic structure. With the increase of cooling rate, the microstructure shaped to needle-like lower bainite.

Table 2 Microstructures of the test alloy under different cooling rates

Test	Cooling rate (°C/s)	Microstructure
1	0.05	Bainite + proeutectoid ferrite
2	0.3	Carbide-free bainite
3	8.0	Lath martensite

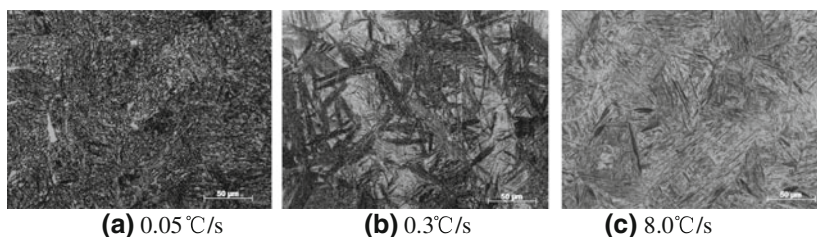


Fig. 2 Microstructures of the steel under different cooling rates

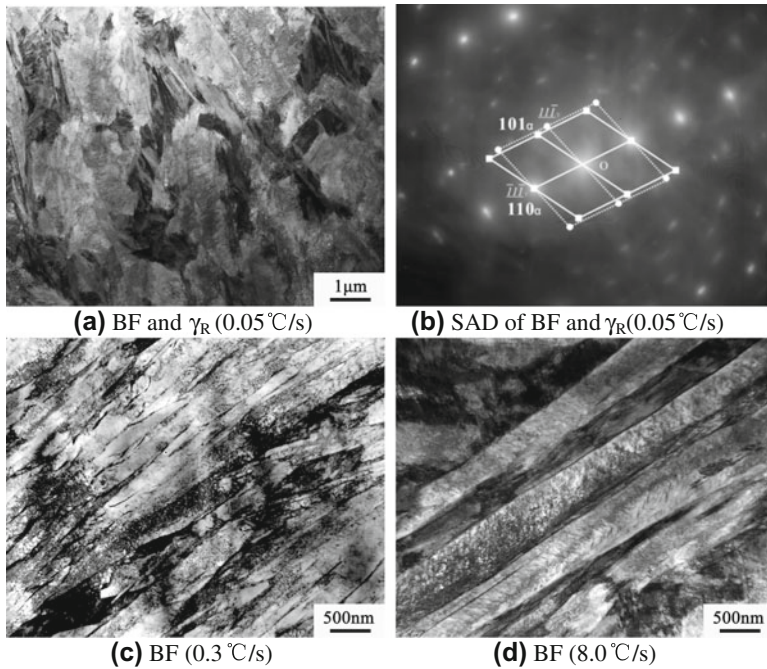


Fig. 3 TEM micrographs of the steel under different cooling rates

3.3 The Effects of Heat Treatment to the Mn–Cr–V–Ni Alloy

3.3.1 Mechanical Properties of the Bainitic Steel After Heat Treatment

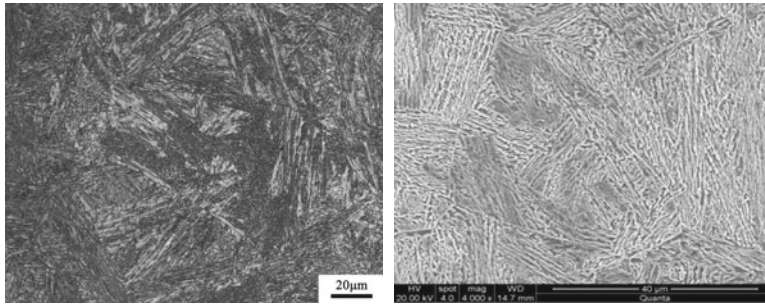
The mechanical properties of the bainitic steel after 920 °C annealing–normalizing–tempering heat treatment are shown in Table 3. For test 4, it can be seen that the alloy has tensile strength of 1353 MPa, yield strength of 1057 MPa, an elongation of 17%, a reduction of 52%, impact toughness of 110 J/cm², and 42.0 HRC. In general, under the condition of test 4, the alloy had a good combination of strength and toughness. It illustrates that annealing improves the elongation and reduction of area. With the addition of silicon, manganese, chromium, nickel, and vanadium, the alloy has high strength and toughness.

