Lecture Notes in Electrical Engineering 288

Limin Jia
Zhigang Liu
Yong Qin
Minghua Zhao
Lijun Diao *Editors*

Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II

Springer

Lecture Notes in Electrical Engineering

Volume 288

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadyay, Palmerston, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Sakyo-ku, Japan
Federica Pascucci, Roma, Italy
Tariq Samad, Minneapolis, USA
Gan Woon Seng, Nanyang Avenue, Singapore
Germano Veiga, Porto, Portugal
Junjie James Zhang, Charlotte, USA

For further volumes:
http://www.springer.com/series/7818

About this Series

"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

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical 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.

Limin Jia · Zhigang Liu · Yong Qin
Minghua Zhao · Lijun Diao
Editors

Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II

Springer

Editors
Limin Jia
Zhigang Liu
Yong Qin
Lijun Diao
Beijing Jiaotong University
Beijing
People's Republic of China

Minghua Zhao
China CNR Changchun Railway
Vehicles Co., Ltd
Changchun
People's Republic of China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-3-642-53750-9 ISBN 978-3-642-53751-6 (eBook)
DOI 10.1007/978-3-642-53751-6
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013957384

© Springer-Verlag Berlin Heidelberg 2014
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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
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.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

EITRT Committees

Honorary Chairs

Shi Zhongheng, Academician of China Academy of Engineering, China
Satoru Sone, Tokyo University, Japan
Buchheit Karlheinz, Siemens, Germany
Qian Qingquan, Southwest Jiaotong University, China
Terasawa Kiyoshi, Hitachi, Japan

General Chair

Jia Limin, Beijing Jiaotong University, China

Program Committee Chairs

Liu Zhigang, Beijing Jiaotong University, China
Li Zhong, University of Hagen, Germany
Li Yaohua, Institute of Electrical Engineering, Chinese Academy of Sciences, China
Holger Hirsch, University of Duisburg-Essen, UK

Organizing Committee Chairs

Zhao Minghua, China CNR Changchun Railway Vehicles Co., Ltd, China
Min An, University of Birmingham, UK
Wolfgang A. Halang, University of Hagen, Germany
Qin Yong, Beijing Jiaotong University, China
Kyandoghere Kyamakya, University of Klagenfurt, UK

Organizing Committee Members

Zhao Minghua, China CNR Changchun Railway Vehicles Co., Ltd, China
Wolfgang A. Halang, University of Hagen, Germany
Qin Yong, Beijing Jiaotong University, China
Min An, University of Birmingham, UK
Wang Shubin, China CNR Changchun Railway Vehicles Co., Ltd, China
Xu Chunmei, Beijing Jiaotong University, China
Diao Lijun, Beijing Jiaotong University, China
Cheng Xiaoqing, Beijing Jiaotong University, China

Technical Program Committee Members

Buchheit Karlheinz, Experts of Siemens, Germany
Cai Changjun, Guangzhou Metro Corporation, China
Chai Jianyun, Tsinghua University, China
Chang Zhenchen, China CNR Changchun Railway Vehicles Co., Ltd, China
Clave Roberts, University of Birmingham, UK
Li Zhong, University of Hagen, Germany
Holger Hirsch, University of Duisburg-Essen, Germany
Fang Youtong, Zhejiang University, China
Feng Jianghua, China CSR Zhuzhou Electric Locomotive Research Institute, China
Gao Shibin, Southwest Jiaotong University, China
Gong Ming, China CSR Sifang Co., Ltd, China
Kyandoghere Kyamakya, University of Klagenfurt, Germany
Jia Limin, Beijing Jiaotong University, China
Jiang Jiuchun, Beijing Jiaotong University, China
Ye Jianqiao, Mechanical Engineering Department of Engineering, Lancaster University, UK
Li Yaohua, Institute of Electrical Engineering, Chinese Academy of Sciences, China
Liu Baoming, China CNR Qingdao Sifang Locomotive & Rolling Stock Research Institute, China
Liu Zhigang, Beijing Jiaotong University, China
Mark Hooper, Faculty of Engineering and Computing, Coventry University, UK
Qin Yong, Beijing Jiaotong University, China
Chen Rui, Loughborough University, UK
Satoru Sone, Tokyo University, Japan
Wang Simon, School of Aeronautical and Automotive Engineering, Loughborough University, UK
Ling Tungchai, University of Birmingham, UK

Wang Litian, China Railway Electrification Survey and Design Institute Co., Ltd, China
Wolfgang A. Halang, University of Hagen, Germany
Yang Zhongpin, Beijing Jiaotong University, China
Zhao Minghua, China CNR Changchun Railway Vehicles Co., Ltd, China

Contents

Part I Electrical Traction and Power Supply of Rail Transportation

1 Application of Affinity Propagation Clustering Algorithm in Fault Diagnosis of Metro Vehicle Auxiliary Inverter 3
Junwei Gao, Zengtao Ma, Yong Qin, Limin Jia and Dechen Yao

2 A Soft-Switching Control Method of Isolated LC Series Resonant Transformer Full Bridge DC–DC Converter. 11
Meng Jiang and Wei Li

3 Study on Efficiency Optimization of Medium Frequency Transformers for Rail Transit Traction System. 21
Juanjuan Zhang, Yumei Du, Zixin Li, Ping Wang and Yaohua Li

4 Primary Permanent Magnet Linear Motors for Rail Transit 31
Ruiwu Cao and Ming Cheng

5 An Maximal Clique Mining Algorithm for Highway Network Optimization Problem 41
Zipeng Zhang, HongGuo Wang, Yanhui Ding and Zengzhen Shao

6 Design of a 6 kW Battery Charger Based on Full-Bridge Phase-Shifted ZVZCS PWM Converter. 51
Tao Sun, Zhigang Liu, Chao Li, Lei Wang and Lijun Diao

7 Diagnostics of Transformer Windings Deformation Based on Transfer Function 65
Lei Yin, Zhensheng Wu and Junfeng Gui

8 Clamped Three-level Inverter Midpoint Potential Control Method 73
Haijie Jia, Xuedong Jiang, Lei Wang, Zheming Jin and Qifeng Li

9 Design and Research for the Low-Loss Passive Snubber Circuit Used in the Metro High Frequency Auxiliary Converter 83
Yao Fu, Lijun Diao, Huiqing Du, Haijie Jia and Lei Wang

10 Research on Parallel Characteristics of Lithium Iron Phosphate Batteries for Dual Electric Multiple Units 95
Yanru Zhang, Jiuchun Jiang, Weige Zhang, Wei Shi, Zeyu Ma and Fangdan Zheng

11 Discussion on the Application of Energy-Saving Traction Power Supply Device 105
Zhiwei Han, Gang Zhang, Zhigang Liu and Fuqiang Mu

12 Research on Applicability of Lithium Titanate Battery for Low-Floor Vehicles 115
Sijia Liu, Jiuchun Jiang, Zhanguo Wang, Zeyu Ma, Wen Chen and Ting Zhao

13 Research of the Lithium Battery-Based Energy Storage System for Light Rail Vehicle 127
Xiaoguang Jia, Zhigang Liu, Wenhui He, Zheming Jin and Lijun Diao

14 Research of Fault Location Method for Metro Traction Power Supply System 137
Xiaoming Song, Jinghan He, Tony Yip, Brian Kirby and Xia Yang

15 Study and Implementation of Closed Loop Control Based on Double Synchronous Rotating Frame for EMU Auxiliary Inverter with Unbalanced Load 147
Huiqing Du, Zhigang Liu, Yao Fu, Lei Wang and Lijun Diao

16 Fluid and Thermal Analysis of Power Li-Ion Battery Pack and Experimental Verification 161
Dafen Chen, Jiuchun Jiang, Yaojuan Duan, Zhanguo Wang and Feng Wen

17 A Novel Real-Time Tension Monitoring Method for Overhead Contact System 171
Xin Meng, Jinghan He, Xiaojun Wang and Guomin Luo

18 Locomotive DC600 V Power Supply Design and Control Redundancy Improvement 179
Yan Xia, Chunmei Xu and Lei Wang

19 The Research on BEV-Based Urban Passenger Transport Environment 189
ZhiQiang Rao, ZiJian Zhuang and Bin Ning

20 Design and Verification of Hybrid Power Box of 100 % Low-Floor LRV 199
Li Ming, Li Minggao, Shi Junjie, Yang Guang and Shao Nan

21 Measurement and Simulation of the Electromagnetic Transients of Lifting Pantograph for an Electric Multiple Units Train 213
Dan Shi, Mingli Wu, Honghe Zhang, Teng Li, Hui Wang and Kejian Song

22 Calculations of Leakage Impedance of Rail to Earth in Ballastless Track by Finite Element Method 223
Teng Li, Mingli Wu, Fan He and Kejian Song

Part II System Safety and Emergency Management

23 Research on Transport Capacity of Urban Rail Transit Based on RailSys 235
Yanjie Wang and Xi Zhang

24 Contribution Rates Calculation for Regulations About Urban Rail Vehicle Inspection and Repair Based on AMSAA Model 243
Jianwei Li, Minzheng Yuan, Yong Qin, Limin Jia and Zongyi Xing

25 Failure Mode Criticality Analysis of Metro Door System 251
Jun Xia, Li sha Pan, Xiao qing Cheng, Yong Qin and Zong yi Xing

26 Reliability Analysis of the Sliding Plug Door System Based on Bayesian Network 259
Lingli Mao, Zhaoyi Su, Jing Long, Limin Jia and Zongyi Xing

27 The Prediction of Derailment Coefficient Based on Neural Networks 267
Xiulian Yu, Guangwu Liu, Yong Qin, Yuan Zhang and Zongyi Xing

28 Research on Hazard Evaluation of Urban Rail Train Based on the Extension Theory 275
Guiling Liao, Yong Qin, Yuan Zhang, Xiaoqing Cheng and Zongyi Xing

29 Reliability Analysis of Metro Door System Based on Fuzzy Reasoning Petri Net 283
Ping Liu, Xiaoqing Cheng, Yong Qin, Yuan Zhang and Zongyi Xing

30 Fault Criticality Evaluation of Metro Door Based on WLSM and FWGM 293
Jin Bao Ren, Jing Long, Yong Qin, Zong Yi Xing and Jun Xia

31 Risk-Based Maintenance Optimization of Metro Vehicle Door System 301
Xiang Guo, Lisha Pan, Limin Jia, Yuan Zhang and Zongyi Xing

32 Using Asynchronous Hot Standby Spare in Time-Stamped, Fault-Tolerant, Real-Time System 309
Mahmoud Jannesari Ladani and Ahmad Kezemi Gazanchaei

33 Formal Modeling and Analysis of Radio Block Center Handover 313
Xiaohui Hu, Jun Ma, Yong Chen and Xin Li

34 Fault Diagnosis for Rail Vehicle Suspension Systems Based on Fisher Discriminant Analysis 321
Xiukun Wei, Sheng Wu, Jianlong Ding, Limin Jia, Qu Sun and Minzhen Yuan

35 Analysis of Related Factors Influencing Reliability of Railway Signaling Systems Based on Fuzzy Analytical Hierarchy Process 333
Hongxia Chen and Min An

36 Application of GA-LSSVM in Fault Diagnosis of Subway Auxiliary Inverter 341
Junwei Gao, Ziwen Leng, Yong Qin, Limin Jia and Dechen Yao

37 The City Rail Safety Detection Sensor Network Platform Using Online Track Detection 349
Shaohuang Pang, Weiping Zhao, Honghui Dong, Wenlong Peng and Jianxiao Chen

Part III Rail Transportation Computer Technology

38 Research for Algorithm of the Super Low Delay Image Coding in Airborne Photo-Electricity Survey Equipment 357
Qihu Li, Yunfeng Wen and Ruicai Jia

39 A Novel Recursive Algorithm for Training RBF Networks 365
Peng Zhou and Zhu Yang

40 Argumentation Pattern: An Approach to Issuing Software Reliability Case 373
Boxuan Wang and Minyan Lu

41 Control Modeling and Signal Processing of a Library Self-Delivery Robot and Its Applications 383
Shuping Dang, Meizun Meng, Daniel Mathews and Rinat Kakimzhanov

42 Analysis of Moore's Law on Intel Processors 391
Jia Chen

43 A Reliable QoE-aware Framework for Cloud Service Monitoring and Ranking 401
Yuchao Zhang, Hongfu Liu, Bo Deng and Fuyang Peng

44 Learning Distance Metrics with Feature Space Performance for Image Retrieval 411
Xin Luo, Guowen Wu and Kenji Kita

45 A Fine-Grained Authentication Model Based on Perceptual Hashing and Grid Descriptor for Remote Sensing Image 423
Kaimeng Ding and Yuhai Wang

46 Adaptive Tuning Algorithm Used in Multi-Join Query Optimization 431
Zhou Jiang, Lianzhong Liu and Zheren Li

47 Dynamic Energy-Efficient Virtual Machine Placement Optimization for Virtualized Clouds 439
Xiaoqing Zhang, Qiang Yue and Zhongtang He

48 The Research on Distortion Correction Algorithm of Unwrapping the Cylinder Image of Panoramic Annular Lens 449
Qiuxiang Liao

49 A Novel Control Strategy of Permanent Magnet Synchronous Machine Drive Under Field-Weakening Operation 457
Xiangyang Wu, Kekang Wei and Shihua Bi

Part IV Rail Transportation and Automatic Control

50 Research on the Balise Uplink Signal Process Method Based on the Noise Feature Extraction and Adaptive Noise Cancellation 469
Yanning Zhang and Huibing Zhao

51 Research on the Relation Between the Wind Direction of Vehicle Air Condition and Cooling Efficiency of the Cab 479
Bin feng Guo, Hong chao Wang and Yan ling Wang

52 Vocational Skills Comprehensive Evaluation Method of Track Maintenance Workers 487
Dongyao Jia and Shengxiong Zou

53 Middleware-Based Distributed CTCS-3 Simulation Platform 497
Lianbao Yang, Tianhua Xu and Zhenxian Wang

54 Study on Express-Slow Routes Mode of Foreign Subway 507
Senhui Li, Xingchen Zhang, Junhua Chen and Jixuan Liu

55 Optimized Design of Urban Rail Vehicle Grounding System 517
Ling Gao, Ruichang Qiu, Lei Wang and Haijie Jia

56 Analysis of the Property of Heavy Haul Railway's Traffic Flow Based on Hybrid Cellular Automaton 525
Wentan Deng and Huibing Zhao

57 Research on AC Drive Test System Based on Vector Control 537
Zhe Zhang, Zhi-gang Liu, Lei-ting Zhao and Li-jun Diao

58 Process Design and Analysis of Emergency Decision Support System for High-Speed Rail Transport Organization 547
Guoxing Han, Yong Qin, Tao Zhu, Jie Xu, Fei Dou and Kai Yan

59 Subway Vehicle Bearing Fault Diagnosis Methodology Research Based on PNN Neural Network and Wavelet Package 555
Xi Li, Yuan Zhang and Limin Jia

60 Study on the Best Transfer Station Selected in Beijing Subway Network 565
Qi Liu, Xingchen Zhang, Junhua Chen and Bin Xu

61 A New Approach to Diagnose Rolling Bearing Faults Based on AFD 573
Yu Liang, Li min Jia, Guo qiang Cai and Jin zhao Liu

62 Train Control Management System Safety Assessment 583
Changyuan Liu, Xiaoming Li and Panpan Yang

63 Research on Disruption Management of Single Machine Scheduling 593
Lei Xiao

64 A Study on Parking Problems and Countermeasures of Urban Central Commercial District 601
Xianyong Miao and Manzhen Duan

65 Research on Structural Modeling Technique of Vehicles Outward Based on AVI 609
Bing Wang, Dudu Guo, Wenlei Sun and Gulbahar Tohti

Part I
Electrical Traction and Power Supply of Rail Transportation

Chapter 1
Application of Affinity Propagation Clustering Algorithm in Fault Diagnosis of Metro Vehicle Auxiliary Inverter

Junwei Gao, Zengtao Ma, Yong Qin, Limin Jia and Dechen Yao

Abstract Auxiliary inverter is one of the most important electrical equipments of metro vehicle; its complex structure causes various faults frequently. In this paper, the fundamental of the Affinity Propagation algorithm is introduced, and its application on fault diagnosis of metro vehicle auxiliary inverter is studied. Fault signals including voltage frequency variation, pulse transient, and power interruption are simulated by using the MATLAB software; clustering center matrix is calculated on the basis of AP algorithm, and the fault samples are classified by calculating the similarity degree between samples and clustering center. The simulation results show that the AP algorithm without initial clustering center can be used in the field of fault diagnosis, and even has better results than FCM algorithm.

Keywords Affinity propagation · Fault diagnosis · Auxiliary inverter · Clustering analysis

1.1 Introduction

Auxiliary inverter is one of the most important electrical equipment of metro vehicle, which provides electrical power for all kinds of AC loads and DC loads. Because of its complex structure, it is the failure-prone equipment on the metro vehicle. Therefore, auxiliary inverter fault diagnosis is of important theoretical and

J. Gao (✉) · Z. Ma
College of Automation Engineering, Qingdao University, No 308 Ningxia Road, Qingdao 266071, China
e-mail: qdgao163@163.com

J. Gao · Y. Qin · L. Jia · D. Yao
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_1,
© Springer-Verlag Berlin Heidelberg 2014

practical significance. In recent years, with the rapid development of fault diagnosis technology, numerous intelligent fault diagnosis algorithms have appeared such as neural network [1, 2], fuzzy clustering [3–5], support vector machine [6], and so on. Good results can be achieved by using these intelligent diagnosis algorithms in the field of fault diagnosis. Clustering analysis theory is a method that divides samples into groups according to the similarity degree between two samples. The samples in the same group are similar while the samples in different groups are different from each other. Clustering analysis theory is widely used in the aspects of data management, data mining, image processing, as well as fault diagnosis. The fuzzy C-means clustering algorithm is applied to fault diagnosis in Ref. [7]. Reference [8] puts the power transformer as the study object and proposes a kind of fault classification which uses the kernel-based possibilistic clustering algorithm. In this paper, the authors apply the Affinity Propagation algorithm to auxiliary inverter fault diagnosis.

1.2 Affinity Propagation Algorithm

Affinity Propagation (AP) algorithm is a new clustering method proposed by Frey and Dueck [9] of Toronto University on *Science* in 2007. It aims to search for the set of clustering center in order to obtain the maximum sum of the similarity degree between samples and nearest clustering center [10].

1.2.1 Fundamental of AP Algorithm

All data points are considered as potential exemplars simultaneously in AP algorithm, for this reason the AP algorithm is suitable for a large number of data points clustering analysis to choose the exemplars to represent the whole class of them. In this paper, the authors choose the Euclidean distance to represent the similarity degree of any two data points x_i and x_j, denoted by similarity matrix $S(i,j)$ calculated as Eq. (1.1). The real number $S(k,k)$, denoted by preference $p(k)$, shows the possibility of data point k to be chosen as an exemplar. In the initialization conditions, the preference $p(k)$ should be set to a common value which means that each data point k has the same possibility to be an exemplar. In this paper, the authors set the median of the similarity matrix $S(i,j)$ as the value of $p(k)$.

$$s(i,j) = \begin{cases} -\left\|x_i - x_j\right\|^2 & i \neq j \\ p(i) & i = j \end{cases} \tag{1.1}$$

The core of AP algorithm is updating the responsibility matrix R and the availability matrix A. It can also be called the message exchanged between data points. There are two kinds of messages exchanged between data points. The

responsibility matrix $R = [r(i,j)]$, send from data point x_i to data points x_j, shows the responsibility of data point x_j to be an exemplar for data point x_i. The availability matrix $A = [a(i,j)]$, send from data point x_j to data point x_i, shows the suitability of data point x_i to choose data point x_j as its exemplar. The schematic diagram of message passing between data point x_i and data point x_j is shown in Fig. 1.1.