3.3.2 Microstructure of the Bainitic Steel After Different Heat Treatment

The microstructures of testing bainitic steel obtained by 920 °C annealing–normalizing–tempering heat treatment are shown in Fig. 4. The steel had a microstructure of lath bainite and a few proeutectoid ferrites. The alloy consisted of

Table 3 Mechanical properties of the test alloy after heat treatment

Test	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation (%)	Reduction of area (%)	Impact toughness (J/cm ²)	Rockwell hardness (HRC)
4	1353	1057	17	52	110	42.0

**(a)** metallographic structure **(b)** scanning electron microscope photograph**Fig. 4** Microstructure of the bainitic steel after heat treatment

parallel lath bainite ferrite and carbon-rich thin-film retained austenite. This was due to silicon, and silicon suppressed carbide precipitation of alloy. Then more and more carbon appeared in austenite, which reduced Martensite start (Ms) temperature under room temperature.

Combined the data of microstructure and mechanical properties, the steel will have good combination of strength, toughness, stabilization, and homogeneity of microstructure through 920 °C annealing–normalizing–tempering as designed.

The well-closed formation could be attributed to heat treatment and addition of alloy elements. The main goals of annealing are making microstructure and composition uniform, refining grain, eliminating internal stress and work hardening, etc. The temperatures annealing and normalizing are similar, but cooling rates are different: cooling rate of normalizing is fast. The strength and hardness are high after normalizing. Tempering helps in decreasing notch sensitivity of steel and obtaining a good combination of high strength and high toughness. And chromium can improve the temper stability and hardenability of bainitic steel, therefore the steel can increase and improve combination of high strength and toughness further.

4 Conclusion

A Mn–Cr–V–Ni forging bainitic steel for railway transportation has been developed. In this paper, the following conclusions have been obtained:

1. A Mn–Cr–V–Ni alloy was prepared, and its composition is: C 0.28–0.29, Mn 1.7–1.8, Si 1.5–1.8, Cr 0.8–1.0, Ni 0.2–0.4, Mo 0.3–0.5, V 0.08–0.12. The alloy was produced with secondary refining, vacuum degassing, blooming casting, billet slow cooling, etc.
2. The Mn–Cr–V–Ni alloy was cooled from 950 °C to room temperature at different cooling rates, microstructure of the alloy consisted of bainite and a small amount of proeutectoid ferrite under cooling rate of 0.05 °C/s, and it consisted of carbide-free bainite under cooling rate of 0.3 °C/s, filminess retained austenite existed discontinuous between bainite ferrite. The test alloy got lath martensite and a little bainite when cooling rate increase to 8.0 °C/s.
3. The Mn–Cr–V–Ni alloy was heat treated by 920 °C annealing–normalizing–tempering, a good combination of tensile strength of 1353 MPa, yield strength of 1057 MPa, an elongation of 17%, a reduction of 52%, impact toughness of 110 J/cm², and 42.0 HRC.

References

1. Esveld C (2001) Modern railway track, 2nd edn. MRT-Productions, Zaltbommel
2. Aglan HA, Liu ZY, Hassan MF, Fateh M (2004) Mechanical and fracture behavior of bainitic rail steel. *J Mater Process Technol* 151(1–3):268–274
3. Yates JK (1996) Innovation in rail steel. *J Sci Parliament* 53:2–3
4. Pacyna, J (2008) The microstructure and properties of the new bainitic rail steels. *J Achiev Mater Manuf Eng* 28(1):19–22
5. Chen X, Luan D, Sujuan SU, Wang H (2008) Study on microstructure and properties of bainitic steel for switch-point component. *Hot Working Technol* 37(2):25–27 (in Chinese)
6. Huang X, Liu W, Huang Y, Chen H, Huang W (2015) Effect of a quenching–long partitioning treatment on the microstructure and mechanical properties of a 0.2 C% bainitic steel. *J Mater Process Technol* 222:181–187
7. Kapito A, Stumpf W, Papo MJ (2013) The role of alloy-ing elements in bainitic rail steels. *J South African Inst Min Metall* 113(2):67–72
8. Cheng J, Liu Z (2011) Research and application of carbon free bainite wear resistant cast steel. *Foundry* 60(4):382–385 (in Chinese)
9. Bracke L, Xu W (2015) Effect of the Cr content and coiling temperature on the properties of hot rolled high strength lower bainitic steel. *ISIJ Int* 55(10):2206–2211

Design of and Research on an Attendance System Based on RFID and WSN Technologies for the Rail Transportation Industry

Chongjun Liu, Kuangang Fan, Yi Pan and Yindong Ren

Abstract To address the serious conflict and poor reliability issues in existing radio-frequency identification (RFID) systems, a complete set of remote wireless noncontact attendance system was designed using a 2.4 GHz active RFID system as the terminal reading device for recording employee attendance information. Combined with a wireless network architecture provided by a wireless sensor network, the system wirelessly transmits attendance information to a remote computer control system, thereby effectively improving the cumbersome wiring, poor mobility, and expansibility faults of the cable-contacted attendance system used in the Rail Transportation industry. Consequently, production safety and personnel attendance management become more convenient. The client management system communicates using LabVIEW development tools with the SQL SERVER 2008 database to achieve attendance management functions. To address the electronic tag collision problem in an active RFID system, an improved framed-slotted ALOHA anti-collision algorithm was proposed by combining the original algorithm with dynamic grouping ideas, thereby improving RFID system performance and efficiency.

Keywords Rail transportation attendance system · RFID · WSN
LabVIEW · Anti-collision algorithm

C. Liu (✉)

System & Electronic Control Unit(NE1), United Automotive
Electronic Systems Co., Ltd, Shanghai, China
e-mail: kuangangfriend@163.com

K. Fan · Y. Pan · Y. Ren

School of Mechanical and Electrical Engineering, Jiangxi University
of Science and Technology, Ganzhou, Jiangxi, China

© Springer Nature Singapore Pte Ltd. 2018

L. Jia et al. (eds.), *Proceedings of the 3rd International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2017*, Lecture Notes in Electrical Engineering 483, https://doi.org/10.1007/978-981-10-7989-4_103

1013

1 Introduction

To address the shortcomings of the attendance system used for employees working in the Rail Transportation industry, such as being error prone [1, 2], exhibiting poor anti-interference capability [3, 4], and using an inconvenient storage system [5, 6], this study mainly introduces an attendance system that combines radio-frequency identification (RFID) and wireless sensor network (WSN) technologies to realize a long-distance and noncontact attendance management system. Managers can check the attendance of operators working in real time, thereby improving production safety and management in the Rail Transportation industry and making the attendance management of operators convenient and scientific.

2 Design and Analysis

The slave computer used in the proposed system is the ARM chip and its peripheral device. Each RFID node connects with each monitoring point to form an information control platform through WSNs. The host computer is an information management system for employee attendance in the Rail Transportation industry. This system is used to manage attendance data inquiries and report forms. The entire system involves wireless communication RFID, database creation, and data control system. The overall diagram of the system is presented in Fig. 1.

A fieldbus system is used for energy transmission; it is connected with the attendance system through a WSN. A reader for rail transportation personnel exchanges background data with the database used for attendance monitoring in the main system to realize attendance management of personnel. The design of the network is shown in Fig. 2.