The responsibility $r(i,j)$ and the availability $a(i,j)$ are calculated according to Eqs. (1.2) and (1.3) in AP algorithm.

$$r(i,j) = s(i,j) - \max_{k\ s.t.\ k \neq j} \{a(i,k) + s(i,k)\} \tag{1.2}$$

$$a(i,j) = \begin{cases} min\{0, r(j,j) + \sum_{k \neq j} \max\{0, r(k,j)\}\} & i \neq j \\ \sum_{k \neq j} \max\{0, r(k,j)\} & i = j \end{cases} \tag{1.3}$$

AP algorithm may not be convergence in the condition of the number of iterations limited. Therefore, the damping factor λ is introduced to AP algorithm as weight to update the responsibility $r(i,j)$ and the availability $a(i,j)$ according to Eqs. (1.4) and (1.5). When the damping factor is larger, the update rate of the responsibility $r(i,j)$ and the availability $a(i,j)$ is slower, the process of the AP algorithm is more steady and the algorithm runs longer; when the damping factor is smaller, the update rate of the responsibility $r(i,j)$ and the availability $a(i,j)$ is faster, the process of the AP algorithm is more oscillate and the algorithm runs shorter on the contrary [11]. For the reason above, choosing a proper damping factor is crucial for the algorithm to improve the convergence rate.

$$R^{(l)} = (1 - \lambda)R^{(l)} + \lambda R^{(l-1)} \tag{1.4}$$

$$A^{(l)} = (1 - \lambda)A^{(l)} + \lambda A^{(l-1)} \tag{1.5}$$

1.2.2 General Steps of AP Algorithm

Concluded from the fundamentals of AP algorithm, the general steps of AP algorithm used in fault diagnosis are shown as follows:

1. Initializing the availability matrix $A^{(0)} = 0$, responsibility $R^{(0)} = 0$, initial clustering center $V^{(0)} = 0$; set the number of clusters C and the damping factor λ.
2. Calculate the similarity matrix S according to Eq. (1.1).
3. Set the maximum iterations *cycle*, and calculate the availability matrix $A^{(l)}$ and responsibility matrix $R^{(l)}$ after the l-th loops according to Eqs. (1.2) and (1.3).

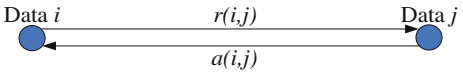


Fig. 1.1 Schematic diagram of message passing between data points

4. Calculate the weighted sum of availability matrix $A^{(l)}$ and $A^{(l-1)}$ according to Eq. (1.4), then obtain the new availability matrix $A^{(l)}$. Calculate the new responsibility $R^{(l)}$ with the same method as above according to Eq. (1.5).
5. Calculate the matrix $E^{(l)}$ according to Eq. (1.6).

$$E^{(l)} = R^{(l)} + A^{(l)} \tag{1.6}$$

6. Search for the k which satisfies $E^{(l)}(k, k) > 0$, then the corresponding sample x_k is an element of the clustering center matrix $V^{(l)}$.
7. Set a small positive number ε as the maximum permissible error, judge whether $\|U(l) - U(l-1)\| < \varepsilon$ is true and the number of clusters reaches the expectation. If true, then stop the loop. Otherwise, return to step 3 until the inequality is true or reaches the maximum loop.
8. Calculate the similarity degree between the samples to be diagnosed and each clustering center according to Eq. (1.7).

$$\sigma(X, V) = \frac{\sum_{k=1}^{m} \min(x_{ik}, v_{jk})}{\sum_{k=1}^{m} \max(x_{ik}, v_{jk})}, \ (i = 1, 2, \cdots n; \ j = 1, 2, \cdots C) \tag{1.7}$$

1.3 Fault Diagnosis Examples

When the auxiliary inverter failed, the spatial distribution of output signal energy is different from it in normal condition. The fault signals are analyzed in each band after multi-level decomposes so that the fault feature characteristics can be shown in the form of energy in different resolution subspace. Thereby, the fault feature can be obtained to distinguish the fault types [12].

Three fault types of auxiliary inverter which include voltage frequency variation, power supply interruption, and pulse transient, are simulated by using MATLAB software in the paper. Voltage frequency variation signal and normal signal are shown in Fig. 1.2, in which fault occurs between 0.5 $\sim$ 0.6 s. The sampling frequency of signals is 4096 Hz.

According to the method above, the authors extract 15 fault samples to calculate the clustering center matrix. Parts of the fault samples which the authors extract and its corresponding fault types are shown in Table 1.1.

In this paper, the authors set the number of cluster $C = 3$, the damping factor $\lambda = 0.58$, the maximum iteration $cycle = 200$ and the maximum permissible error

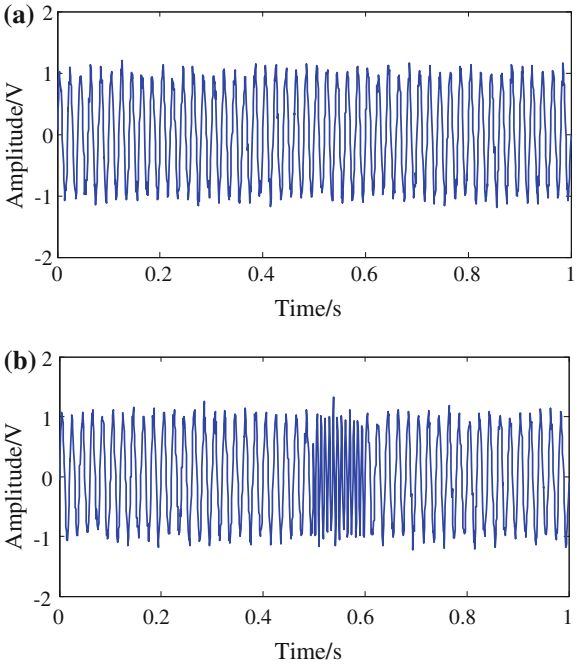


Fig. 1.2 The waveform of normal signal (**a**) and voltage frequency variation signal (**b**) in time domain

Table 1.1 Fault samples and the corresponding fault types

Sample numbers	Sample characteristics data	Fault type
1	(0.9160, 0.0342, 0.0025, 0.0351, 0.0024, 0.0040, 0.0028, 0.0029)	Voltage frequency variation
2	(0.8567, 0.0301, 0.0030, 0.0976, 0.0043, 0.0048, 0.0025, 0.0025)	
3	(0.9146, 0.0287, 0.0030, 0.0027, 0.0420, 0.0030, 0.0030, 0.0030)	
4	(0.4274, 0.3801, 0.0390, 0.1063, 0.0051, 0.0068, 0.0291, 0.0062)	Pulse transient
5	(0.4385, 0.4197, 0.0370, 0.0617, 0.0111, 0.0150, 0.0083, 0.0087)	
6	(0.4797, 0.3779, 0.0309, 0.0838, 0.0061, 0.0072, 0.0093, 0.0050)	
7	(0.5950, 0.1232, 0.0533, 0.0974, 0.0018, 0.0119, 0.0858, 0.0317)	Power interruption
8	(0.5929, 0.1188, 0.0505, 0.1072, 0.0020, 0.0118, 0.0836, 0.0332)	
9	(0.6019, 0.1284, 0.0510, 0.0916, 0.0019, 0.0122, 0.0833, 0.0297)	

Table 1.2 Comparison of fault diagnosis results

Fault type	Total of samples	FCM		AP	
		Accurate number	Accurate ratio (%)	Accurate number	Accurate ratio (%)
Voltage Frequency Variation	40	38	95.00	40	100.00
Pulse Transient	40	36	90.00	38	95.00
Power Interruption	40	36	90.00	38	95.00
Sum Accurate Ratio	120	91.67 %		96.67 %	

$\varepsilon = 0.01$ for AP algorithm. After 44 loops, the final clustering center matrix is obtained as the following matrix V which is constituted by the 3th, 6th, and 7th fault samples in Table 1.1.

$$V = \begin{bmatrix} x_3 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 0.9146 & 0.0287 & 0.0030 & 0.0027 & 0.0420 & 0.0030 & 0.0030 & 0.0030 \\ 0.4797 & 0.3779 & 0.0309 & 0.0838 & 0.0061 & 0.0072 & 0.0093 & 0.0050 \\ 0.5950 & 0.1232 & 0.0533 & 0.0974 & 0.0018 & 0.0119 & 0.0858 & 0.0317 \end{bmatrix}$$

There are 120 fault samples which include three fault types of auxiliary inverter. The authors achieve the fault types of samples by calculating the similar degree between samples and clustering center according to Eq. (1.7). The fault diagnosis results are shown in Table 1.2.

1.4 Conclusions

In this paper, the authors apply the AP algorithm in the field of fault diagnosis of metro vehicle auxiliary inverter. The initial clustering center is not needed for the AP algorithm, which is different from the FCM algorithm. Concluding from the results of simulation, the fault samples can be divided into groups by using the AP algorithm, and it has higher accuracy rate than FCM. However, the damping factor, which has influence on the convergence of AP algorithm, is set to be a constant in the paper. Therefore, how to modify the value of damping factor at any time to improve the convergence rate of AP algorithm is the key to the future research.

Acknowledgments This work is partially supported by the National Science and Technology Pillar Program (2011BAG01B05), the Foundation of Shandong Province (BS2011DX008, ZR2011FM008), and the State Key Laboratory of Rail Traffic Control and Safety Foundation (RCS2011K005), Beijing Jiaotong University.

References

1. Yao DC, Yang JW, Yin YF et al (2010) Fault diagnosis of railway bearing based on muti-method fusion techniques. Mach Des Res 26(3):70–73 (in Chinese)
2. Huang YC (2003) Condition assessment of power transformers using genetic-based neural networks. IEE Proc Sci Meas Technol 150(1):19–24
3. Bezdek JC (1974) Cluster validity with fuzzy sets. J Cybern 3(3):58–72
4. Ma ZT, Gao JW, Leng ZW et al (2013) A method of metro vehicle auxiliary inverter fault diagnosis based on fuzzy clustering arithmetic. J Qingdao Univ Eng Technol Ed 28(3):8–14 (in Chinese)
5. Ma ZT, Gao JW, Zhang B et al (2013) A method of metro vehicle auxiliary inverter fault diagnosis based on weighted fuzzy clustering algorithm. Lect Notes Electr Eng 254:633–639
6. Gao JW, Yu JP, Leng ZW et al (2013) The application of PSO-LSSVM in fault diagnosis of subway auxiliary inverter. ICIC Express Lett 4(3):777–784
7. Meng XY, Han XJ (2007) The fuzzy C-means clustering algorithm and its application in the fault diagnosis of ships. Shipbuild China 48(4):98–103 (in Chinese)
8. Xiong H, Sun CX, Liao RJ et al (2005) Study on kernel-based possibilistic clustering and dissolved gas analysis for fault diagnosis of power transformer. Proc CSEE 25(20):163–167 (in Chinese)
9. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315(5814):972–976
10. Xiao Y, Yu J (2008) Semi-Supervised clustering based on affinity propagation algorithm. J Softw 19(11):2803–2813 (in Chinese)
11. Wang KJ, Zhang JY, Li D et al (2007) Adaptive affinity propagation clustering. Acta Automatica Sin 33(12):1242–1246 (in Chinese)
12. Yang JW, Cai GQ, Yao DC et al (2010) Fault diagnosis method for the rolling bearing of railway vehicle based on wavelet packet transform and BP neural network. China Railw Sci 31(6):68–73 (in Chinese)

Chapter 2
A Soft-Switching Control Method of Isolated LC Series Resonant Transformer Full Bridge DC–DC Converter

Meng Jiang and Wei Li

Abstract For the different characteristics of nonresonant and resonant isolated bidirectional full bridge DC–DC converter, a unified expression of power transmissions is derived from two DC–DC converters. The power transfer characteristics could be unified described through the power expression. The problem of isolated bidirectional DC–DC converter is that the switching loss increases and the converter efficiency declines with the forced turn-on or turn-off of switch devices in high-frequency situation. In this paper, to solve this problem, an isolated LC series resonant transformer full bridge DC–DC converter is taken as research object, a phase-shift control strategy which could realize zero voltage turn-on and decrease the turn-off current of the power devices to decrease switching loss and increase the efficiency is proposed. The validity of proposed control strategy is verified through simulation and experiment results.

Keywords Bidirectional full bridge DC–DC converter · Series resonant · Soft-switching technology · First harmonic analysis · Voltage gain

2.1 Introduction

The isolated bidirectional full bridge DC/DC converter, known as Dual Active Bridge (DAB), has the advantages that medium (or high) frequency transformer could realize electrical isolation, which improves system reliability, and soft

M. Jiang (✉)
Management Department, Tianjin University, 92 Weijin Road, Nankai District, Tianjin, People's Republic of China
e-mail: jiangmeng168@163.com

M. Jiang · W. Li
Rescuing Center for Mineral Disaster of Henan, No. 116, East Second Street, Zhengzhou, People's Republic of China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_2,
© Springer-Verlag Berlin Heidelberg 2014

switch control of both primary side and secondary side of transformer that could reduce system loss. Bidirectional energy flow of converter could be achieved. High frequency transformer is used to replace the power frequency transformer, which could reduce the size and mass and improve power density of the system. Such converters are widely used in applications like power electronic transformer [1], locomotive traction [2], renewable energy power generation [3], high voltage motor drive [4, 5].

DC–DC converter, which is an important unit of power electronic transformer is utilized to achieve electrical isolation and voltage level conversion. According to the different topologies, it could be divided into nonresonant and resonant DC–DC converters. According the different control strategies, it could be divided into phase-shift control and combination of phase-shift and PWM modulation [6–12].

The main work in this paper is listed as:

(1) The unified power transfer expression, which could describe the relationship of resonant and nonresonant DC–DC converter, is derived in this paper. The derived expression helps in modeling and Analysis of DC–DC converter.
(2) Through the comparison and analysis of nonresonant and resonant DC–DC converter, a control strategy based on phase-shift control for resonant DC–DC converter is proposed in this paper.

2.2 The Principle of DC/DC Converter

Figure 2.1 shows the topology of an isolated LC series resonant transformer full bridge DC/DC converter. It consists of two active full bridges, a series resonance unit and an isolation transformer. L_r is the equivalent of primary and secondary leakage inductance of high frequency transformer. U_{ab} and U_{cd} are 50 % duty cycle square waves. ϕ is the phase-shift angle of U_{ab} and U_{cd}.

Considering the affection of line impedance of LC series resonance circuit and magnetic inductance of transformer, the approximate equivalent circuit and its phasor diagram of DC–DC converter is shown in Figs. 2.2 and 2.3.

where θ, β, ϕ are respectively phase differences among current i_r, voltage U_{ab}, u'_{cd}. U_{ab_f} and u'_{cd_f} are respectively the fundamental component of voltage U_{ab} and u'_{cd}. Z_{eq} is the equivalent impedance of R_o in the equivalent circuit.

Set ω_r, ω_s are resonant angular frequency and switching frequency, f_r, f_s are resonant frequency and switching frequency, n is turns ratio of transformer.

The fundamental of u_{ab} and u_{cd}, which are defined as u_{ab_f} and u_{cd_f}, could be expressed as follows:

$$u_{ab_f}(t) = \frac{4}{\pi} U_i \cos \omega_s t \tag{2.1}$$

$$u_{cd_f}(t) = \frac{4}{\pi} U_o \cos(\omega_s t - \phi) \tag{2.2}$$

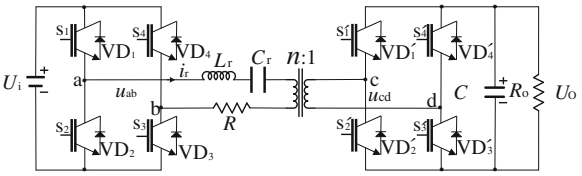


Fig. 2.1 Isolated bidirectional full bridge DC/DC converter

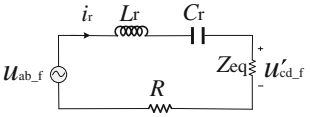


Fig. 2.2 Approximate equivalent circuit of DC–DC converter

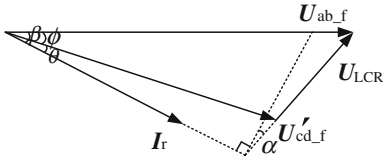


Fig. 2.3 Phasor diagram of the equivalent circuit

In the approximate equivalent circuit, u'_{cd_f} could be expressed as

$$u'_{cd_f}(t) = \frac{4}{\pi} U'_o \cos(\omega_s t - \phi) \tag{2.3}$$

The Impedance of R, L_r and C_r in the LRC resonant circuit could be derived as

$$Z = R + j(\omega_s L_r - \frac{1}{\omega_s C_r}) = |Z| \angle (90^o - \alpha) \tag{2.4}$$

$$|Z| = \sqrt{R^2 + (\omega_s L_r - \frac{1}{\omega_s C_r})^2} \tag{2.5}$$

So

$$\dot{I}_r = \frac{\dot{U}_{ab_f} - \dot{U}'_{cd_f}}{Z} = \frac{2\sqrt{2}}{\pi |Z|} (U_i \angle (\omega_s t - 90^o + \alpha) - U'_o \angle (\omega_s t - \phi - 90^o + \alpha)) \tag{2.6}$$

The output instantaneous power p_o,which is transferred through the secondary of transformer, could be derived as

$$p_o = u'_{cd_f}(t) i_r(t) = \frac{16 n U_o}{\pi^2 \sqrt{R^2 + (\omega_s L_r - \frac{1}{\omega_s C_r})^2}} \cdot$$

$$\left\{\frac{1}{2}U_i[-\sin(2\omega t+\alpha-\phi)+\sin(\phi+\alpha)]-\frac{1}{2}U_o'[-\sin(2\omega t+\alpha-2\phi)+\sin\alpha]\right\} \tag{2.7}$$

Considering $U_o' = nU_o$, in the $0 \sim T$ period, the average power P_o of DC–DC converter could re-expressed as

$$P_o = \frac{8nU_o[U_i\sin(\phi+\alpha)-nU_o\sin\alpha]}{\pi^2\sqrt{R^2+(\omega_s L_r-\frac{1}{\omega_s C_r})^2}} \tag{2.8}$$

Set $F = \frac{\omega_s}{\omega_r} = \frac{f_s}{f_r}$, $Q = \frac{\omega_r L_r}{R_o'}$, so $\frac{\omega_s L_r}{R_o'} = QF$ $\frac{1}{\omega_s C_r R_o'} = \frac{Q}{F}$

The average output power P_o is derived as

$$P_o = \frac{8n\cdot U_o\cdot[U_i\cdot\sin(\phi+\alpha)-nU_o\sin\alpha]}{\pi^2\sqrt{R^2+(w_r L_r)^2\cdot(F-1/F)^2}} \tag{2.9}$$

Formula (2.9) indicates that the average transfer power of DC/DC converter is determined by phase-shift ϕ and output voltage U_o.

2.3 Traditional Control Method of DC–DC Converter and Existing Problem

Traditional topology of DC–DC converter is shown in Fig. 2.4, where R is the circuit impedance, L is the leakage inductance of high frequency isolated transformer.

From formula (2.9), in nonresonant situation, the converter output power P_o is re-expressed as

$$P_o = \frac{8n\cdot U_o\cdot U_i}{\pi^2\cdot w_s\cdot L}\sin\phi \tag{2.10}$$

Under PWM modulation, the typical waveforms of voltage U_{ab} and current i_r is shown in Fig. 2.5.

From Fig. 2.5, the turn-off current is very high in hard-switching mode. The forced shut-off mode of switching devices is called hard-switching. The circuit structure is simple, however, the switching loss is greater.

2.4 Improved Control Method of DC–DC Converter

To decrease the switching loss and improve converter efficiency, in this paper, an isolated LC series resonant transformer full bridge DC/DC converter, which is

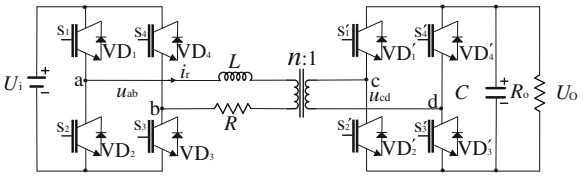


Fig. 2.4 Nonresonant bidirectional full bridge DC–DC converter

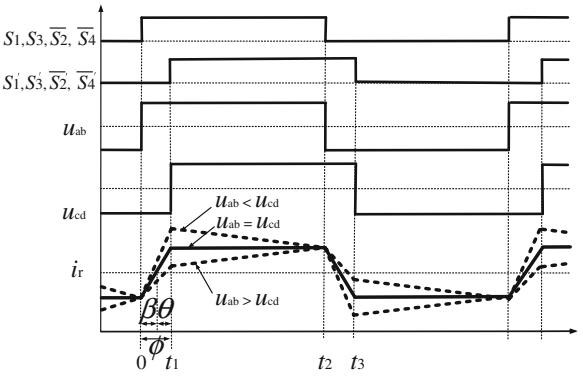


Fig. 2.5 Waveforms of U_{ab}, U_{cd} and i_r in hard-switching mode

shown in Fig. 2.1, is taken into research. Two power transmission modes are given based on analysis of formula (2.9).

2.4.1 Synchronous Control Mode

If $F = 1$, $\phi = 0$, formula (2.9) is expressed as

$$P_o = \frac{8n \cdot U_o}{\pi^2 \cdot R}(U_i - n \cdot U_o) \tag{2.11}$$

From the output load power:

$$P_{load} = \frac{U_o^2}{R_o} \tag{2.12}$$

From formula (2.11) and (2.12), the transmission power could be re-expressed as

$$P_o = \frac{8n \cdot U_i \cdot R_o}{8n^2 R_o + \pi^2 R} \tag{2.13}$$

Formula (2.13) indicates that power transmission is determined by differential of input and output voltage in synchronous mode. The output voltage U_o is determined by load R_o. Assuming the turn ratio of transformer is 1, the typical waveforms of voltage U_{ab} and current i_r is shown in Fig. 2.6.