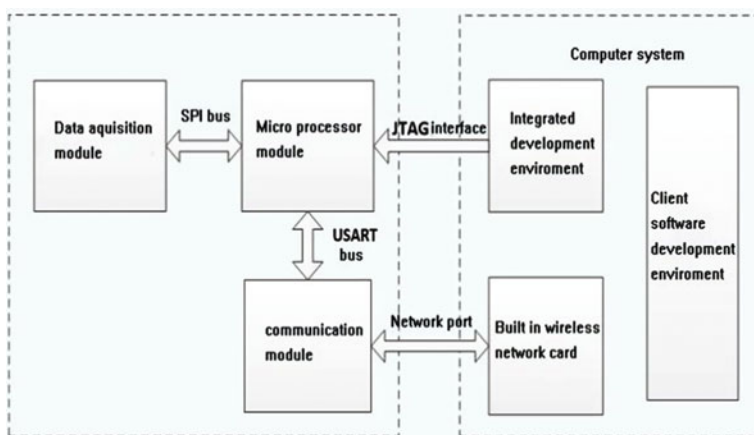


Fig. 1 Overall block diagram of the system

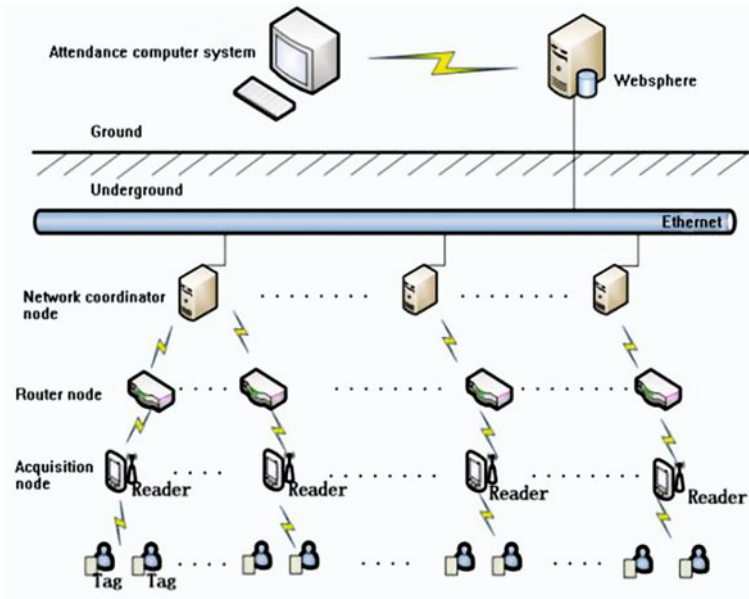


Fig. 2 Overall network structure of the system

3 The Design of Software Model

The slave computer, i.e., LabVIEW, is connected with the database; it indexes personnel information from the host computer to realize attendance function. The diagram of the system software modules is provided in Fig. 3.

The main part of the software design is the LabVIEW client, which communicates with the ARM chip and the database to realize certain attendance functions. The database is established using the database management system software SQL SERVER 2008. The E-R diagram of the database is shown in Fig. 4.

LabVIEW and the database use LabSQL module to achieve interconnected communication.

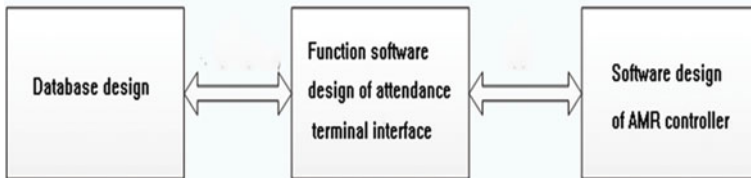


Fig. 3 Block diagram of system software module

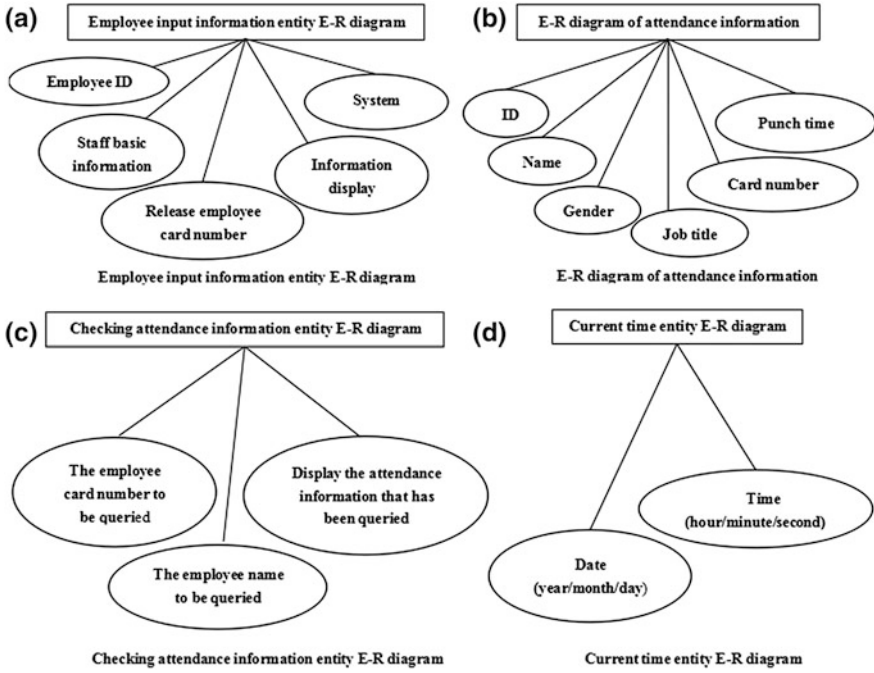


Fig. 4 Database entity E-R diagram

4 Improvement of the ALOHA Algorithm

While the RFID system is working, the integrity of anti-collision algorithm determines the system performance largely [3, 4]. When the frame length is L and the label is N , the system reads the correct slot rate, s shown in Formula (1):

$$P_l = \frac{n}{L} \left(1 - \frac{1}{L}\right)^{n-1} \tag{1}$$

The idle slot rate is shown in Formula (2):

$$P_0 = \left(1 - \frac{1}{L}\right)^n \tag{2}$$

The collision slot rate is shown in Formula (3):

$$P_k = 1 - P_0 - P_l = 1 - \left(1 - \frac{1}{L}\right)^n - \frac{n}{L} \left(1 - \frac{1}{L}\right)^{n-1} \tag{3}$$

From the preceding mathematical models, the simulation of the framed-slotted ALOHA algorithm is illustrated in Fig. 5.

To address the shortcoming of the framed-slotted ALOHA algorithm, a dynamic frame-slotted ALOHA (DFSA) algorithm is proposed to improve the efficiency of label recognition in the framed-slotted ALOHA algorithm [6], which effectively enhances system efficiency [7]. Moreover, the DFSA algorithm should estimate the number of labels during the reading process, and several ideas for estimating labels [8], such as estimating the total number or the number of unrecognized labels, are presented [9–11].

The DFSA algorithm dynamically adjusts frame length according to the current idle and collision rates to improve efficiency [12, 13]. Compared with the fixed framed-slotted ALOHA algorithm, the proposed algorithm exhibits higher recognition rate and stability [14]. The main objective of the improved algorithm is to address the poor recognition performance of the framed-slotted ALOHA algorithm when the number of labels exceeds 256. A grouping framed-slotted ALOHA (GFSA) anti-collision algorithm, which plays an important role when the frame length is 1.7 times the number of labels, is proposed. On the basis of this conclusion, the improved packet DFSA algorithm is more convenient for determining the optimal frame length and adjusting it in real time [15]. The concept of the improved GFSA algorithm is illustrated in Fig. 6, and the mathematical model for the time-GFSA algorithm simulation is shown in Fig. 7. The throughput rate increases as frame length increases. Meanwhile, the recognition rate gradually increases as the idle rate declines. The improved DFSA algorithm performs best when F reaches 256.

The relation between the number of slots and the number of labels in the improved DFSA algorithm is shown in Fig. 8. When the frame length is longer than

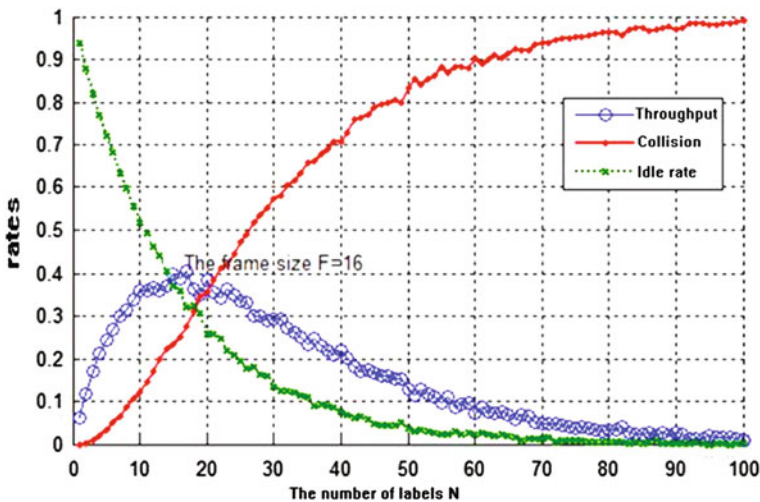


Fig. 5 Throughput, collision rate, idle rate, and label number of frame time slot ALOHA