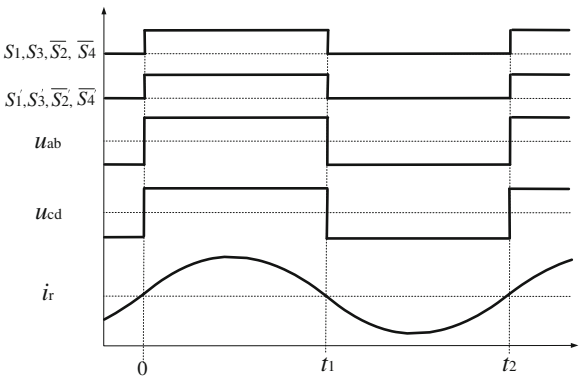


Fig. 2.6 Waveforms of U_{ab}, U_{cd} and i_r in synchronous control mode

Figure 2.6 indicates that it is a synchronization action for switching devices of both primary and secondary sides under such power transmission mode. The output voltage U_o is determined by load R_o. This method is suitable for the situation that the variation range of U_o is not great.

2.4.2 Phase-Shift Control Method

If $F \neq 1$, $\phi \neq 0$, neglecting circuit impedance, $\alpha = 0$, the average transmission power P_o could be described as

$$P_o = \frac{8n \cdot U_o \cdot U_i}{\pi^2 \cdot (w_s L_r - \frac{1}{w_s C_r})} \tag{2.14}$$

Via adjusting phase-shift ϕ, the output power P_o and output voltage U_o could be closed-loop controlled. Assuming the turn ratio of transformer is 1, the typical waveforms of voltage U_{ab} and current i_r is shown in Fig. 2.7.

The voltage gain of the equivalent circuit, namely the ratio of output voltage and input voltage could be defined as

$$M = \frac{\left|U'_{cd_f}\right|}{\left|U_{ab_f}\right|} \tag{2.15}$$

Figure 2.7 shows that when $\omega_s t = 0$, $i_r(0) < 0$, namely $\sin(-\beta) < 0$, the primary side achieved ZVS and when $\omega_s t = \phi$, $i_r(t_1) > 0$, namely $\sin(-\theta) < 0$, the secondary side achieved ZVS. So constraint equations of soft switch could be derived as follows:

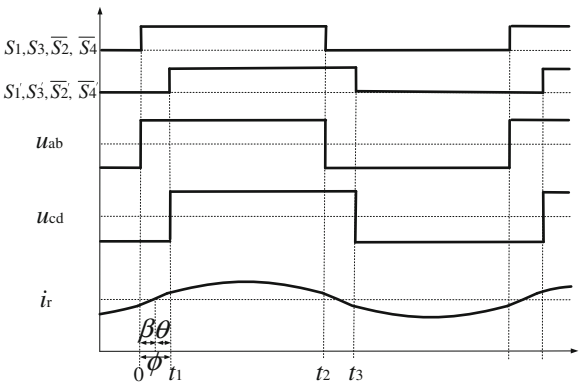


Fig. 2.7 Waveforms of U_{ab}, U_{cd} and i_r in phase-shift control mode

$$\begin{cases} M\cos\phi - 1 < 0 \\ M - \cos\phi > 0 \end{cases} \tag{2.16}$$

If only $M = 1$, the equations could be satisfied. In this mode, the realization of soft-switching is independent of the load. It means that ZVS of primary side and secondary side could be achieved over a wide load range.

2.5 Simulation and Experiment

2.5.1 Simulation

From the above theoretical analysis, PSIM is utilized to verify the proposed strategy. The simulation parameters are list as the following.

In nonresonant situation, L is 30 μH, switching frequency f_s is 6 kHz, load R_o is 7 Ω. The input voltage U_i is, respectively, set to 50 V and 100 V, (corresponding output power 0.36 kW and 1.43 kW), the waveforms of driving signal, U_{ab}, U_{cd} and i_r are shown in Figs. 2.8 and 2.9.

In resonant situation, the resonant inductor L_r is 30 μH, resonant capacitor C_r is 20 μF, resonant frequency f_r is 5.3 kHz, switching frequency f_s is 6 kHz, load R_o is 7 Ω. When input voltage U_i is, respectively, set to 50 and 100 V, (corresponding output power is 0.36 and 1.43 kW), the waveforms of driving signal, U_{ab}, U_{cd} and i_r are shown in Figs. 2.10 and 2.11.

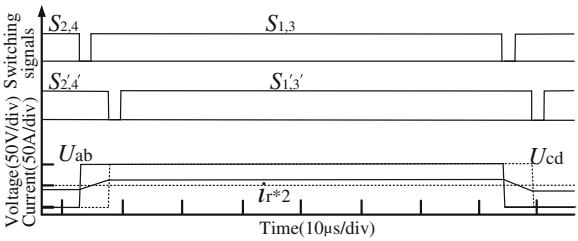


Fig. 2.8 Waveforms of U_{ab}, U_{cd} and i_r when $U_i = 50$ V

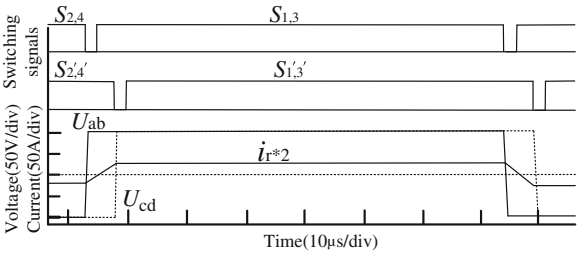


Fig. 2.9 Waveforms of U_{ab}, U_{cd} and i_r when $U_i = 100$ V

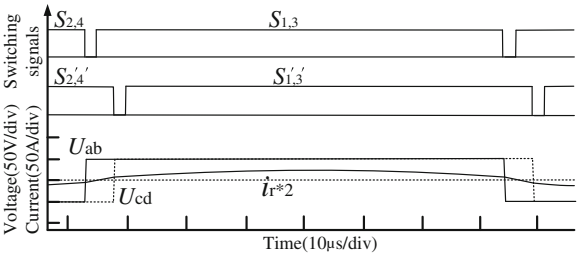


Fig. 2.10 Waveforms of U_{ab}, U_{cd} and i_r when $U_i = 50$ V

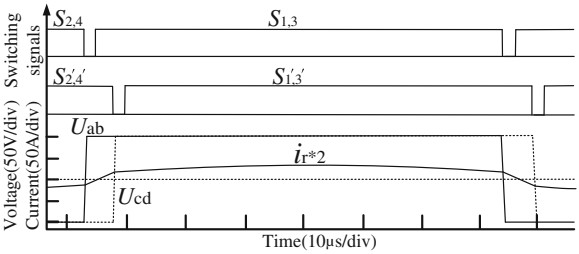


Fig. 2.11 Waveforms of U_{ab}, U_{cd} and i_r when $U_i = 100$ V

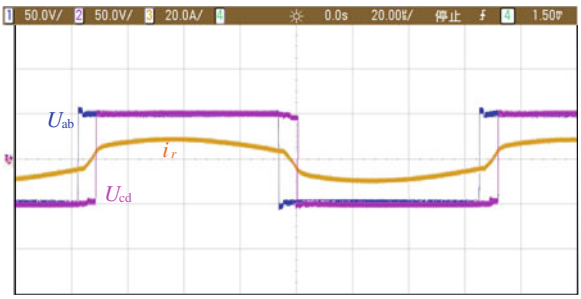


Fig. 2.12 Waveforms of U_{ab}, U_{cd} and i_r when $U_i = 50$ V

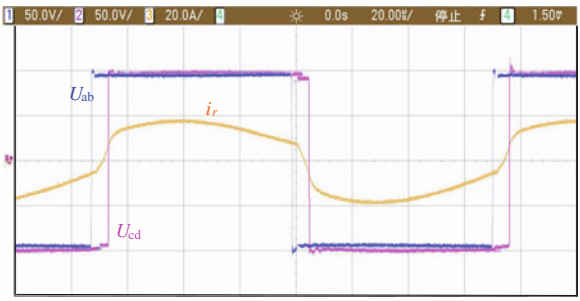


Fig. 2.13 U_{ab}, U_{cd} and i_r when $U_i = 50$ V

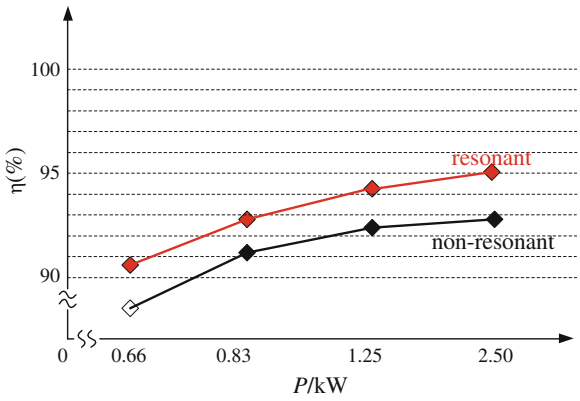


Fig. 2.14 Efficiency curves of DC–DC converter

2.5.2 Experiment

The parameters of actual circuit are as follows: transformer turns ratio r = 1, $L_s = 30$ μH, $P_s = 10$ kW, $L_r = 15$ μH, $C_r = 20$ μF, $R_o = 7$ Ω, $f_s = 6$ kHz, $f_r = 5.3$ kHz, $t_{dead} = 2$ μs.

Where r is transformer ratio, P_s is rated power, t_{dead} is switching tube dead-time.

Figures 2.12 and 2.13 shows the waveforms of U_{ab}, U_{cd} and i_r when input voltage is respectively 50 and 100 V.

In order to verify the efficiency of soft-switching of the system, experiment curves are depicted to calculate system efficiency in non-resonant mode and resonant mode. The output power are, respectively 0.36, 1.43, 3.21 and 5.71 kW. Comparison results are shown in Fig. 2.14.

Experiments results indicate that the efficiency is higher under resonant mode.

2.6 Conclusion

Based on the comparison of power transfer mode and control features for non-resonant and LC resonant DC–DC converter, the conclusion is:

(1) The proposed power transfer unified expression could describe the power transmission characteristics of nonresonant and resonant DC–DC converter in one formula. This expression is helpful for analysis and modeling of DC–DC converter

(2) Under phase-shift control method for resonant mode of DC–DC converter, the turn-off current could be decreased, and the efficiency could be improved.

References

1. Zhang M, Liu J, Jin X (2011) Research on the FREEDM micro-grid and its relay protection. Power Syst Prot Control 39(7):95–99
2. Li L, Liu G (2011) Development of bi-directional DC–DC converter in multiple battery energy storage system. Power Syst Prot Control 39(3):90–94
3. Du C, Zhang C, Chen A et al (2011) Digital control and implementation of photovoltaic soft-switching DC–DC converter with high-frequency step-up transformer isolation. Trans Chin Electrotech Soc 26(8):57–63
4. Liu H, Mao C, Lu J et al (2010) Energy storage system of electronic power transformer and its optimal control. Trans Chin Electrotech Soc 25(3):54–60
5. Demetriades GD, Nee HP (2008) Small-signal analysis of the half-bridge soft-swithing uni-directional converter employing extended state-space averaging. In: Power electronics specialists conference, Rhodes, pp 385–391
6. Demetriades GD, Nee HP (2008) Characterization of the dual-active bridge topology for high-power applications employing a duty-cycle modulation. In: Power electronics specialists conference, Rhodes, pp 2791–2798
7. Bai H, Nie Z, Chris CM (2010) Experimental comparison of traditional phase-shift, dual-phase-shift, and model-based control of isolated bidirectional DC–DC converters. IEEE Trans Power Electron 25(6):1444–1449
8. Oggier GG, Gar′cia GO, Oliva AR (2009) Switching control strategy to minimize dual active bridge converter losses. IEEE Trans Power Electron 24(7):1826–1838
9. Ortiz G, Biela J, Bortis D et al (2010) 1 megawatt, 20 kHz, isolated, bidirectional 12 kV to 1.2 kV DC–DC converter for renewable energy applications. In: Power electronics conference, Singapore, pp 3212–3219
10. Lenke R, Mura F, De Doncker RW (2009) Comparison of non-resonant and super-resonant dual-active ZVS-operated high-power DC–DC converters. In: European conference on power electronics and applications, 2009, pp 1–10
11. Li X, Ashoka KSB (2010) Analysis and design of high-frequency isolated dual-bridge series resonant DC–DC converter. IEEE Trans Power Electron 25(4):850–862
12. Wu L, Zhang Y, Li Z, Wang P, Li Y, Liu Z (2012) A control strategy of isolated bidirectional full bridge DC/DC converter. Electric Mach Control 16(12):21–27

Chapter 3
Study on Efficiency Optimization of Medium Frequency Transformers for Rail Transit Traction System

Juanjuan Zhang, Yumei Du, Zixin Li, Ping Wang and Yaohua Li

Abstract Power electronics transformers (PETs) have the advantages of small volume and light weight, so they have promising applications in rail transit traction systems substituting traditional transformers. The PET consists of the medium frequency transformer (MF transformer) and power electronic converters on the primary and secondary sides. Among these components, the MF transformer efficiency affects overall system power transfer capacity, and plays an important role as it determines whether the PET system could work with high performance. This paper aims at efficiency optimization of MF transformers and studies the relation between efficiency and core materials, operating parameters, and geometric parameters to give MF transformer designs some useful references.

Keywords PETs · MF transformer · Efficiency optimization

3.1 Introduction

Power frequency transformers are widely used in the power supply installation of traditional rail transit traction systems, but they have the disadvantages of large volume and high weight. With the fast development of power electronic devices, power electronics transformers (PETs) have gained more and more attention.

Figure 3.1 demonstrates the structure of PET. It can be divided into three components: the power electronic converter on the primary side transmits 50 Hz AC power to medium frequency AC power; the power electronic converter on the secondary side transmits medium frequency AC power back to the lower frequency power; medium frequency transformer (MF transformer) connects these

J. Zhang (✉) · Y. Du · Z. Li · P. Wang · Y. Li
Key Laboratory of Power Electronics and Electric Drive, Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing, China
e-mail: zhangjuanjuan@mail.iee.ac.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_3,
© Springer-Verlag Berlin Heidelberg 2014

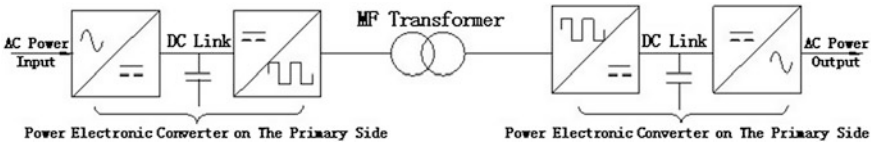


Fig. 3.1 Schematic of power electronics transformer

two converters and acts as voltage conversion, energy transfer, and electrical isolation. From the input and output PET works like a traditional transformer, but the volume and weight have decreased significantly due to operating at a higher frequency. It has promising applications in rail transit traction systems substituting traditional transformers [1].

On the downside, transformer loss at medium frequency is higher so as to reduce the efficiency. For high efficiency and high power density of PET, optimization design of MF transformer is of great importance. In [2], operating frequency influences on the volume and weight of MF transformers were studied. In [3, 4] the overall design of MF transformers were analyzed and efficiency, weight, and volume variation effected by isolation and thermal structure are considered.

In this paper, the relation between efficiency and core materials, operating parameters (magnetic inductance B and current density J), and geometric parameters of transformers are analyzed aiming at raising the MF transformer's efficiency. It provides some useful references for MF transformer design to some extent.

3.2 Transformer Model

Figure 3.2 shows the configuration of the MF transformer. Two paralleled C cores looking like an E core and coaxial-foil-type windings form the shell-type single-phase transformer. The primary winding is inside and the secondary winding is outside. Foil-type windings are usually modeled as a whole component as Fig. 3.2 to simply analyze with bearable error.

The efficiency of transformer is calculated from Eq. (3.1), where P_{loss} consists of core loss P_c and winding loss P_w, and P_1 is the input power.

$$\eta = 1 - P_{\text{loss}}/P_1 \tag{3.1}$$

in (3.1), $P_{\text{loss}} = P_c + P_w$

Core losses: when calculating core loss, the modified Steinmetz equation is applied considering the rate of change of magnetic induction *(dB(t)/dt)* impacts on core loss [5]:

$$P_v = f_r K f_{\text{eq}}^{\alpha-1} B_m^{\beta} \tag{3.2}$$

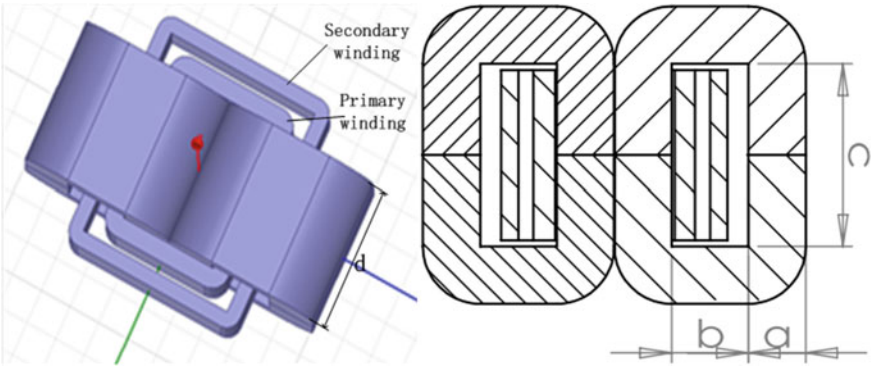


Fig. 3.2 Configuration of medium frequency transformer

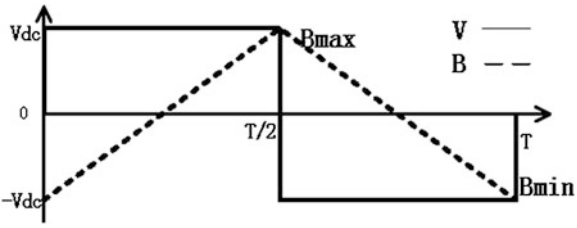


Fig. 3.3 Voltage and magnetic induction waveforms

in (3.2), $f_{\text{eq}} = \frac{2}{\Delta B^2 \pi^2} \int_0^T \left(\frac{dB(t)}{dt}\right)^2 dt$

In Eq. (3.2), P_v represents core loss density (core loss per unit volume), f_r is fundamental frequency of the input voltage, K, α, β are characteristic parameters of core materials given by core manufacture, B_m is absolute maximum operating magnetic induction, and f_{eq} is equivalent frequency. Under the conditions specified in this paper, the voltage is square wave so magnetic induction is triangular wave as in Fig. 3.3. From Fig. 3.3 we can get:

$$\frac{dB(t)}{d(t)} = \frac{2\Delta B}{T}, \quad f_{\text{eq}} = \frac{\pi^2}{8}$$

into (3.2),

$$P_v = \left(\frac{\pi^2}{8}\right)^{\alpha-1} K f_r^{\alpha} B_m^{\beta} \tag{3.3}$$

Equation (3.3) represents core loss density of the MF transformer under square-wave voltage input.

Winding losses: On high frequency operating conditions, skin and proximity effects inside and between the windings increase and they will influence the winding resistance. Therefore, high frequency effects should be taken into

consideration in winding loss calculation. Dowell model is introduced because it calculates AC resistance precisely considering high frequency impacts on winding resistance [6]. The ratio of AC resistance to DC resistance is expressed in Eq. (3.4), where winding thickness is per layer, is skin depth, and m is the number of winding layers.

Equation (3.7) calculates the core loss using Dowell model. Since the current is quasi-sine wave when PET operates on series resonant condition, winding loss of harmonic currents are very small and can be neglected.

$$F_R = \frac{R_{ac}}{R_{dc}} = y[M(y) + \frac{2}{3}(m^2 - 1)D(y)] \tag{3.4}$$

In (3.4),

$$M(y) = \frac{\sinh(2y) + \sin(2y)}{\cosh(2y) - \cos(2y)} \tag{3.5}$$

$$D(y) = \frac{\sinh(y) - \sin(y)}{\cosh(y) + \cos(y)} \tag{3.6}$$

$$P_w = R_{dc}F_R(I_{rms})^2 = R_{ac}(I_{rms})^2 \tag{3.7}$$

From Eqs. (3.3) and (3.7), it can be seen that core loss only depends on characteristic parameters of core materials and magnetic inductance when operating frequency is given. As for winding loss, I_{rms} can be calculated from specified output voltage and output power, while R_{ac} is influenced by parameters like current density J, transformer geometric structure, etc. This paper will give some advice on core material selection and study how the operating parameters (magnetic inductance B and current density J) and geometric parameters affect transformer efficiency.