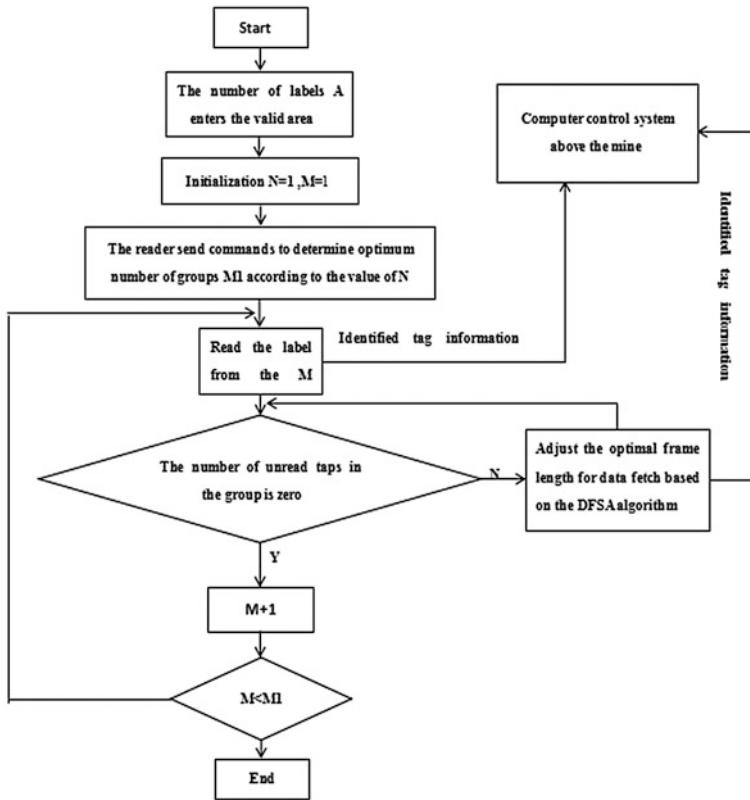


Fig. 6 The improved idea of DFSA algorithm

Fig. 7 The relationship of throughput rate, collision rate idle rate, and the number of tags in frame time slot time-GFSA

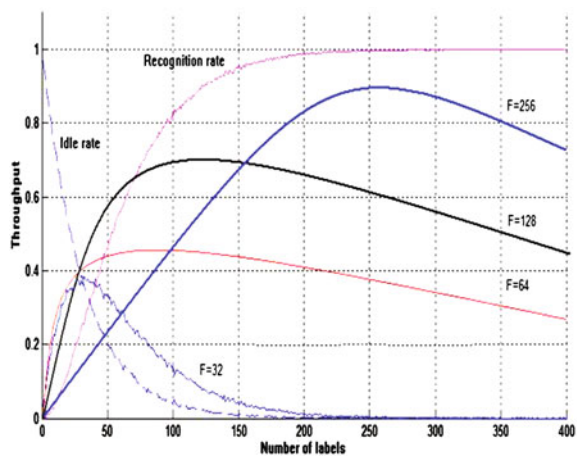
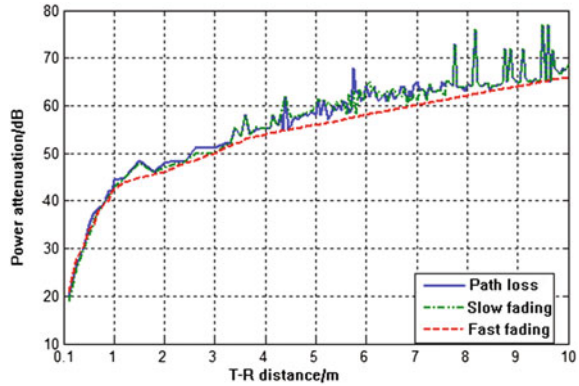


Fig. 8 The relation between the number of slots and the number of tags in the improved DFSA



256, performance sharply deteriorates. When the number of labels reaches 500, the number of slots sharply increases. The number of readable labels is increased by at least 9.36%.

5 Performance Analysis of RFID System

The RFID signal will experience loses during the channel transmission process due to various factors. The receiving power formula for the power transmission signal is given in Formula (4)

$$P(d) = d^{-n} \cdot S(d) \cdot R(d) \tag{4}$$

In this formula, “*d*” is the distance between the sending and receiving ends.

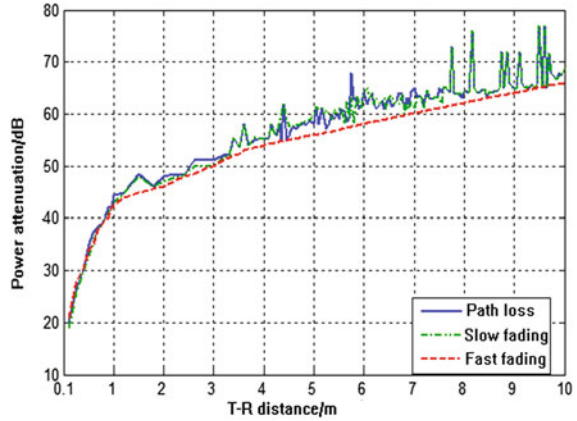
Figure 9 shows the path loss, slow decline, and rapid decline of the RFID signal passing through the RFID channel. In the simulation, we set the reference distance *d*₀ as 1 m and the other parameter settings are as follows: frequency = 2.4 GHz, *n* = 2.4, and *X*_σ = 9.6 dB. The distance in the figure is represented by a logarithm (which corresponds to UHF RFID systems).

We assume that the binary phase-shift keying (BPSK) signal, which applies BPSK modulation, is similar to Formula (5).

$$s_{\text{BPSK}}(t) = A \left[\sum_{n=-\infty}^{\infty} a_n g_T(t - nT_b) \right] \cos \omega_c t \tag{5}$$

In this formula, {*a_n*} is a binary number sequence and the two electrical levels exhibit the same probability to appear. *T_b* is the binary symbol interval, and *g_T*(*t*) is impulse response of Baseband Transmitting Forming Filter, and we assume that *g_T*(*t*) has a nonzero rectangular shape.

Fig. 9 The relationship of path loss, slow decline, and fast decline in power transmission



When a BPSK signal is demodulated using a matching filter, we sample the matched filter output under the condition of sending +1 and $t = T_b$. The sampling value is given in Formula (6).

$$y(T_b) = \int_0^{T_b} [s_1(t) + n_w(t)]s_1(t)d(t) = E_b + Z \tag{6}$$

In Formula (6), $Z = \int_0^{T_b} n_w(t)s_1(t)d(t)$

When sending -1, the sampling value is given in Formula (7)

$$y(T_b) = \int_0^{T_b} [s_1(t) + n_w(t)]s_1(t)d(t) = -E_b + Z \tag{7}$$

We assume that $y = y(T_b)$, and its sampling value is given in Formula (8–9).

$$p(y|s_1) = \frac{1}{\sqrt{\pi N_0 E_0}} \exp \left[-\frac{(y - E_b)^2}{N_0 E_b} \right] \tag{8}$$

$$p(y|s_1) = \frac{1}{\sqrt{\pi N_0 E_0}} \exp \left[-\frac{(y + E_b)^2}{N_0 E_b} \right] \tag{9}$$

When $P(s_1) = P(s_2) = \frac{1}{2}$, the average bit error ratio (BER) A is given in Formula (10)

$$P_b = P(e|s_1) = P(e|s_2) \tag{10}$$

In this formula, $P(e|s_1) \int_{-\infty}^0 P(y|s_1) dy = \frac{1}{2} \operatorname{erfc} \left[\sqrt{\frac{E_b}{N_0}} \right]$

Thus, the optimal obtained BER is given in Formula (11).

$$P_b = \frac{1}{2} \operatorname{erfc} \left[\sqrt{\frac{E_b}{N_0}} \right] = Q \left[\sqrt{\frac{2E_b}{N_0}} \right] \quad (11)$$

The wall transmission environment or other obstacles in radio obscuration will lead to a decline in the accepted signal intensity of path loss d^{-n} , which varies with distance in a large-scale range and with the slow change in the median of the signal level in medium-range slow fading $S(d)$. Multipath scattering produces fast fading, in which the instantaneous value of the signal field strength varies rapidly within a small range. Under the same signal-to-noise ratio (SNR), the BER performance of the signal in the additive Gaussian white noise channel is better than that in the multipath channel. When the SNR is 14 dB, the additive Gaussian white noise channel has a BER of less than 10^{-3} , and the error rate of the multipath Rayleigh fading channel is approximately 5%. To obtain a transmission effect that is equivalent to the additive Gaussian white noise channel, SNR should be increased.

6 System Experimental Test

When the attendance system for the Rail Transportation industry is opened, the corresponding log-in interface will be displayed. The main interface, including the current time, welcome remarks, and description design, along with the three function modules, namely, staff information entry, real-time attendance table, and attendance query, will pop up after successful landing is achieved. First, in the staff information entry module, the basic information, entry post, and card number of a new employee should be inputted into the left column. When data entry is successful, the aforementioned information, together with the recording time, will be displayed in the right column. Second, the real-time attendance table module can be used to display the real-time information of workers under the current credit card and stored in the database for managers to query. Third, in the attendance query module, the employee name or card number to be queried is selected through the query option button, and the corresponding query information is entered in the query box, the staff attendance situation will then be presented.

Various algorithms are used to perform a comprehensive test on the RFID attendance system, with multiple interference sources, to verify the corresponding work situation. The results are provided in Table 1.