3.3 Optimization Design of Medium Frequency Transformers

Generally, optimization design of MF transformers starts from selecting core and wind materials and designing structure according to required technical specification. Then the geometric parameters and operating parameters are adjusted to satisfy the temperature ascension or loss requirement. The design flowchart is shown in Fig. 3.4. In this paper, a 100 kW, 5 kHz transformer with a ratio 1700 V:750 V is discussed.

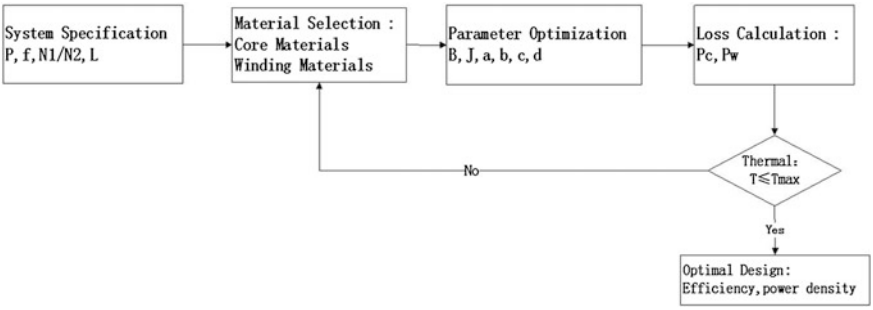


Fig. 3.4 Design procedure of the MF transformer

3.3.1 Selection of Core Materials

Silicon steel sheet and permalloy are often selected as core materials of power frequency transformer, but MF transformers require those core materials with high saturation magnetic inductance, low loss, and high curie point. Table 3.1 lists the performance of several soft magnetic materials for MF transformers. The core loss density P_v at 5 kHz is also given.

The typical soft-core materials at medium frequency are ferrite, amorphous alloy, and nanocrystalline. It can be seen from Table 3.1 that ferrite core loss is low at 5 kHz. It also has the advantages of various core shape, low cost, etc. [7]. However, saturation flux density of ferrite is only 0.5 T, which will cause a larger transformer volume. Besides, the magnetic performance at high temperature is not stable: saturation magnetic inductance has decreased under 0.4 T at 100 °C.

Amorphous materials have excellent performance: high saturation magnetic inductance, high permeability, and high mechanical robustness. Moreover, cut C-cores of multiple sizes are commercially available and so they provide much convenience to optimization design. On the downside, the core loss is higher than the other two materials.

The nanocrystalline core material is especially suitable for high frequency and high power situation with high saturation magnetic inductance and extreme low loss. However, the majority of nanocrystalline cores available are toroidal and cannot undertake too large power transfer.

To summarize, nanocrystalline cores geometry are limited and the cost will be high if ordering specific C-core, so nanocrystalline cannot be the best choice in spite of low loss. The ferrite leads to a larger volume, lower power density, and fluctuations in performance with the temperature. Synthesizing all the factors like efficiency, power density, and cost, amorphous materials are suggested to be selected for the MF transformer cores.

Table 3.1 Several core materials for MF transformers

Material	Series	Saturation density/T	Frequency/ Hz	Permeability	Curie point (°C)	Pv (mW/ cm^3)
Ferrite	3C96	0.5	<400 k	2,200	>240	32.7
Amorphous alloy	2605	1.56	5–100 k	4,000–80,000	395	107.7
	2705	0.77	5–250 k	5,000–50,000	365	70.2
Nanocrystalline	VAC 500F	1.2	20–200 k	20,000–30,000	600	6.7
	Namglass4	1.23	<200 k	10,000–100,000	570	9.7

3.3.2 *Efficiency Variation with Flux Density (B_m) and Current Density (J)*

This section focuses on eight transformers of different core sizes and studies efficiency variation with magnetic Inductance (B_m) and current Density (J). All the core sizes are demonstrated in Table 3.2, where A_p is the product area of the two paralleled C-cores. It is the product of cross-section area A_e and window area A_w. Parameters *a, b, c* and *d* are, namely core width, window width, window height, and core thickness (See Fig. 3.2).

The saturation magnetic inductance (B) of amorphous material is 1.5 T. To ensure transformer normal operation, the magnetic inductance is selected from 0.5 to 1 T leaving enough margin. Generally, current density J ranges from 3 to 5 A/m^2. Efficiency changes with magnetic inductance B specifically are calculated and demonstrated in Fig. 3.5.

It can be seen from Fig. 3.5 that when current density is fixed, transformer efficiency decreases with increasing magnetic inductance. However, when current density changes, from Fig. 3.5a–c, efficiency of eight sizes transformers does not change significantly. Take the series **3** which has the highest efficiency, for example (the curves marked by green triangles in Fig. 3.5). Efficiency variation ranges at three types of current density are, respectively, 99.55–99.82 %, 99.55–99.82 %, and 99.56 ~ 99.82 %. It is concluded that current density has small influences on efficiency under the operating condition required in this paper.

3.3.3 *Efficiency Variation with Core Shape*

In this section, efficiency variation with different core shapes is studied by changing the values of parameters *a, b, c* and *d* of eight sizes transformers while keeping A_p, A_e, A_w constant [8]. In addition, the values of magnetic inductance and current density of each transformer have been selected based on the above section's conclusion.

Table 3.2 Geometric parameters of cores

Series	A_p(cm^4)	a(mm)	b (mm)	c(mm)	d(mm)
1	1026.7	25	40	88	70
2	1101.8	32.5	45	113	40
3	1231.5	20	45.5	203	40
4	1490.4	27	46	90	80
5	1750.0	40	50	210	25
6	1811.3	25	45	161	60
7	1833.8	25	45	163	60
8	2002.4	32	76	124	40

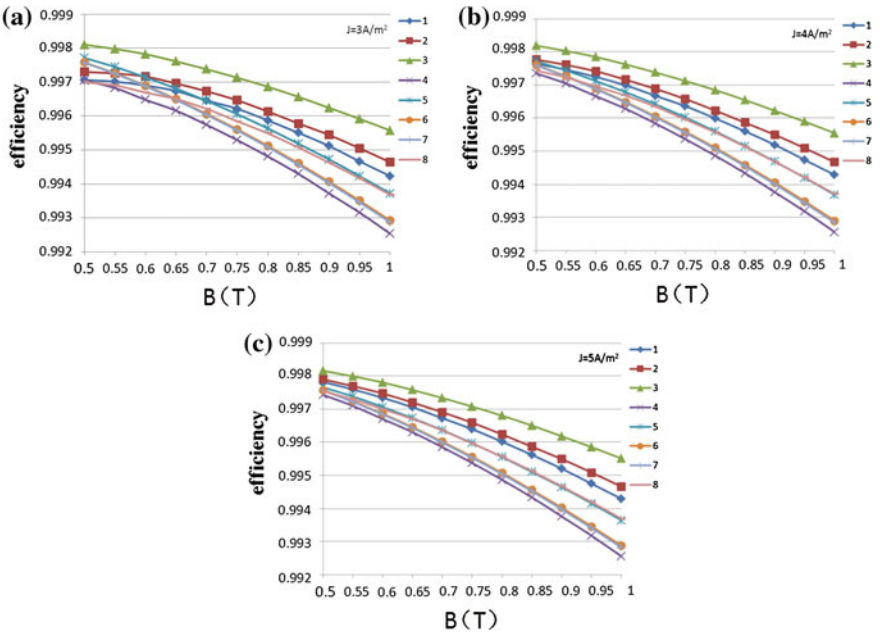


Fig. 3.5 Efficiency variation with magnetic inductance at different current densities. **a** $J = 3$ A/m^2, **b** $J = 4$ A/m^2, **c** $J = 5$ A/m^2

Figure 3.6 shows efficiency variation with core width a. It can be seen that efficiency rises faster when core width value is small. Then efficiency reaches the maximum point around 15 mm. After that, efficiency decreases slowly.

From the perspective of the transformer manufacture, the core window width b is related to winding size, turns, and so on. So the utility ratio of the core window usually has empirical value ranges. This section analyzes efficiency variation with window width b (shown in Fig. 3.7) under the condition that all transformers have reasonable utility ratio of core window areas. From Fig. 3.7, for all these eight sizes transformers, efficiency first increases and then decreases and the decline is greater when b is larger.

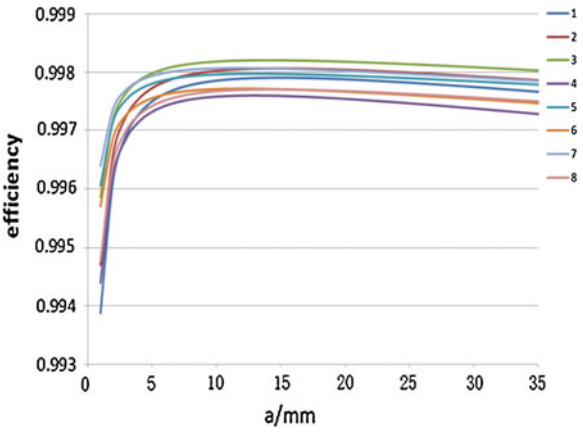


Fig. 3.6 Efficiency variation with core width *a*

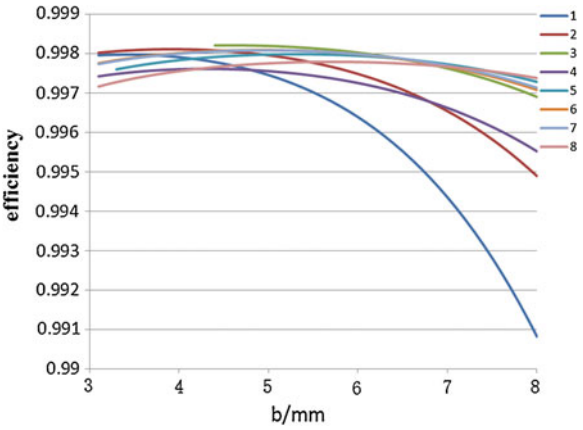


Fig. 3.7 Efficiency variation with window width *b*

Table 3.3 Optimum structure and maximum efficiency of different transformers

Series	*a*/mm	*b*/mm	Efficiency/ %
1	15	34	99.80
2	15	39	99.81
3	14	46	99.82
4	13	53	99.77
5	13	53	99.8
6	12	49	99.81
7	12	50	99.81
8	14	57	99.78

By the above results, there exists an optimum structure when the product area of transformer A_p is given. That is to say, efficiency can get maximum if parameters *a, b, c*, and *d* take specific values. These values of eight sizes transformers are listed in Table 3.3.

3.4 Conclusions

This paper aims at efficiency optimization of MF transformers and studies how efficiency is influenced by core materials, operating parameters, and geometric parameters of transformers. The following conclusions can be obtained:

- Amorphous materials have excellent performance and multiple C-core sizes are commercially available. Although the core loss is higher, it is enough to satisfy the power transition requirement. Therefore, amorphous cores are the better choice for MF transformers.
- Under the operating condition required in this paper, small magnetic inductance value leads to high efficiency, while current density has small influences on efficiency.
- When the product area of transformer A_p is given, there exists an optimum structure that makes efficiency maximum.
- Among the eight types of transformer sizes analyzed, the best design for optimal efficiency is given as: efficiency can reach 99.82 % with optimized series **3**, when magnetic inductance $B = 0.5$ T and current density $J = 4.6$ A/m^2

References

1. Reinold SH (2007) Medium frequency topology in railway applications. In: 2007 European Conference on Power Electronics and Applications, pp. 1–10
2. Gu W-J, Liu R (1993) A study of volume and weight vs. frequency for high-frequency transformers. In: 24th Annual IEEE Power Electronics Specialists Conference PESC'93 Record, pp. 1123–1129
3. Drofenik U (2012) A 150 kW medium frequency transformer optimized for maximum power density. In: 7th International Conference on Integrated Power Electronics Systems (CIPS), pp. 1–6
4. Ortiz G, Biela J, Kolar JW (2010) Optimized design of medium frequency transformers with high isolation requirements. In: IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society, pp. 631–638
5. Villar I, Rufer A, Viscarret U, Zurkinden F, Etxeberria-Otadui I (2008) Analysis of empirical core loss evaluation methods for non-sinusoidally fed medium frequency power transformers. In: IEEE International Symposium on Industrial Electronics ISIE 2008, pp. 208–213
6. Dowell P (1966) Effects of eddy currents in transformer windings. In: Proceedings of the Institution of Electrical Engineers, vol. 113, pp. 1387–1394
7. Du Y, Baek S, Bhattacharya S, Huang AQ (2010) High-voltage high-frequency transformer design for a 7.2 kV to 120 V/240 V 20kVA solid state transformer. In: IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society, pp. 493–498
8. Pavlovsky M, Haan SWD, Ferreira JA (2006) Integral design of 50 kW, 25 kHz dc–dc converter based on ZVS, Quasi-ZCS topology. In: 4th International Conference on Integrated Power Systems (CIPS), 2006, pp. 1–6

Chapter 4
Primary Permanent Magnet Linear Motors for Rail Transit

Ruiwu Cao and Ming Cheng

Abstract Permanent magnet (PM) linear motor having magnets and armature windings in the primary mover (the so-called primary-PM machines) have attracted more and more attention due to its definite advantages of robust structure, high power density, high efficiency, and low cost in long stator application. In this paper, the structure, operation principle, characteristics, and electromagnetic performance of two kinds of the primary-PM motors in the application of the rail transit will be presented. Also, some experiment results are given to validate the study.

Keywords Linear motor doubly salient · Flux-switching · Primary permanent magnet motor · Urban rail transit

4.1 Introduction

Currently, there are two kinds of drive motors are used in urban rail transit, namely rotary motors and linear motors. Comparing to rotary drive method, the linear motor drive method can offer many merits. First, electricity energy is directly converted to the linear motion mechanical energy. Hence, it has the benefits of simple structure and long lifetime. Second, the thrust force of linear motor is independent of the adhesion factor between wheel and rail. Hence, it has good climbing ability. Third, because the wheel is only used to support the car body, so the small wheel diameter is enough, which can reduce the tunnel radius to save construction cost. Furthermore, linear motor drive train can shorten the length of the rail and reduce noise due to the elimination of unnecessary energy conversion from rotary to linear motion.

R. Cao · M. Cheng (✉)
School of Electrical Engineering, Southeast University, Nanjing 210096, China
e-mail: mcheng@seu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_4,
© Springer-Verlag Berlin Heidelberg 2014

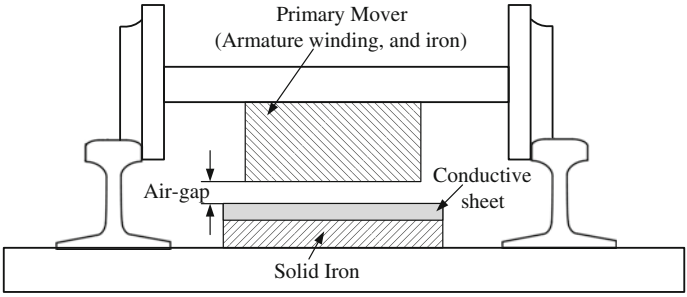


Fig. 4.1 Principle of the linear induction motor in rail transit

Currently, the rail transits driving by linear induction motors have been operated in Japan, Canada, the USA, etc. In China, the first linear motor car line, the Guangzhou Subway Line No. 4, already partly put into operation.

The conventional permanent magnet (PM) linear motor exhibits higher efficiency and higher power factor than induction linear motor [1]. However, in long stator applications such as urban rail transit, this solution inevitably results in significant cost increase due to a large amount of magnets or armature windings set on the long stator. Hence, it has the important theory significance and the practical application value to research a new linear motor with high power density and low system cost.

4.2 Characteristics of Existing Linear Motor for Rail Transit

Currently, there are three kinds of linear motors, namely, linear induction motors, linear permanent magnet motor, and linear switched reluctance motor (LSRM). In this section, the advantages and disadvantages of the three motors in the application of rail transit will be discussed.

4.2.1 Linear Induction Motor

Figure 4.1 shows the principle of the linear induction motor used for railway vehicle, in which the short primary mover is fixed under the bogie of a railway vehicle and the secondary conductive sheet stator is fixed between the iron rails. The operation principle of a linear induction motor is the same with the rotational induction motor. The multiphase winding of the mover produces a traveling electromagnetic field. So, some current will be induced in the conductive sheet, which in turn develops a thrust due to the interaction of the traveling field and induces currents. Currently, linear induction motor for urban rail transit has

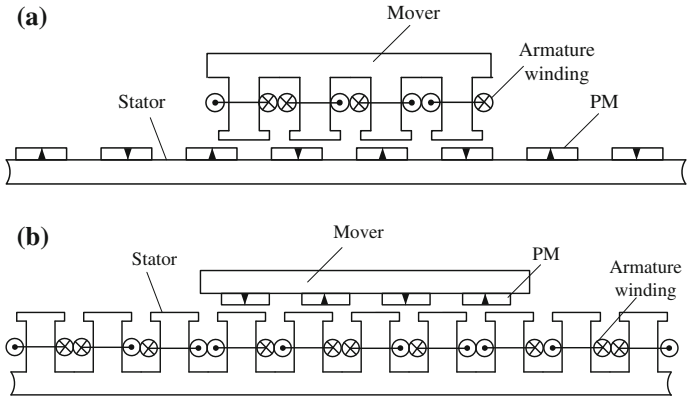


Fig. 4.2 Cross-sections of two linear permanent magnet motors. **a** Short armature winding mover. **b** Short permanent magnet mover

attracted more and more attention due to its simple and inexpensive structure. However, the efficiency and power factor of induction motor are lower due to the large air gap caused by the tolerances for driving dynamics.

4.2.2 Linear Permanent Magnet Synchronous Motor

As we know, the linear permanent magnet synchronous (LPMS) motors exhibit higher efficiency than induction linear motor. If this kind of motor can be used in the rail transit system, the system efficiency will be raised greatly. Generally, linear permanent magnet synchronous (LPMS) motors can be classified into short armature winding mover and short permanent magnet types. For the short armature winding mover LPMS motors, the armature winding mover is set under the train, while a lot of high price permanent magnets are set along the rail. The main drawbacks of this type LPMS motors are high system cost and the inconvenience of maintaining. For the short permanent magnet mover linear motor, permanent magnets are set under the short mover, while the multiphase armature windings are set along the rail. Hence, the LPMS motors inevitably result in significant cost increase due to a large amount of magnets or armature windings set on the stator (Fig. 4.2).

4.2.3 Linear Switched Reluctance Motor

Figure 4.3 shows the cross-section of a linear switched reluctance (LSR) motor. It can be seen that there are no armature winding and permanent magnet on the long secondary stator, which only made up with low price of iron. Hence, it is less

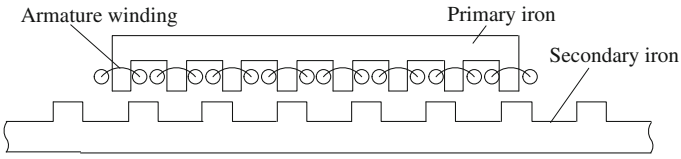


Fig. 4.3 Cross-sections of linear switched reluctance motor

expensive, more robust, and fault tolerant, and has less thermal problems. Also, the armature windings are concentrated, and easier to manufacture and maintain. Because of these merits, LSR motor has been investigated for rail transit in [2]. However, comparing with the LPMFS motors, it suffers from some drawbacks such as high force ripple and low power density.

4.3 Primary Permanent Magnet Linear Motors

In recent years, a new class of PM brushless motors with magnets and contracted armature winding all located in the stator, namely the flux-switching permanent magnet (FSPM) and doubly salient permanent magnet (DSPM) motors [3, 4], have received wide attention due to its high power density, fault-tolerant property [5], robust mechanical integrity, and free from thermal stress problem. Recently, the linear structure of DSPM motor [6] and FSPM motors [7] have attracted much attention, in which both the PMs and armature windings are all located in the short mover, while the long stator only consists of iron core. Hence, these linear FSPM (LFSPM) and linear DSPM (LDSPM) motors are suitable for long stator applications.