Table 1 Comparison of error rate in system operating state

Test times	Pure ALOHA		Framed-slotted ALOHA		DFSA		Improved	
	Interference (error count)	Normal condition (error count)	Interference (error count)	Zero interaction (error count)	Interference (error count)	Zero interaction (error count)	Interference (error count)	Zero interaction (error count)
100	91	65	61	52	22	11	6	2
200	180	132	128	100	38	20	10	3
500	467	342	318	263	118	49	27	8
1000	908	656	605	523	216	112	59	21

7 Conclusion

A long-distance attendance system is designed after considering the problems in existing attendance systems. The RFID technique combines WSNs and the features of an RFID signal to identify a target while simultaneously realizing wireless communication. The active perception and communication functions can realize no-contact, long-distance attendance monitoring and record attendance data in real time. The host computer of the client and the terminal reader communicate data transmission through WSNs, which is convenient for gathering statistics and inquiring about the attendance data of employees working in the Rail Transportation industry, thereby enabling an enterprise to improve accuracy and efficiency in the production safety management of employees. An improved DFSA anti-collision algorithm that is suitable for the attendance system is proposed based on the time during which a staff member enters the scan range. In addition, a mathematical model image is drawn through Matlab to compare the advantages and disadvantages of various algorithms. Subsequently, the improved DFSA algorithm presents the improved method and programming concept of the algorithm, and then verifies the applicability and feasibility of the system. Ultimately, the number of effective slots read by the system is increased, i.e., it is 9.36% higher than the previous value, thereby proving that the system exhibits a high degree of determination and conflict prevention functions. In addition, the system is convenient for gathering statistics and inquiring about the attendance data of employees who are working in the Rail Transportation industry. Accordingly, an enterprise can better manage the production safety of its employees.

Acknowledgements This work is supported by the Support Program of China National Science and Technology (No. 61763018, No. 2006BAK03A21, No. 2009BAK58B03) and Jiangxi Provincial Department of Science and Technology(No. 20171BAB206030) and the Foundation of Education Committee of Jiangxi(GJJ150674) and the Key Program of Nature Foundation of Jiangxi University of Science and Technology(NSFJ2015-K09).

References

1. Yun P (2014) Research and design of remote attendance system based on RFID. Jiangsu University of Science and Technology, Zhenjiang (in Chinese)
2. Jianhong Li (2013) Design and implementation of 2.4 GHz large capacity active RFID tag system based on Cortex_M3. Chengdu: University of Electronic Science and Technology (in Chinese)
3. Occhiuzzi C, Marrocco G (2016) Precision and accuracy in UHF-RFID power measurements for passive sensing. *IEEE Sens J* 3901–3908
4. Zhao B, He J, Huang N, Qu H, Liu G (2014) RFID Anti—collision algorithm based on Flat N—tree search. *J Beijing Univ Posts Telecommun* 37:50–55 (in Chinese)
5. Wang X, Qian Z, Liu X, Cheng C (2015) Improved anti—collision algorithm for tree structure. *J Commun* 36:129–137 (in Chinese)

6. Cao G, Lu J, Wang X (2011) Research on Anti—information collision of dynamic frame time slot ALOHA algorithm. *J Xi'an Univ Post Telecommun* 23–25 (in Chinese)
7. Cui Y, Zhao Y (2009) A modified Q-parameter anti-collision scheme for RFID systems. 2009 International conference on ultra-modern telecommunications & workshops. 09:197–203 (in Chinese)
8. Wang B, Zhang Q (2009) Research on Anti—collision performance of UHF RFID system with parallel recognition. *J Commun* 30:108–113 (in Chinese)
9. Ferdous RM, Reza AW, Siddiqui MF (2016) Renewable energy harvesting for wireless sensors using passive RFID tag technology: A review. *J Renew Sustain Energy Rev* 58:1114–1128
10. Tang L, Zheng L, Cao H, Huang N (2016) An improved multi-objective genetic algorithm for heterogeneous coverage RFID network planning. *Int J Prod Res* 54(8):2227–2240
11. Sunny AI, Yun T, Zhang J (2016) Low frequency (LF)RFID sensors and selective transient feature extraction for corrosion characterization. *Sens Actuators A-Phys* 241:34–43
12. Guo Z, Linlin C, Zhou Y, Gu J (2012) Improvement of dynamic frame time slot ALOHA algorithm. *Comput Appl Res* 907–909
13. Yu J (2010) An improved dynamic frame time slot aloha algorithm. *Data Commun* 04:25–27
14. Guidi F, Decarli N, Dardari D (2016) A low complexity scheme for passive UWB-RFID: proof of concept. *IEEE Commun Lett* 20:676–679
15. Shan J, Xie J, Zhuang Q (2011) Research on ALOHA Anti—collision algorithm based on grouping dynamic frame timeslot. 39–45