4.3.1 Linear Doubly Salient Permanent Magnet Motor

Figure 4.4 shows a modular linear doubly salient permanent magnet (LDSPM) motor, in which its mover consists of two modules whose positions are mutually four and one half of the stator pole pitch (namely 180° electrical degree) apart and there is a flux barrier between the two adjacent modules. In addition, only one PM is inserted in the mover iron of each mover part. The two magnets are magnetized in alternate directions. Similar to a LSRM, concentrated armature windings are adopted, which are wound around the mover teeth of each mover part. Also, each phase winding is composed of four concentrated coils connected in series, e.g., coil A1–A4 for phase A. However, since two mover parts have 180° electrical degree

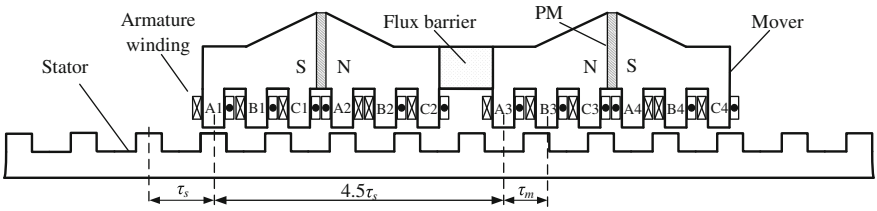


Fig. 4.4 The cross-section of modular LDSPM motor

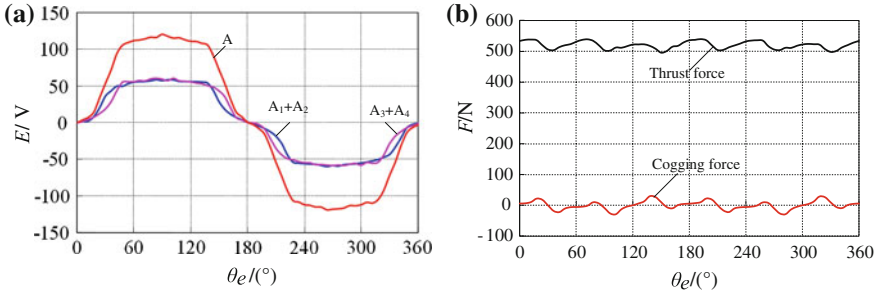


Fig. 4.5 Back EMF and force performance of LDSPM motor. **a** Back EMF of coils and phase winding. **b** Thrust force and cogging force

shift, this LDSPM motor can offer symmetrical back EMF, lower inductance variation range with mover position, bigger average thrust force as well as greatly reduced cogging force and reluctance force ripple especially under BLAC operation. Figure 4.5 shows the three phases back EMF and thrust force waveforms under brushless direct current (BLDC) operation module. It can be seen that the phase back EMF is symmetrical and the thrust force ripple is small. The detailed performance of this motor has been published in [6].

4.3.2 Linear Flux-Switching Permanent Magnet Motor

Figure 4.6 shows one "E"-shaped module of a linear FSPM motor. Each "E"-shaped module consists of two "U"-shaped mover iron, between which a PM is sandwiched. The armature winding coils are located in the slot and wound around the adjacent teeth of the two "U" modules. When the mover moves one stator pole pitch at constant speed, a sinusoidal back EMF will be induced in the armature winding.

A three-phase LFSPM motor with different mover and stator pole pitch τ_m/τ_s can be obtained directly using the following relationships:

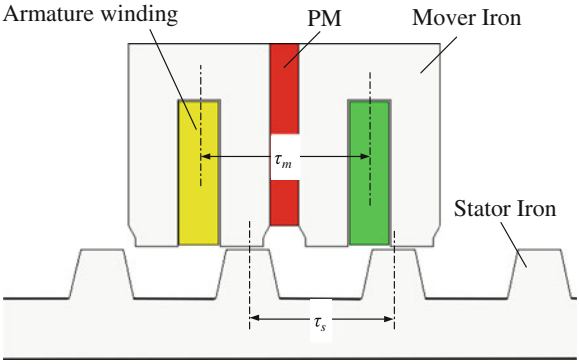


Fig. 4.6 The cross-section of one "E"-*shaped* module

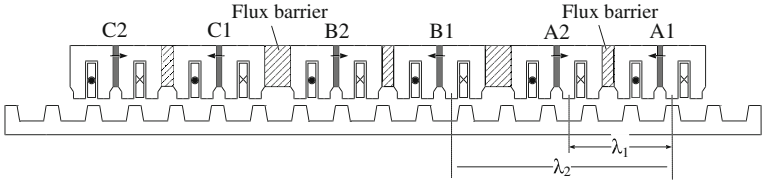


Fig. 4.7 The cross-section of LFSPM motor

- To obtain complementary structure, the relative displacement between the two E-shaped modules of one phase need to satisfy the following relationships:

$$\lambda_1 = (j \pm 1/2)\tau_s \tag{4.1}$$

- For 3-phase motor, the relative displacement between the two E-shaped modules of the adjacent two phases must satisfy:

$$\lambda_2 = (k \pm 1/3)\tau_s \text{ or } \lambda_2 = (k \pm 1/6)\tau_s \tag{4.2}$$

where j and k are all positive integer. Based on (4.1) and (4.2), a three phase LFSPM motor can be obtained as shown in Fig. 4.7. By using finite element analysis, the back EMF, cogging force, and thrust force waveforms are calculated and shown in Fig. 4.8. It can be seen that this LFSPM motor can offer sinusoidal and symmetrical back EMF, small force ripple of thrust force and cogging force. The mathematics modeling, performance of different pole pitch, general design principle of this kind of LFSPM motors have been investigated in [7–9].

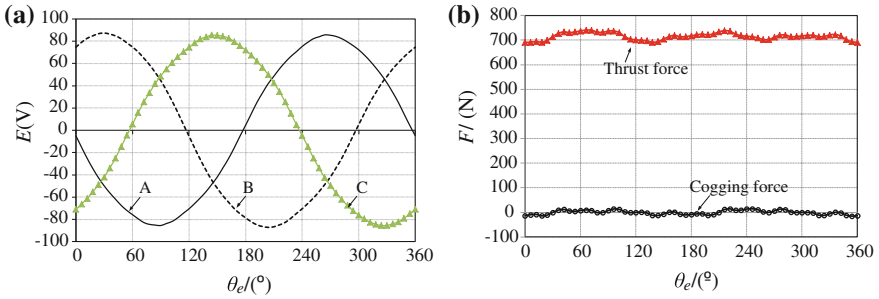


Fig. 4.8 Back-EMF and force waveforms of LFSPM motor. **a** Back-EMF. **b** Thrust force and cogging force

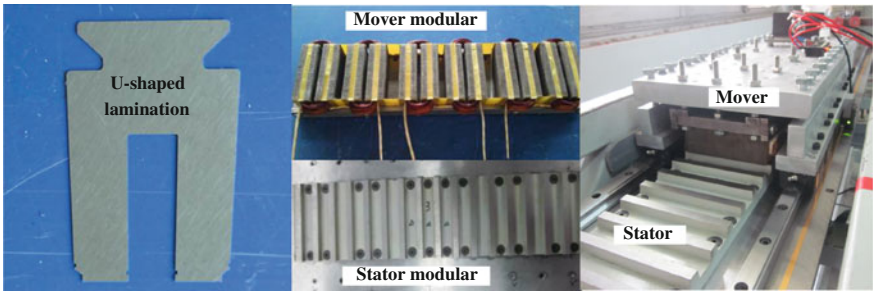


Fig. 4.9 The prototype of a LFSPM motor

4.4 Experiment Research

To validate the associated FEM results of the LFSPM motor as shown in Fig. 4.7, a three-phase motor prototype has been built. The detailed structure including the mover, stator, U-shaped laminated segment, and the prototype motor is shown in Fig. 4.9.

Figure 4.10a shows the responses of speed, current, and position when the motor operating at a step given speed, which can be expressed as "0 m/s-0.3 m/s-0.5 m/s-1 m/s-0 m/s." It can be seen that the response speed is match well with the given speed. The phase current at high speed is bigger than the one at low speed. The reason is that the friction force at high speed is bigger. Figure 4.10b shows the responses of position, current, and position errors when the motor operating at a step given position, which can be expressed as "0 m/s-0.5 m/s-1 m/s-1.5 m/s-2 m/s-0 m/s." It can be seen that the response position is match well with the given position. The position errors are nearly to zero when the motor is stop, while there are some little errors when the given position changed. This is caused by the heavy mover weight.

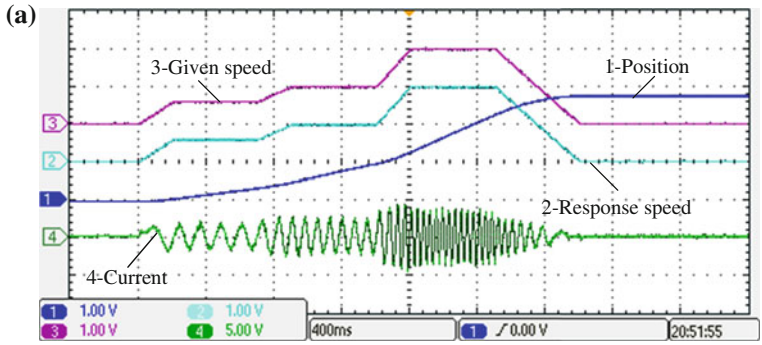


Fig. 4.10 Characteristics of the prototype motor at speed and position control method. **a** Speed control. **b** Position control

4.5 Conclusion

- The primary PM linear motors incorporate the simple stator of LSR motors and high power density of Linear synchronous PM motors, which is very suitable for long stator application to reduce system cost.
- The proposed LFSPM motor and LDSPM motor can offer symmetrical and balanced back EMF, smaller cogging force and smaller thrust force ripple.
- Different from the conventional LFSPM motor, the LFSPM motor has independent three-phase armature windings, and a smaller mutual inductance, thus providing high fault-tolerant ability.
- The proposed two primary PM linear motors have modular mover and stator, which are easy to assemble. Also, it can be a component of a multi machine system.

References

1. Stumberger G, Zarko D, Aydemir MT, Lipo TA (2003) Design and comparison of linear synchronous motor and linear induction motor for electromagnetic aircraft launch system. In: Proceedings of IEMDC2003, Madison, USA, pp 494–500
2. Kolomeitsev L, Kraynov D, Pakhomin S et al (2008) Linear switched reluctance motor as a high efficiency propulsion system for railway vehicles. In: International symposium on power electronics, electrical drives, automation and motion, pp 155–160
3. Liao Y, Liang F, Lipo TA (1992) A novel permanent magnet machine with doubly salient structure. In: Proceedings of IEEE industry applications society annual meeting, pp 308–314
4. Hoang E, Ben-Ahmed AH Lucidarme J (1997) Switching flux permanent magnet polyphased machines. In: Proceedings of European conference on power electronics, pp 903–908
5. Zhao W, Cheng M, Hua W, Jia H, Cao R (2011) Back-EMF harmonic analysis and fault-tolerant control of flux-switching permanent-magnet machine with redundancy. IEEE Trans Ind Electron 58(5):1926–1935
6. Cao R, Cheng M, Mi C, Hua W, Zhao W (2011) A linear doubly salient permanent magnet motor with modular and complementary structure. IEEE Trans Magn 47(12):4809–4821
7. Cao R, Cheng M, Mi C, Hua W, Wang X, Zhao W (2012) Modeling of a complementary and modular linear flux-switching permanent magnet motor. IEEE Trans Energy Convers 27(2):489–497
8. Cao R, Cheng M, Hua W (2013) Investigation and general design principle of a new series of complementary and modular linear FSPM motors. IEEE Trans Ind Electron 60(12):5436–5446
9. Cao R, Cheng M, Mi C, Hua W, Zhao W (2013) Comparison of complementary and modular linear flux-switching motors with different mover and stator pole pitch. IEEE Trans Magn 49(4):1493–1504

Chapter 5
An Maximal Clique Mining Algorithm for Highway Network Optimization Problem

Zipeng Zhang, HongGuo Wang, Yanhui Ding and Zengzhen Shao

Abstract In view of the realistic site and distance information of the highway network and its complex topological structure, a new degree-based maximal clique mining algorithm which containing a series of pruning strategies and dictionary ordering strategies is proposed in this paper according to the top-down method to optimize the expressway network information reasonably. The efficiency of searching and clustering the highway network is improved further. In this paper, a simplified model of the highway network is designed, and based on the example of Shandong province highway network database, the new algorithm shows better results.

Keywords Mining maximal clique · Degree-based algorithm · Highway network optimization

5.1 Introduction

In recent years, with the increasing popularity of highway network in China, the graph mining problem which is based on the topological structure and properties attracts more and more researchers' interest. Thus, the aspects of the reliability, the hierarchical division, the evolutionary pattern, and the capacity of the highway network [1–3] are becoming a new fields of graph mining.

Maximal clique mining problem is a branch of graph mining problem, maximal clique mining problem can be earliest traced back to a theoretical article which has

Z. Zhang (✉) · H. Wang · Y. Ding · Z. Shao
School of Management Science and Engineering, Shandong Normal University,
No. 88 Wenhua Road, JiNan, ShanDong Province, China
e-mail: zhangzipeng20@126.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_5,
© Springer-Verlag Berlin Heidelberg 2014

published by the England scientists Luce and Perry in 1973 [3] Since it was first put forward until recent years, maximal clique mining problem has always been attracting much attention by the researchers in the group mining community. Due to the objects of applications by the mathematic and natural model in the group theory is different, various maximal clique mining algorithms have been put forward by experts in various field in recent years [4–6], such as the social network problems, the Biological information network, the intelligent transportation network, the wireless sensor network problems, the DNA interactive network problems and so on.

At present, scientists have put forward some effective maximal clique algorithm, such as the local search algorithm-MCODE [7] which based on density in a biological network in order to find out the complex organism enginery of protein effectively and used in the field of social networks. K-medoids algorithm [8] for mining network connection structure based on node clustering is proposed in order to enumerate the community associated with the association. Although these algorithms have their own advantages.

The content of this paper put the reality highway traffic network topological structure as a starting point, the maximal mining algorithm is applied to the highway network to simplify the road network and reduce the computational complexity of the objective. Finally makes an empirical study of mining and analysis of the maximum clique which the highway contains.

5.2 The Definition of the Problem

The highway network is a large graph data model which contains a large number of entries and exits of the stations, service stations, and overpass. In the graph-theoretic problems considered, the highway network is defined as a tuple G = (V, E, W), where V denotes the set of entries or exits of the stations and the service stations in the highway network, and E is a set of disorder two-tuples composed of the elements in V, which represents a set of existing highway connected between two stations. W is the weight of edge E, which represents the distance between the station. E[i][j] = 0 says that Vi and V_j is not connected. W[i][j] = 2 points out V_i connected with V_j and the distance between them is 2. Figure 5.1 is a graph theory representation of traffic network information data. Figure 5.2 is an adjacent matrix representation of Fig. 5.1.

In the undirected graph of the highway network G, if the subgroup G_s = (v, e) meets the condition that v ∈ V(G), e ∈ E(G), then G_s is called the subgraph of G. If $C \in V(G), a, b \in C, E[a][b] = 1$, and $E_{ab} \in E(G)$, then C is called complete subgraph of graph G, it is also called clique. A maximal clique is a clique that cannot be extended by adding one more adjacent vertex, that is, a clique is not a subset of any other cliques. For example, in Group2 of Fig. 5.1, the maximal clique is C{a, b, c, d, e} and C{d, e, f}.

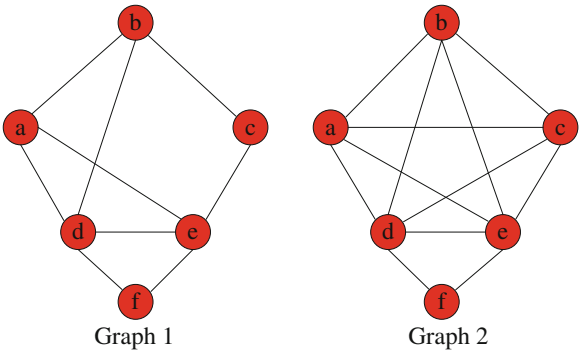


Fig. 5.1 An example for graph database

	a	b	c	d	e	f
a	∞	1	0	1	1	0
b	1	∞	1	1	0	0
c	0	1	∞	0	1	0
d	1	1	0	∞	1	1
e	1	0	1	1	∞	1
f	0	0	0	1	1	∞

Adjacency matrix for Graph1

	a	b	c	d	e	f
a	∞	1	1	1	1	0
b	1	∞	1	1	1	0
c	1	1	∞	1	1	0
d	1	1	1	∞	1	1
e	1	1	1	1	∞	1
f	0	0	0	1	1	∞

Adjacency matrix for Graph2

Fig. 5.2 Adjacency matrix representation of G_1, G_2

The structure of overpass in the expressway network can be divided into T-shaped overpass and cross overpass. On the basis of connectivity between point and point of highway network, the topology structure of the overpass can be turned into the data in the graph theory model. This model is also the starting point of maximal clique mining in highway network. Figures 5.3 and 5.4 show the conversion process mentioned above.

5.3 The Realization of the Algorithm

5.3.1 The Basic Definition of the Degree-Based Algorithm

In this section, we will discuss the algorithm designed for the special topology of the highway network in detail, that is, a maximal clique mining algorithm which is a top-down algorithm, beginning from the vertex with maximum degree and then mining recursively all figure from the undirected graph, that is the degree method. The following is a list of coincidences and abbreviations of the algorithm.

Symbol	Definition
Gs	subgraph
$V(G)$	the set of nodes in G
$E(G)$	the set of edges in G
$N(V_i)$	$N(V_i) = \{V_j \vert E_{ij}\} \in E(G),\ V_i,\ V_j \in\ V(G)$
$N(V_i)'$	$N(V_i)\prime = N(V_i) - V(C_s)$
Di	$D_i = \vert N(V_i)\vert$, the degree of the node V_i
$D(G)$	the set of nodes with degree $= D_i$ in G
$V(G)'$	a set of node V_i by descending Di
$N_L(V_i)$	$N_L(V_i) = \{V_j \vert\ V_j \in N(V_i),\ i < j\}$, $N_L(V_i) \in N(V_i)$
$\Delta(V_i, V_j)$	the set of nodes which can form a triangle with node V_i, V_j
C(G)	the set of all maximal clique in G
P_i	the number of occurrences of the node i in C(G)
$E_{i,j}$	the edge connected by node i and j
$C_i\{i, j, k, m\ldots\ldots\}$	presentation of the maximal clique
$E(U)$	number of edges in maximal clique
$\sum W(E_{i,j})$	weight of edge $E_{i,j}$
$W(U)$	average weight
S	the size of the maximal clique
$V(C)$	the set of the nodes which belong to the maximal clique

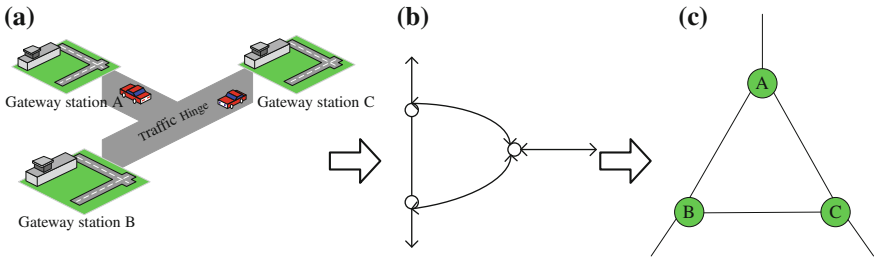


Fig. 5.3 The clique model of T-shaped overpass. **a** T-shaped overpass, **b** Model, **c** Clique

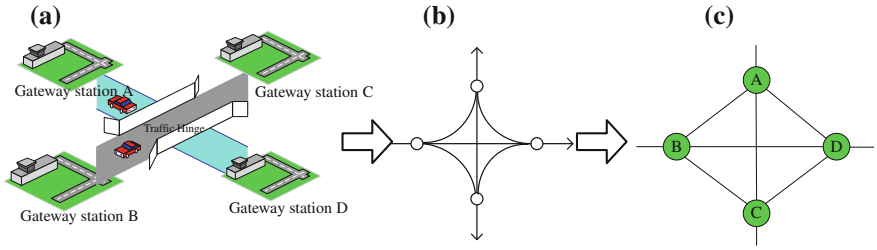


Fig. 5.4 The clique model of cross overpass. **a** Cross overpass, **b** Model, **c** Clique

- *Theorem* 1: For any candidate, sub-clique structure $C_s\{i, j, k\ldots\ldots\}$ in the undirected graph G, if $\forall \boldsymbol{Vi} \in \boldsymbol{C_S}$ and only if there is no nodes $u \in N(Vi) - V(C_S)$ meet the condition $V(C_S) \subset N(u)$, then the structure of the candidate sub-clique C_s is the maximal clique.
- *Theorem* 2: In all candidate sub-clique structure $C_s\{i, j, k,\ldots\ldots\}$, for $\forall Vi \in C_S$, always have $V(C_S) \subset N(u)$.
- *Pruning rule* 1: In the sub-clique structure $C_s\{i, j, k,\ldots\ldots\}$, we will select the node $Vi \in V(C_S)$ with the minimum degree D_{min} to use the decision rule of theorem 1 to determine whether the sub-clique C_s is the maximal clique.
- *Pruning rule* 2: In the process of extending the candidate sub-clique structure $C_s\{i, j, k\ldots\ldots\}$, for the node $\forall u \in N(Vi) - V(C_S)$, we always choose the node with $D_i \geq |V(C_3)|$ to expand the candidate sub-clique structure further.
- *Pruning rule* 3: In the process of mining maximal clique, the nodes in set V(C), contained by maximal clique C(G), cannot be used as the roots of the search trees, that is $V(G)' = V(G)' - C(G)$.
- *Pruning rule* 4: In order to avoid finding the maximum clique in the set $N(V_i)'$, if $|N(V_i)'| \geq 2$, we will go on mining maximum clique.
- *Dictionary order* 1: The node whose degree is higher than 2 in the set $V(G)'$ is in descending order. This dictionary order sets the degree of node in descending order, which can benefit for matching the top-down method of the maximal clique mining.
- *Dictionary order* 2: The nodes in adjacency node set $N(V_i)$ are ordered by ascend according to their node label. Every time, for a given node V_i, the node with minimum label number is selected from the set $N(V_i)$, so the entire search space are organized according to the dictionary order.

5.3.2 The Process of the Degree-Based Algorithm

- Step 1: With the depth first search scanning method, we first calculate the degree D_i of every node V_i in the database. According to the dictionary order 1, after the degree of the nodes in descending order, we select nodes who have the degree $D_i > 2$ to form a new node set $V(G)'$. If $D_i = D_j = D_p = D_q\ldots\ldots$, then the nodes are in descending order according to the node label. Find out the adjacency node set $N(V_i)$ of the node V_i with a maximum degree.
- Step 2: According to the dictionary order 2, remove minimum number in the adjacency node set $N(V_i)$ of the node V_i, and select the node V_j, which meets the condition that $E[k][i] = 1 \wedge E[k][j] = 1$ from the rest node, and then select the node V_k whose node label is minimum and who meets the condition that $E[k][i] = 1 \wedge E[k][j] = 1$, and then V_k can be added to the maximal clique C{i, j, k}. Followed by recursion, we will finish until cannot find the node meets the conditions, we will finally record the maximal clique C{i, j, k, m……}.

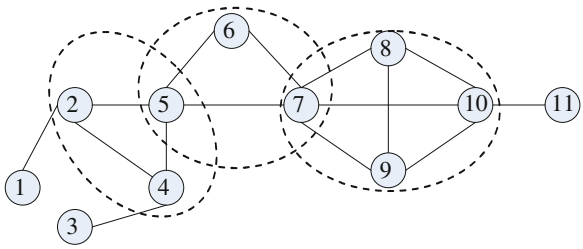


Fig. 5.5 The model of an example algorithm

- Step 3: According to the pruning rule 3, for nodes in the adjacency node set $N(V_i)$ of the node V_i, we remove the nodes contained in the maximal clique to obtain the remaining node set $N(V_i)'$. Then return to step 3, until the rest node sets can no longer form any maximal clique. Now, we can conclude that all cliques which consist of node V_i have been found.
- Step 4: Remove V_i and the set $N(V_i)$ in $V(G)'$, and then select another V_i according to degree D_i in descending order again.
- Step 5: Repeat step 2, step 3, step 4, until the node set $V(G)'$ is Φ. The maximal clique set mined is sequentially numbered, then end.

Figure 5.5 is an example of the algorithm mentioned above, at the end the maximal clique $U_1\{V_5, V_6, V_7\}$, $U_2(V_7, V_8, V_9, V_{10})$, $U_3(V_2, V_4, V_5)$ are found.

5.4 The Algorithm of the Simplification in Highway Network

In order to simplify the highway network by the maximal clique mining algorithm, we should solve the problem of overlap and interaction of the maximal clique. Considering that we propose nonexistent node V' to replace a true one, we see that the average weight of a maximal clique in (U) satisfies the formula (5.1):

$$W(U) = \frac{\sum_{i,j \in U} W\left(E_{i,j}\right)}{E(U)}. \tag{5.1}$$

Function:	make the Simplified Group {G(V,E,W}
Input:	the set of the maximal clique C(G) = {C_i, C_j, C_k, C_m ……}
Output:	make the simplified highway network G′
Step 1:	traverse the set of the date C(G), record P_i as the number of occurrences of V_i(V_i ∈ C(G))
Step 2:	if $P_i = 2$, add V_i' as one of the node V_i, satisfied $V_i \in C_i$, $V_i' \in C_j$ and met the condition that E[i][j] = 1, W[i][j] = 0 If $P_i = 3$, Add V_i' as one of the node V_i, V_i'' as one of the other node V_i, let $V_i \in C_i$, condition that E[i][i′] = 1, E[i][i″] = 1, E[i′][i″] = 1, W[i][i′] = 0, W[i′][i″] = 0, W[i][i″] = 0 //select the node with $P_i \in [2, 3]$, update the maximal clique
Step 3:	calculate the edges of E(U) = s(s−1)/2 which in the maximal clique U_i(i, j, k, m……)
Step 4:	According to the following formula to calculate the average weight of the maximal cliques: $W(U) = \frac{\sum_{i,j \in U} W(E_{i,j})}{E(U)}$
Step 5:	let V_i as U_i(i, j, k, m……), renumber of the set of V_i

Figure 5.6 shows a demonstration model of simplified network algorithm, the nodes with same color representing maximal clique structure. We simplify the 20 nodes of the original network to 10 nodes by our network optimization algorithm. Finally, we will find that the scale of the network is reduced by 50 %, showing the efficiency and necessary of simplifying the network.

5.5 Experimental Results and Analysis

In order to display the effectiveness and performance of maximum clique mining algorithm which is based on the method of degree better, we use 443 site information and 1198 section information of highway road network database in the Shandong province, the experiment is conducted based on the following computer platform: Intel (R) Celeron (R) CPU E3300, 2.50 GHz, dual core, 2G memory, the operating system of Windows XP Professional.

It can be very efficient digging out the maximal clique in the Shandong highway network with the degree-based algorithm. From the Figs. 5.7 and 5.8, we can see that the complexity of the highway and the station in the simplified highway network is more small.

It is a comparison of efficiency between original and simplify network in Fig. 5.9, the first group is the comparison of time in depth first search method, the second one is a comparison between the degree-based algorithm and enumeration method, we can see that the degree of law and the introduction of a variety of pruning strategy lexicographic strategy greatly reducing the size of the algorithm running.

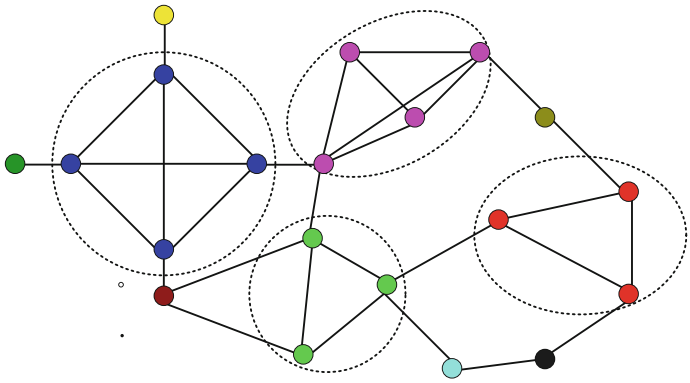


Fig. 5.6 The simplified model of the highway network



Fig. 5.7 The original highway network



Fig. 5.8 The simplified network

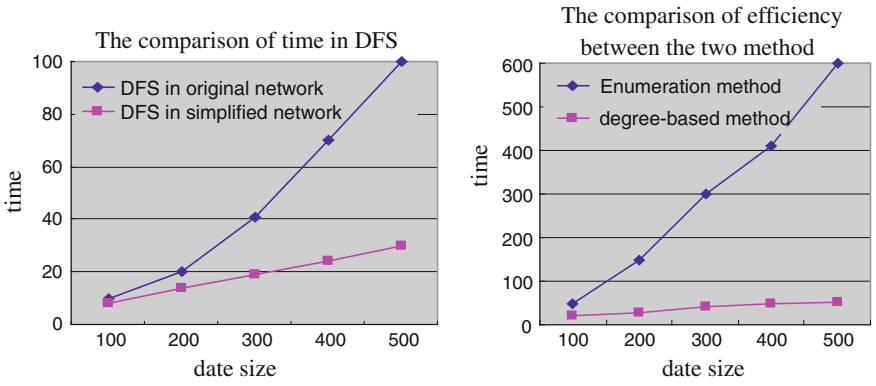


Fig. 5.9 The comparison of efficiency between original and simplify network

5.6 Conclusion

The highway maximum clique mining algorithm based on degree proposed in the paper is effective for practical problem. The algorithm introduces the top-down search method innovatively, iteratively using pruning strategy, at the same time using parallel computing, so as to fulfill the processing of the data from the large-

scale highway network on the basis of reducing the complexity of time and space. Using the maximal clique mining algorithm to simplify the national expressway network, and monitoring highway information in real time, are the main research direction of this field.

References

1. Cheng J, Ke Y, Fu A W-C, Yu JX, Zhu L (2011) Finding maximalcliques in massive networks. ACM Trans Database Syst 36(4):21:1–21:34
2. David E, Darren S (2011) Listing all maximal cliques in large sparse real-world graphs. CoRR abs/1103.0:150–159
3. Bron C, Kerbosch J (1973) Finding all clique of an un directed graph. Commum ACM 16(9):575–577
4. Kose F et al (2001) Visualizing plant metabolomic correlation net works using clique-metabolite matrices. Bioinformatics 17(12):1198–1208
5. Chang L, Yu JF (2013) Fast maximal cliques enumeration in sparse graphs. Algorithmica 66(1):173–186
6. Yan X-F, Han J-W (2002) gSpan: graph-based substructure pattern mining. In: Proceedings of the 2002 IEEE International Conference on Data Mining, IEEE Computer Society, Washington, DC, pp 721–724
7. Eblen JD, Phillips CA, Rogers GL, Langston MA (2012) The maximum clique enumeration problem: algorithms, applications, and implementations. BMC Bioinformatics 13(10):S5
8. Li Z-W, Yang X-H (2013) Dynamically weighted clique evolution model in clique networks. Math Probl Eng pp 778–783

Chapter 6
Design of a 6 kW Battery Charger Based on Full-Bridge Phase-Shifted ZVZCS PWM Converter

Tao Sun, Zhigang Liu, Chao Li, Lei Wang and Lijun Diao

Abstract Based on full-bridge phase-shifted zero voltage and zero current switching (FB-PS-ZVZCS) PWM converter, this paper presents the design procedure of a 6 kW battery charger. In the converter, one blocking capacitor and two diodes are added in the primary side of the transformer to realize ZVS for the leading-leg and ZCS for the lagging-leg. The very detailed analysis of the operation principles is proposed. According to the results, some simplification works are done and the design procedure of the main parameters, including the value of the switches snubber capacitor, the blocking capacitor and the delay time between gates drive signals, also is drown. Finally, a 6 kW battery charger prototype is built. The simulation and experiment results verify the parameters as well as the proposed design procedure.

Keywords Battery charger · ZVZCS · Simplification · Design procedure

6.1 Introduction

Battery chargers are basic components of urban rail transit vehicles. They are typically employed to charge the battery as the characteristic curve and also provide DC power for control circuits, cabin lightening, contactor, magnetic track breaking, and other DC loads in the vehicles [1]. Considering the high input voltage and high power level, in nowadays, battery chargers are mostly designed based on full-bridge DC–DC converter with isolation transformer and choose the IGBTs as switching devices.

T. Sun (✉) · Z. Liu · C. Li · L. Wang · L. Diao
School of Electrical Engineering, Beijing Jiaotong University, 3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 11121660@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_6,
© Springer-Verlag Berlin Heidelberg 2014

To reduce the size and weight of the chargers, the switching frequency of IGBTs is usually required at 20 kHz or above. However, it is difficult to achieve as the current tail will cause higher switching loss when IGBTs turn off. In order to overcome this problem, a ZVZCS technique, which means ZVS operation for leading-lag switches and ZCS operation for lagging-lag switches, has been introduced. The ZVS operation can be implemented by the specific order of the phase-shifted gate drive signals, while the ZCS operation can be implemented by resetting the primary current during the freewheeling period. In the past few years, a series of FB-PS-ZVZCS PWM converters have been proposed and the primary current is mainly reset in these different ways: (a) adding a blocking capacitor and a saturable inductor in the primary side [2]; (b) adding secondary active clamp [3]; (c) adding energy recovery snubber include two capacitors and three diodes in the secondary side [4]; (d) using transformer auxiliary winding [5]; (e) using a simple auxiliary circuit include one capacitor and two diodes in the secondary side [6]; (f) using coupled output inductor [7]; (g) adding two diodes in series with the lagging-leg [8]. Among all the topologies, the last one is the most attractive since it is simple and easy to implement. The blocking capacitor is placed in series with the primary winding of the transformer making the primary current reduced to zero during the freewheeling period. That will ensure the ZCS operation for the switches of the lagging-leg. Then the diodes in the lagging-leg are used to prevent the current from reversing. The converter has good ZVZCS characteristics and analyses well in Ruan and Yan [8].

This paper presents the design procedure of a 6 kW battery charger based on the FB-PS-ZVZCS PWM converter using two diodes in series with the lagging-leg. The very detailed analysis of the operation principles, the equivalent simplification, and the design procedure of the main parameters are proposed. After that, a battery charger prototype is built. The simulation and experiment results are also included to verify the converter design.

6.2 Operation Principles

The main topology of the battery charger is presented in Fig. 6.1. During a whole switching period, there are 12 operation modes in the steady state and Fig. 6.2 shows the waveforms of the key parameters, such as the primary current and gate drive signals, etc. In order to simplify the analysis, the following assumptions are made:

(1) All the switches, diodes, capacitors, inductors are ideal;
(2) $C_1 = C_3$, $C_1 + C_3 = C_{eq}$;
(3) The input filter capacitor C_{in} is large enough to be treated as the input power supply and its voltage V_{in} is considered unchanged during a switching period;

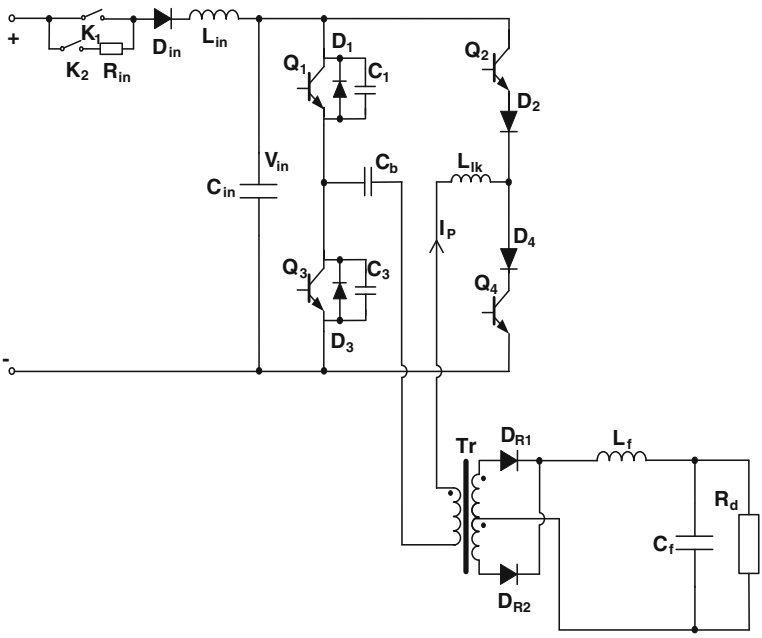


Fig. 6.1 The main topology of the battery charger

(4) The output filter inductor L_f and capacitor C_f are large enough that the output current I_o and output voltage V_o are treated as constant during a switching period.

Mode 0 [t_0]: At the t_0 moment, Q_1 and Q_4 are conducting and all the current flows through D_{R1}. The energy is transferred from the input to the output. At this time, the primary current is $I_p(t_0) = I_o/K$, where K is the turns radio of the transformer.

Mode 1 [$t_0 \sim t_1$]: Q_1 is turned off at t_0, changing the primary current flow path from Q_1 to C_1 and C_3. As a result, C_1 is charged and C_3 is discharged, which means the voltage across Q_1 increases slowly. So Q_1 is turned off with ZVS. Since the transformer leakage inductance L_{lk} and output filter inductor L_f are in series relationship during mode 1 and L_f is large enough, the primary current is considered constant approximately. Therefore, the voltage of blocking capacitor V_{C_b} increases linearly and V_{ab} decreases linearly at the same time.

$$V_{C_b}(t) = V_{C_b}(t_0) + \frac{I_o}{KC_b}(t - t_0) \tag{6.1}$$

$$V_{C_1}(t) = \frac{I_o}{KC_{\text{eq}}}(t - t_0) \tag{6.2}$$

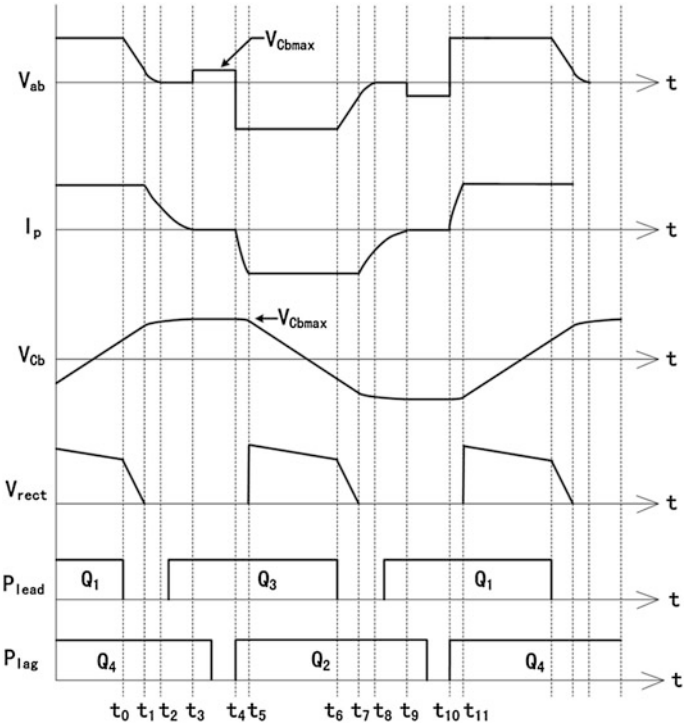


Fig. 6.2 The operation waveforms of the key parameters

$$V_{\mathrm{ab}}(t) = V_{C_3}(t) = V_{\mathrm{in}} - \frac{I_o}{KC_{\mathrm{eq}}}(t - t_0) \tag{6.3}$$

At t_1 moment, the voltage of C_3 (same with V_{ab}) is equal to V_{C_b} and mode 1 is over. The lasting time is:

$$t_{01} = \frac{KC_{\mathrm{eq}}(V_{\mathrm{in}} - V_{\mathrm{C_b}}(t_1))}{I_o} \tag{6.4}$$

And V_{C_b} is:

$$V_{C_b}(t_1) = V_{C_b}(t_0) + \frac{I_o}{KC_b} t_{01} \tag{6.5}$$

Mode 2 [$t_1 \sim t_2$]: At t_1, the voltage of the transformer primary winding becomes zero and the next moment it will be negative. The voltage direction of the transformer secondary winding is flipped too. $\mathrm{D_{R1}}$ and $\mathrm{D_{R2}}$ are conducting at the same time and both the voltages of the transformer primary and secondary winding are clamped to zero. Therefore, C_1, C_3, C_b and L_{lk} start to resonate. At the moment of t_2, V_{ab} is zero and so is the voltage of C_3. The anti-parallel diode of Q_3 conducts naturally and mode 2 ends.

$$V_{\mathrm{ab}}(t) = \frac{I_o}{K}\left(\frac{1}{wC_b} - wL_{\mathrm{lk}}\right)\sin[w(t-t_1)] + V_{C_b}(t_1) \tag{6.6}$$

$$V_{C_b}(t) = \frac{I_o}{KwC_b}\sin[w(t-t_1)] + V_{C_b}(t_1) \tag{6.7}$$

$$I_p(t) = \frac{I_o}{K}\cos[w(t-t_1)] \tag{6.8}$$

where $w = \sqrt{\frac{C_{\mathrm{eq}}+C_b}{L_{\mathrm{lk}}C_{\mathrm{eq}}C_b}}$.

Mode 3 [$t_2 \sim t_3$]: After D_3 turns on, the flow path of I_p is changed from C_1 and C_3 to D_3. L_{lk} and C_b continue to resonate and the primary current is reduced to zero at t_3. If Q_3 is turned on during the mode 3 time, it is obviously under ZVS condition.

$$V_{C_b}(t) = wL_{\mathrm{lk}}I_p(t_2)\sin[w(t-t_2)] + V_{C_b}(t_2)\cos[w(t-t_2)] \tag{6.9}$$

$$I_p(t) = I_p(t_2)\cos[w(t-t_2)] + \frac{V_{C_b}(t_2)}{wL_{\mathrm{lk}}}\sin[w(t-t_2)] \tag{6.10}$$

where $w = \frac{1}{\sqrt{L_{\mathrm{lk}}C_b}}$.

Mode 4 [$t_3 \sim t_4$]: Because of the diodes in series with the lagging-leg, the primary current cannot be reversed. V_{Cb} will remain unchanged and both of the rectifier diodes in the secondary side are conducting. If Q_4 is turned off at this time, as there is no current, Q_4 is turned off with ZCS.

Mode 5 [$t_4 \sim t_5$]: Q_2 is turned on at t_4. L_{lk}, C_b and V_{in} start to resonate and the primary current increases reversely. Both of the rectifier diodes still conduct. Q_2 is turned on approximately with ZCS. When the primary current reaches the load current at t_5, the voltage direction of the transformer winding is reversed. D_{R1} turns off and the entire load current flows through D_{R2}. Then Mode 5 is over.

$$I_p(t) = wC_b[V_{\mathrm{in}} + V_{C_b}(t_3)]\sin[w(t-t_4)] \tag{6.11}$$

$$V_{C_b}(t) = [V_{\mathrm{in}} + V_{C_b}(t_3)]\cos[w(t-t_4)] - V_{\mathrm{in}} \tag{6.12}$$

where $w = \frac{1}{\sqrt{L_{\mathrm{lk}}C_b}}$.

Mode 6 [$t_5 \sim t_6$]: From the time of t_5, the energy is delivered form input to output continuously. Since the I_o is considered constant, V_{C_b} will increases linearly negative.

$$V_{C_b}(t) = V_{C_b}(t_5) - \frac{I_o}{KC_b}(t-t_5) \tag{6.13}$$

Q_3 is turned off at t_6. The first half of the switching period is over and the other haft-cycle is similar with the first one.

From the analysis above, it's easy to know the operation principles of the main topology are quite complicated. The trigonometric formulas make it difficult to solve the equations and that's the reason why some simplification work should be

done. Considering the resonate time is very short and C_b is usually several uF, it is acceptable to treat V_{C_b} as constant approximately during the freewheeling period (t_1–t_5) and the value is kept as $V_{Cb\max}$. The formulas of each operation mode are simplified as follows:

Mode 1 [$t_0 \sim t_1$]: Eqs. (6.1), (6.2), (6.3), (6.5) and

$$t_{01} = \frac{KC_{\text{eq}}(V_{\text{in}} - V_{Cb\max})}{I_o} \tag{6.14}$$

Mode 2 [$t_1 \sim t_2$]:

$$V_{\text{ab}}(t) = V_{cb\max} - \frac{I_o}{KwC_{\text{eq}}}\sin[w(t - t_1)] \tag{6.15}$$

$$I_p(t) = \frac{I_o}{K}\cos[w(t - t_1)] \tag{6.16}$$

$$t_{12} = \frac{1}{\text{w}}\arcsin(\frac{V_{cb\max}KwC_{\text{eq}}}{I_o}) \tag{6.17}$$

Based on the operation principles, V_{ab} need decrease to zero first and after that, I_p flows through D_3 and makes it turn on naturally which creates the ZVS condition for Q_3. So there is:

$$\frac{V_{cb\max}KwC_{\text{eq}}}{I_o} < 1 \tag{6.18}$$

Mode 3 [$t_2 \sim t_3$]:

$$I_p(t) = I_p(t_2) - \frac{V_{cb\max}}{L_{\text{lk}}}(t - t_2) \tag{6.19}$$

$$t_{23} = \frac{I_p(t_2)L_{\text{lk}}}{V_{cb\max}} \tag{6.20}$$

Mode 5 [$t_4 \sim t_5$]:

$$I_p(t) = -\frac{V_{\text{in}} + V_{Cb\max}}{L_{\text{lk}}}(t - t_4) \tag{6.21}$$

$$t_{45} = \frac{I_o L_{\text{lk}}}{K(V_{\text{in}} + V_{cb\max})} \tag{6.22}$$

Mode 6 [$\text{t}_5 \sim \text{t}_6$]:

$$V_{Cb}(t) = V_{Cb\max} - \frac{I_o}{KC_b}(t - t_5) \tag{6.23}$$

6.3 ZVZCS Inequalities

In this part, some ZVZCS inequalities are presented. They are very helpful and necessary for the design procedure.

(1) The maximum value of V_{Cb}

It's discussed in Ruan and Yan [8] and the result is:

$$V_{Cb\max} \approx \frac{I_o}{2KC_b} t_{56} \approx \frac{I_o T_s D}{4KC_b} \tag{6.24}$$

In order to be close with the exact value and deduce the deviation, C_b should be as large as possible within the acceptable range.

(2) The ZVZCS conditions

For the ZVS operation of leading-lag IGBTs, Eqs. (6.18) and (6.25) should be satisfied. The former one ensures V_{ab} decreases to zero before I_p in Mode 2 and Mode 3 will exist. The posterior one guarantees the IGBTs are turned on accurately in Mode 3 by setting up the dead time (t_{dlead}).

$$t_{01} + t_{12} \le t_{\text{dlead}} \le t_{01} + t_{12} + t_{23} \tag{6.25}$$

Take t_{01}, t_{12}, t_{23} into 6.25 and it is seen that the expression is complicated. It's difficult to get the value of t_{dlead} by Eq. (6.25). Therefore, the inequality can be approximated as:

$$t_{01} + t_{12} = \frac{KC_{eq}(V_{in} - V_{cb\max})}{I_o} + \frac{1}{w}\arcsin\left(\frac{V_{cb\max} K w C_{eq}}{I_o}\right) \tag{6.26}$$

Take Eq. (6.24) into Eq. (6.26):

$$t_{01} + t_{12} \approx \frac{KC_{\text{eq}} V_{\text{in}}}{I_o} - \frac{C_{eq} T_s D}{4C_b} + \frac{1}{w}\arcsin\left(\frac{w C_{\text{eq}} T_s D}{4C_b}\right) \tag{6.27}$$

In general, $C_{\text{eq}} \ll C_b$ and Eq. (6.27) can be simplified as:

$$t_{01} + t_{12} \approx \frac{KC_{\text{eq}} V_{\text{in}}}{I_o} \tag{6.28}$$

Analysis Eq. (6.28) and it can be known if Mode 1 is extended to the end of Mode 2, then $t_{01} + t_{12}$ is exactly equal to Eq. (6.28). So Mode 1 and Mode 2 can be merged as Mode 1' and the new formulas are:

Mode 1′ [t_0–t_2]: Eqs. (6.1)–(6.3) are unchanged and the time of Mode 1′ is:

$$t_1 = \frac{KC_{\text{eq}} V_{\text{in}}}{I_o} \tag{6.29}$$

Mode 2′ [t_2–t_3]:

$$I_p(t) = \frac{I_o}{K} - \frac{V_{cb\max}}{L_{lk}}(t - t_2) \tag{6.30}$$

$$t_2 = \frac{I_o L_{\mathrm{lk}}}{K V_{cb\max}} \tag{6.31}$$

After the simplification above, Eq. (6.25) is expressed as:

$$t_1 \le t_{\mathrm{dlead}} \le t_1 + t_2 \tag{6.32}$$

To ensure the ZCS operation of lagging-lag IGBTs, the following inequalities should be required:

$$t_{\mathrm{dlag}} > t_{\mathrm{off}} \tag{6.33}$$

$$D \le D_{\max} \tag{6.34}$$

Equation (6.33) means the dead time of lagging-leg must be longer than t_{off}. Otherwise, the short-circuit fault may happen. Equation (6.34) indicates the maximum of the duty cycle. It is discussed well in [8] and $D_{\max}$ can be expressed as:

$$D_{\max} \approx 1 - D_2 - D_{\mathrm{zcs}} \tag{6.35}$$

where $D_2 \approx \frac{2I_o L_{\mathrm{lk}}}{K V_{cb\max} T_s}$ and$D_{\mathrm{ZCS}} = \frac{2t_{\mathrm{dlag}}}{T_s}$.

6.4 Design Procedure

The main technical indexes of the converter are given in Table 6.1, by which the turns radio of the transformer can be determined. As a result, K is 19.5:1:1 and the duty cycle range is 0.58–0.62. Design and manufacture the transformer as the parameters above and the leakage inductor measured is 55 uH.

A. The blocking capacitor C_b

Based on Eq. (6.24), the value of C_b can be designed to keep V_{Cbmax} around 10 % of V_{in} [8]. Therefore, C_b is calculated as:

$$0.1V_{\mathrm{in}} = V_{\mathrm{Cbmax}} \approx \frac{I_o T_s D}{4KC_b} \tag{6.36}$$

To meet the requirement of V_{Cbmax}, D should be the maximum value 0.62. V_{in} can be 780 V and I_{O} is 250 A. Then

$$C_b \approx \frac{I_o T_s D}{0.4KV_{\mathrm{in}}} = \frac{250 \times 50 \times 10^{-6} \times 0.62}{0.4 \times 19.5 \times 780} \approx 1.3\ \mathrm{uF} \tag{6.37}$$

Table 6.1 Technical indexes of the converter

Rated input voltage	DC780 V
Input voltage range	DC750 V–DC800 V
Rated output voltage	DC24 V
Highest output current	250 A
Switching frequency	20 kHz

Limited by the experiment conditions, we choose $C_b = 0.5\,\text{uF}$ and V_{Cbmax} is

$$V_{cb\max} \approx \frac{I_o T_s D}{4KC_b} = \frac{250 \times 50 \times 10^{-6} \times 0.62}{4 \times 19.5 \times 0.5 \times 10^{-6}} \approx 200\,\text{V} \tag{6.38}$$

B. The snubber capacitor C_{eq}

Combined with Eq. (6.24) and (6.18) can be expressed as

$$C_{\text{eq}} < \frac{16C_b^2 L_{\text{lk}}}{T_s^2 D^2} \tag{6.39}$$

In order to make C_{eq} satisfy Eq. (6.39) at all time, D should be the maximum value to minimize the right part of the inequality. So the range of C_{eq} is

$$C_{\text{eq}} < \frac{16 \times 0.25 \times 10^{-12} \times 55 \times 10^{-6}}{2500 \times 10^{-12} \times 0.62^2} \approx 229\,\text{nF} \tag{6.40}$$

As we can see, the range is large. In fact, C_{eq} is usually 5–30 nF in engineering, because smaller C_{eq} is, the formulas simplified before are more accurate. However, C_{eq} cannot be too small at the same time as that will weaken the ZVS effect when IGBTs are turned off in the leading-leg. So after careful consideration, we choose $C_{\text{eq}} = 12\,\text{nF}$ at last.

C. The dead time t_{dlead} and t_{dlag}

Based on Eq. (6.32), the range of t_1 should be calculated first. It is easy to get the minimum value of t_1 as follows:

$$t_{1\min} = \frac{KC_{\text{eq}} V_{\text{in}}}{I_{\text{omax}}} = \frac{19.5 \times 12 \times 10^{-9} \times 800}{250} \approx 0.75\,\text{us} \tag{6.41}$$

With the decrease of I_{o}, t_1 will approach infinity. Obviously t_{dlead} will not satisfy Eq. (6.32) at that time. So we need set up a ZVZCS limit load and once I_{o} is smaller than the load, the IGBTs are turned on while the snubber capacitors in parallel are not discharged entirely. Then the IGBTs become hard-switching. The range can't be too large (I_{o} is small and the deviation caused by approximation will increase) and too small (that will debase the ZVZCS performance of the converter). Therefore, we select the ZVZCS limit load initially as 20 % I_{omax} and the t_{dlead} is

$$t_{\text{dlead}} = \frac{KC_{\text{eq}}V_{\text{in}}}{I_{\text{o}}'} = \frac{19.5 \times 12 \times 10^{-9} \times 800}{0.2 \times 250} \approx 3.75\,\text{us} \tag{6.42}$$

As a result, within the range of the ZVZCS load, the left part of Eq. (6.32) will be satisfied.

Take Eq. (6.24) into Eq. (6.31) and t_2 is expressed as

$$t_2 = \frac{4C_bL_{\text{lk}}}{T_sD} \tag{6.43}$$

From Eq. (6.43) we can know, the variation regular of t_2 is the same with t_1. Both of them reach the minimum value when output current is 250 A and $t_{1\text{min}} + t_{2\text{min}}$ is calculated as follows:

$$t_{2\text{min}} = \frac{4C_bL_{\text{lk}}}{T_sD_{\text{min}}} = \frac{4 \times 0.5 \times 10^{-6} \times 55 \times 10^{-6}}{50 \times 10^{-6} \times 0.62} \approx 3.55\,\text{us} \tag{6.44}$$

$$t_{1\text{min}} + t_{2\text{min}} = 0.75 + 3.55 = 4.3\text{us} \tag{6.45}$$

In order to meet the requirement of the right part of Eq.(6.32), t_{dlead} should be smaller than $t_{1\text{min}} + t_{2\text{min}}$. Otherwise, the primary current will decrease to zero and the IGBTs of leading-leg are still not turned on. The snubber capacitor in parallel is going to be charged again which means the ZVS condition will be lost.

Based on the analysis above, we choose $t_{\text{dlead}} = 3.75\,\text{us}$. Considering the generation principles of gate drive signals, t_{dlead} is approximately set as 3.5 us at last.

Compared with t_{dlead}, t_{dlag} can be a smaller value so that D_{max} will increase based on Eq. (6.35). That is favorable to the converter and we choose $t_{\text{dlag}} = 2\,\text{us}$ which satisfies Eq. (6.34) at the same time.

6.5 Simulation and Experiment Results

Following the design procedure above, a 6 kW battery charger model is proposed and simulated in Matlab/Simulink. The following parameters are used: $V_{\text{in}} = \text{DC780 V}$, $L_{\text{in}} = 1\text{mH}$, $C_{\text{in}} = 1360\ \text{uF}$, $C_{\text{eq}} = 12\,\text{nF}$, $C_b = 0.5\,\text{uF}$, $L_{lk} = 55\,\text{uH}$, $K = 19.5:1:1$, $L_f = 12\,\text{uH}$, $C_f = 13600\,\text{uF}$, $t_{\text{dlead}} = 3.5\,\text{us}$ and $t_{\text{dlag}} = 2\,\text{us}$. The simulation results are shown in Fig. 6.3 and the experiment results are shown in Figs. 6.4 and 6.5.

Figure 6.3 shows the simulation waveforms at 100 % load are exactly the same with the theoretical analysis. There are six apparent operation modes during the half cycle and the deviation caused by approximation is quite small. Figures 6.4 and 6.5 shows the actual experiment waveforms and we can see the IGBTs of the leading-leg are turned on when I_p is not zero and the IGBTs of the lagging-leg are turned off after I_p reaches zero. Therefore, the design procedure presented is rational and the battery charger can achieve desired ZVZCS operation

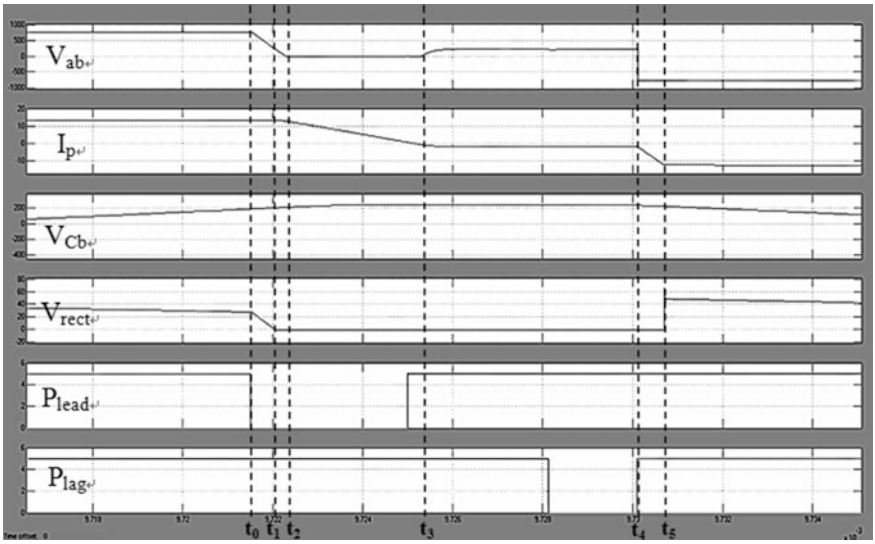


Fig. 6.3 The operation waveforms at 100 % load

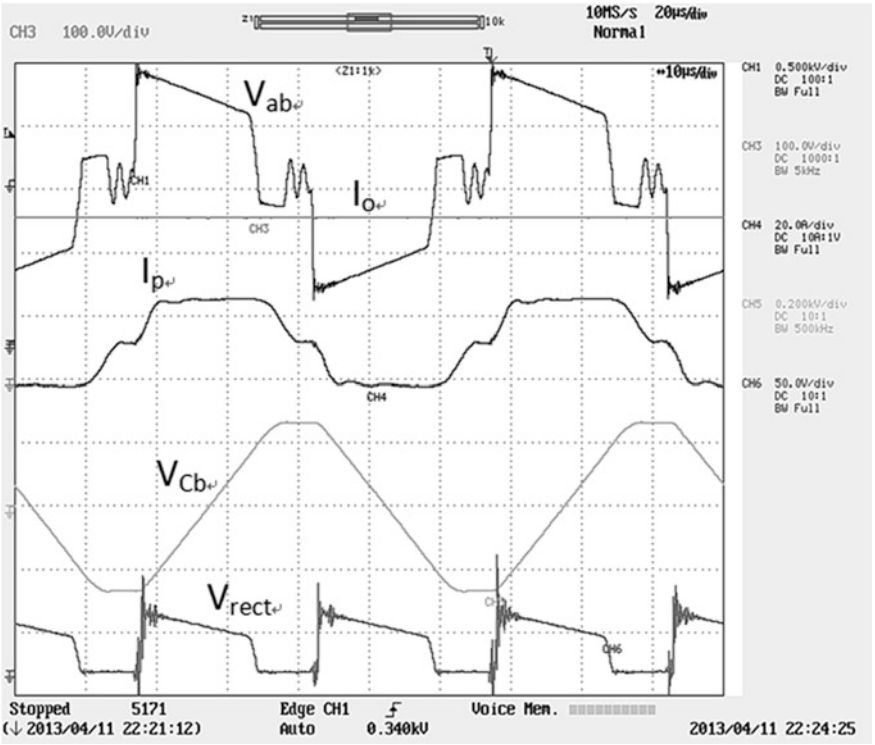


Fig. 6.4 The experiment results at 100 % load

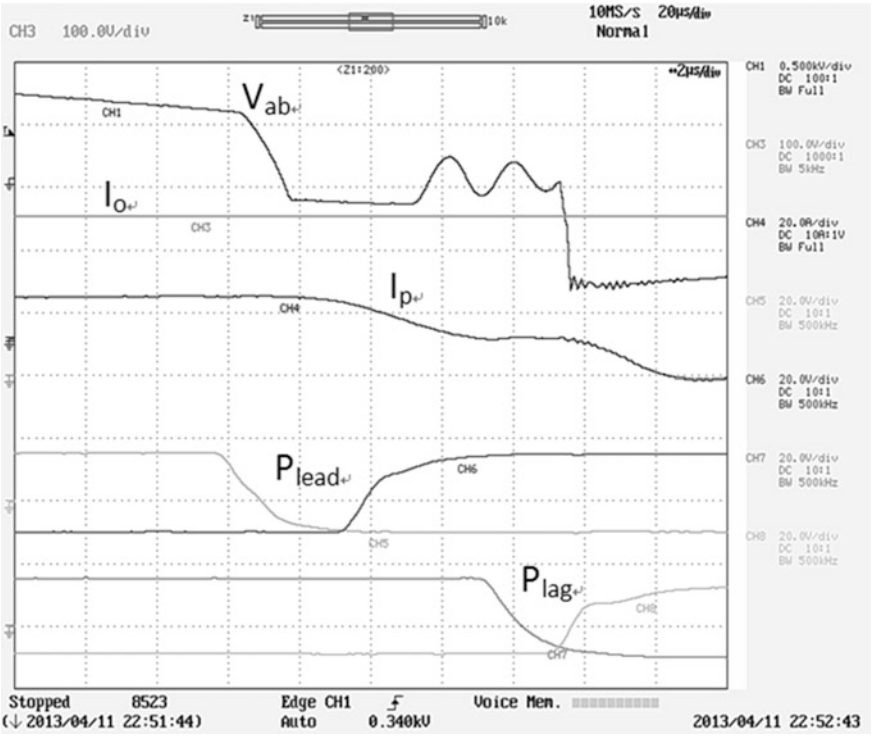


Fig. 6.5 The experiment results with gate drive signals at 100 % load

characteristics well with satisfying all the technical indexes. The efficiency is 94 % at full load.

6.6 Conclusions

This paper presents the design procedure of a 6 kW battery charger based on the FB-PS-ZVZCS PWM converter using two diodes in series with the lagging-leg. The leading-leg can achieve ZVS within a wide load range and the lagging-leg can achieve ZCS at any load. The proposed design procedure is easy and reliable. The parameters designed are certified suitable by simulation and experiment.

Acknowledgments This work was supported in part by the China National Science and Technology Support Program under Grant (2013BAG21QB00).

References

1. Wenli L, Ying M (2008) Design of a soft switched 8 kW battery charging converter for 100 % low floor light rail vehicle. In: Vehicle power and propulsion conference, pp 1–5
2. Cho JG, Sabate JA, Hua G, Lee FC (1994) Zero-voltage and zero-current-switching full bridge PWM converter for high power applications. In: Power electronics specialists conference, PESC '94 record, 25th annual IEEE, vol 1. pp 102–108
3. Cho JG, Rim GH (1996) Zero voltage and zero current switching full bridge PWM converter using secondary active clamp. In: Power electronics specialists conference, vol 1, pp 657–663
4. Kim ES, Joe KY, Kye MH (1997) An improved ZVZCS PWM FB DC/DC converter using energy recovery snubber. In: Applied power electronics conference and exposition, vol 2, pp 1014–1019
5. Cho JG, Baek JW, Yoo DW, Lee HS, Rim GH (1997) Novel zero-voltage and zero-current-switching (ZVZCS) full bridge PWM converter using transformer auxiliary winding. In: Power electronics specialists conference, vol 1, pp 227–232
6. Cho JG, Baek JW, Jeong CY, Yoo DW, Lee HS, Rim GH (1998) Novel zero-voltage and zero-current-switching (ZVZCS) full bridge PWM converter using a simple auxiliary circuit. In: Applied power electronics conference and exposition, vol 2, pp 834–839
7. Choi HS, Kim JW, Cho BH (2001) Novel zero-voltage and zero-current-switching (ZVZCS) full-bridge PWM converter using coupled output inductor. In: Applied power electronics conference and exposition, vol 2, pp 967–973
8. Xinbo R, Yangguang Y (2001) A novel zero-voltage and zero-current-switching PWM full-bridge converter using two diodes in series with the lagging leg. IEEE Trans Industr Electron 48(4):777–785

Chapter 7
Diagnostics of Transformer Windings Deformation Based on Transfer Function

Lei Yin, Zhensheng Wu and Junfeng Gui

Abstract The deformation of transformer windings after a short circuit was analyzed by studying the criterion for the transfer function. First, the lumped parameter method was used to establish the equivalent circuit of transformer winding, the losses, and mutual inductance was added in it. Second, a single-layer continuous transformer winding was used as an example to calculate the changes in equivalent circuit parameters due to the changes of the structure and geometry size of the transformer winding, when the transformer windings occurred different forms and degrees deformation. Finally, the common axial and radial deformations of transformer winding was studied to calculate the transfer function of the deformed winding, with the degree of winding deformation judged by the mean square root of the spectrum difference, a diagnostic criterion was given to judge the deformation degree.

Keywords Transformer · Winding deformation · Transfer function · Deformation diagnostic

7.1 Introduction

Power transformer is an important part of power system, it will cause significant economic losses and even serious consequences in the event of failure. Transformers in power system often suffer the impact of short-circuit current. The winding may be deformed by huge electrodynamic stress. Therefore, the test and

L. Yin (✉) · Z. Wu · J. Gui
Beijing Jiaotong University, Beijing, China
e-mail: 11121704@bjtu.edu.cn

Z. Wu
e-mail: zhshwu@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_7,
© Springer-Verlag Berlin Heidelberg 2014

diagnostic of winding deformation is greatly valuable for the safe operation of the transformers.

The common test methods of transformer winding deformation are Frequency Response Analysis and Low Voltage Impulse. However, these methods has been greatly limited in the practical application due to lack of testing and judgment standard and the influence of field test environment [1]. Transfer function method has been in practice for several years [2–5]. Winding Deformation diagnosis using transfer function has the advantages of high sensitivity and can recognize the winding deformation type and degree of judgment.

7.2 The Establishment of the Transformer Winding Model

7.2.1 The Equivalent Model of Transformer Winding

According to the electro technique theory [6], all of these parameters rely on dimensions and positions of conductor in the transformer winding [7]. It can be suitable for different winding structure types and can reflect the pulse propagation in the winding process accurately [8]. The lumped-element model of the transformer winding is constructed as below: the winding is divided into many cells, and each cells includes several discs. There is mutual inductance between each two cells and grounding capacitance at the ends of a unit. A single-layer continuous transformer winding was used as an example to construct the equivalent circuit model [9], the equivalent circuit model of the transformer winding with the single cake as a unit as shown in Fig. 7.1.

Where, C_{20}, C_{40}, … is the grounding capacitance, C_{12}, $C_{23,}$ $C_{34,}$ … is the equivalent capacitance of interturn; $C_{13,}$ $C_{24,}$ $C_{35,}$ … is the equivalent capacitance of interdisc; L_{11}, $L_{22,}$ … is the self-inductance. There is mutual inductance between each two cells.

7.2.2 The Determination of Simulation Model Parameters

Because of the distance between the turn is far less than turn higher and the distance between the disc is far less than its radial width, the geometric capacitance of interturn and interdisc can be estimated according to the plate capacitor formula [10], namely

$$C_g = \frac{\pi \varepsilon_0 \varepsilon_p B d}{\varepsilon_p} \tag{7.1}$$

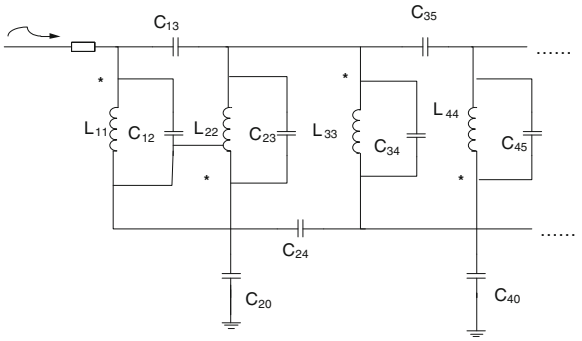


Fig. 7.1 Equivalent circuit model of the transformer winding with the single cake as a unit

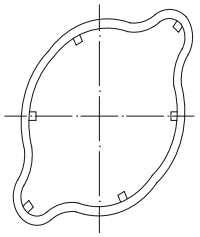


Fig. 7.2 Winding deformation under radial electrodynamic force

where, ε_p, the equivalent dielectric coefficient of the combination insulation; B, disc radial width or turn higher; d, the average diameter of disc or turn; δ_p, combination insulation spacing.

7.2.3 Parameter Calculation of the Deformed Winding

When the winding suffers radial direction electrodynamic stress, compression stress acts on the inner winding. This form of deformation is called radial distortion as shown in Fig. 7.2. For the parameters in the equivalent circuit, this deformation mostly influences the geometric capacitance of interturn and interdisc. The capacitance can be estimated by:

$$C = \frac{\varepsilon S}{d} \tag{7.2}$$

where S is the area and d is the space between the effective electrode.

On the assumption that the shape of disc-cell keeps and its position changes, parameters in the circuit-unit are no change and relevant parameters among adjacent units are recalculated by:

$$C_d = \frac{2\pi\varepsilon L}{\ln\left(\frac{r_1}{r_2}\right)} \tag{7.3}$$

where r_1 and r_2 are the outer and inner diameter of the coaxial electrodes, L is the height, ε is the equivalent dielectric coefficient of the combined insulation in the transformer.

7.3 Transfer Function Method

7.3.1 Theory

Transformer winding itself can be regard as a passive four-terminal components. The transfer function only depends on the structure of the system and components, it has nothing to do with the external action. At one end of the transformer winding exert an excitation signal, measuring at the other end of the response. Using Fast Fourier Transform (FFT) and turn them into frequency domain function, transfer function is obtained.

Define the transfer voltage transfer function for power transformer:

$$H\left(e^{jw}\right) = \frac{U_2(e^{jw})}{U_1(e^{jw})} \tag{7.4}$$

where, $U_1(e^{jw})$, $U_2(e^{jw})$ is at the ends of the frequency domain characteristics.

For linear circuit network, the transfer function reflects the characteristics of the circuit itself. It is determined by the network topology and component parameters has nothing to do with the external action. For the actual power transformer winding, its equivalent circuit topology and element parameters are determined. If the transformer is of winding deformation, the equivalent circuit parameters will change, which would reflect on the transfer function of the change. Therefore, changes in the transfer function can characterize the deformation of the transformer winding.

7.3.2 Deformation Analysis and Judgment

The deformation was judged by comparing pre and post the fault of the transfer function curve through the simulation of winding deformation. With X_i, $Y_i(i = 1,2,\ldots,n)$ represents the two sampling point sequence of transfer function, the difference can be expressed by the mean square root of the spectrum difference (E):

$$E = \sqrt{\frac{1}{n}\sum\nolimits_{k=1}^{n}(x_i - y_i)^2} \tag{7.5}$$

In the practical application of the transfer function for winding deformation diagnosis, we usually judge the deformation by comparing the differences between different periods of winding transfer function curve, it has a vertical comparison and horizontal comparison method. The former is measured by comparing the transfer function original curves before operation of the transformer with the curve measured. The latter is measured by comparing the transfer function curve of three-phase due to the three-phase consistency of transform. The mean square root of the transfer function spectrum difference was used a criterion to judge the deformation degree.

7.4 Analysis of Each Deformation Degree

In theory, the transfer function has nothing to do with the excitation signal. In fact, the bandwidth, sampling rate, and sampling time of the excitation signal can affect the transfer function. The narrow pulse rectangular wave was adopted as the excitation signal in this paper due to the broadband.

7.4.1 Deformation of the Same Type Different Locations

The transfer function curve when transform winding occurred radial distortion of the same type at different locations is shown in Fig. 7.3. When the winding occurred radial distortion in the 1th disc, the amplitude of transfer function poles at 1.45, 1.7, and 2.3 MHz have larger changes, the pole position at 2.3 MHz Shift to the right. When the winding occurred radial distortion in the 10th disc, the amplitude of transfer function poles at 1.45 and 1.7 MHz have changes, its changes magnitude smaller than radial distortion in the 1th disc.

The mean square root of the spectrum difference is calculated by formula (7.5), when the winding occurred radial distortion in the 1th disc, $E1 = 2.0529$(dB) and when the winding occurred radial distortion in the 10th disc, $E2 = 1.9595$(dB). Through the analysis we can draw a conclusion that the changes of the transfer function is more obvious than in the middle position when the deformation occurred in the head end.

The cause of the transfer function spectrum fluctuate is that the transfer function of the transform winding is determined by the network topology and component parameters. When the component parameter changes, the corresponding performance on transfer function curve have two changes:

(1) Frequency offset, that is the pole of transfer function will deviate from the original position;
(2) Pole height variation, that is the changes in the amplitude transfer function poles.

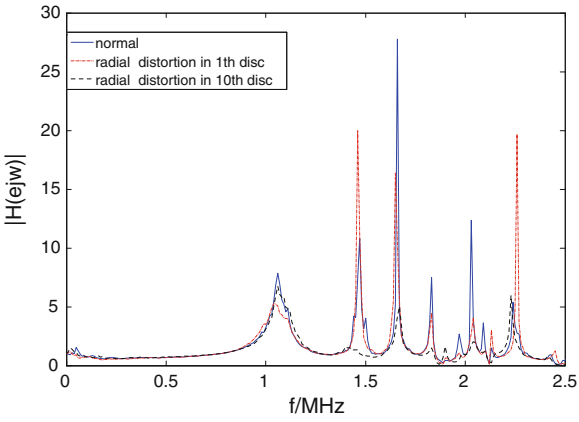


Fig. 7.3 Transfer function curve when deformation occurring at different positions of the same type

7.4.2 Deformation of the Same Locations Different Type

Figure 7.4 shows the transfer function curve when the transform winding occurred radial and axial distortion in the same disc (the 10th disc). When the winding occurred axial distortion in the 10th disc, the amplitude of transfer function poles at 1.1 and 1.45 MHz have a great decline and the pole position at 2.3 MHz Shift to the right. When the winding occurred radial distortion in the 10th disc, the amplitude of transfer function poles at 1.45 and 1.7 MHz have changes, it changes magnitude smaller than axial distortion in the 10th disc.

The mean square root of the spectrum difference is calculated by formula (7.5), when the winding occurred radial distortion in the 10th disc, $E1 = 1.9595$(dB) and when the winding occurred axial distortion in the 10th disc, $E2 = 2.0765$(dB).

From these data we can reason out that the axial distortion has a greater influence on the transfer function than the radial distortion when transform windings deformed on the same disc. The axial distortion is the axial displacement of the whole disc, it primarily affects the equivalent capacitance of interdisc for the equivalent circuit model of the transformer winding. We can see that the axial displacement of the whole disc has a great influence on the transform windings than the radial distortion.

7.4.3 Deformation of the Different Locations Different Type

Based on the above analysis, the common axial and radial deformations of transformer winding was studied to judge the degree of winding deformation. According to the number of deformed discs, the degree of winding deformation is divided into three kinds of circumstances: slight deformation, moderate deformation, severe deformation. Figure 7.5 and Table 7.1 show that the E values after

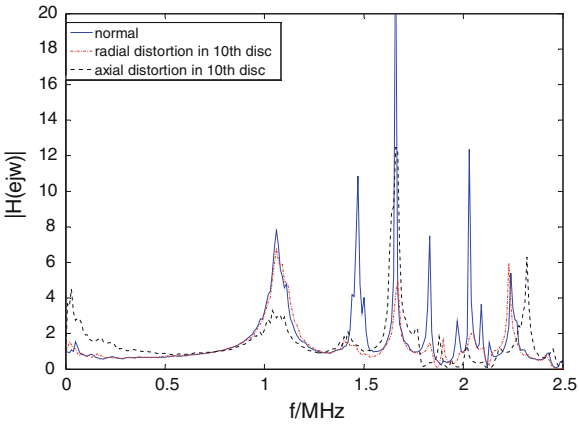


Fig. 7.4 Transfer function curve when deformation occurring at same positions of the different type

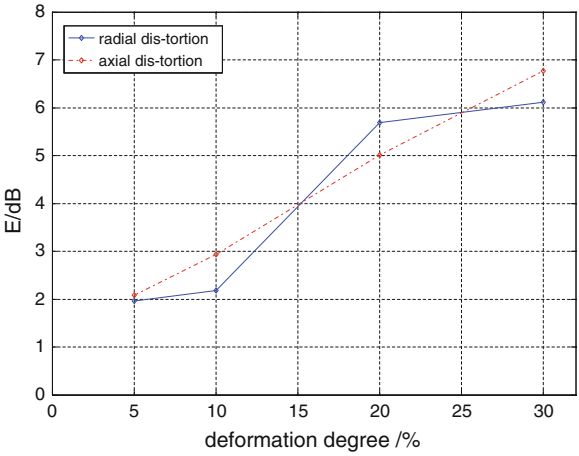


Fig. 7.5 Transformer E(dB) values under different deformation conditions

Table 7.1 Transformer E(dB) values under different deformation conditions

Deformation mode	Deformation degree/ %			
	5	10	20	30
Radial distortion	1.96	2.18	5.69	6.12
Axial distortion	2.08	2.43	5.01	6.77

different forms and degree of deformation. We can see from it that the mean square root of the spectrum difference will increase with the increase of deformation degree. Therefore, the mean square root of the spectrum difference can be used as a criterion to judge the deformation of the transform windings.

7.5 Conclusions

The lumped parameter method was used to establish the equivalent circuit of transformer winding in this paper. It established a basis for calculating the transfer function of the transformer. Three kinds of deformation was analyzed to study the different deformation by simulation. Through the analysis we can see that the changes of the transfer function is more obvious than in the middle position when the deformation occurred in the head end; the axial distortion has a greater influence on the transfer function than the radial distortion when transform windings deformed on the same disc. The mean square root of the spectrum difference was proposed as a criterion to judge the deformation of the transform windings. The changes of the transfer function characteristic provide a basis for monitoring and diagnosis the deformation of the transform windings, it can perfect relevant transformer winding deformation test method and criterion by it.

Acknowledgments This work is supported by basic scientific research funded projects of Beijing Jiao Tong University (No.2012JBM100).

References

1. Diek EP, Erven CC (1978) Transformer diagnostic testing by frequeney response analysis.·IEEE Trans·Power APP Aratus Syst 97(6):2144–2153
2. Malewski R, Poulin B (1988) Impulse testing of power transformers using the transfer function method. IEEE Trans Power Deliv 3(2):476–489
3. Christian J, Feser K (1999) Diagnostics of power transformers by using the transfer function method. In: IEEE conference publication proceedings of the 1999 11th international symposium on high voltage engineering (ISH99), London
4. Rajasekaran S, Yu X, Fetherstone F (2000) Diagnosis and identification of transformer faults from frequency response data. In: Conference record of the 2000 IEEE power engineering society winter meeting, Singapore
5. Xuechang Y, Yongping W (1997) Calculation of transient voltage distribution in transformers using bergeron's method. Tsinghua Sci Technol 2(4):838–841 (in Chinese)
6. Kezunovic M, Guo Y (2000) Modeling and simulation of the power transformer faults and related protective relay behavior. IEEE Trans Power Deliv 15(1):44–50
7. Wang Z (1996) Modeling for simulating very fast transient in transformer windings. Proc CSEE 16(5):299–305 (in Chinese)
8. Gui J (2003) Fault diagnosis of power transformer based on the principle of pulse propagation along the winding. Department of Electrical Engineering, Tsinghua University, Beijing (in Chinese)
9. Qi Y, Yang X, Zhou W (2002) Based on the transfer function of power transformer insulation fault and diagnostic methods. Automation of electric power systems (1). (in Chinese)
10. Gui J,Gao W,Tan K (2004) Criterion for transformer winding deformation calculated by the structural parameter method. J Tsinghua University (Science and Technology), vo1 44,N.1. (in Chinese)

Chapter 8
Clamped Three-level Inverter Midpoint Potential Control Method

Haijie Jia, Xuedong Jiang, Lei Wang, Zheming Jin and Qifeng Li

Abstract From the perspective of the current relationship, the reason of midpoint potential fluctuations for three-level inverter was analyzed. And based on redundant small vectors compensation, the action time correction methods for the basic voltage vectors and the original fundamental voltage vector have been proposed. Finally a simulation model was built in MATLAB for analysis, which showed that the three-level neutral point potential fluctuations can be suppressed effectively by the midpoint potential compensation algorithm.

Keywords Three-level inverter · Midpoint potential · Control method

8.1 Introduction

With the increasing use of the high-voltage and high-power inverter, outstanding advantages of multilevel inverter have gradually been recognized. Compared with the two-level inverter, three-level inverter halved the voltage level for switching device, improved utilization of the voltage of DC link, which also reduced the switching frequency to a certain extent, and hence reduces the device switching losses, meanwhile the output voltage harmonics was reduced. Therefore, it is more suitable for medium and high-voltage high-power applications. Three-level inverter midpoint voltage fluctuation problem has always been an important factor restricting its development. In neutral point clamped three-level inverter, two

H. Jia (✉) · X. Jiang · L. Wang · Z. Jin
School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China
e-mail: 13125928@bjtu.edu.cn

Q. Li
Combat and training division, Public Security Fire Department of Gansu Province, Lanzhou 730000, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_8,
© Springer-Verlag Berlin Heidelberg 2014

capacitors in series are used for partial pressure. As using SVPWM modulation algorithm will inevitably cause voltage fluctuations in the midpoint, and will increase the output low-order harmonics, which will also lead to an increased switch voltage stress, so to achieve the effective control of the midpoint voltage fluctuations has important significance to enhance the applicability for three-level inverter [1]. The reason of midpoint potential fluctuations for three-level inverter was analyzed from the perspective of the current relationship and then the action time correction methods for the basic voltage vectors and the original fundamental voltage vector have been proposed in this paper, which has been verified by a Simulation Model built in MATLAB.

8.2 Cause Analysis of Midpoint Potential Fluctuation

The current relationship of three-level inverter was showed in Fig. 8.1. Analysis of specific current relationship:

$$i_{np} = (1 - |S_a|) \cdot i_a + (1 - |S_b|) \cdot i_b + (1 - |S_c|) \cdot i_c \tag{8.1}$$

i_{np} is the equivalent current outflowing the midpoint, i_a, i_b, i_c are the AC current instantaneous value of each phase [2, 3].

Taking the respective switching state corresponded to the fundamental voltage vector in the fundamental voltage vector set into the equation above, we can get the equivalent current values outflowing the midpoint shown in Table 8.1.

Obviously, when the fundamental voltage vector is small sector or medium vector, because $i_{np} \neq 0$, so we can get:

$$i_{np} = i_{c1} - i_{c2} = C_1 \cdot \frac{d}{dt} u_{c1} - C_2 \cdot \frac{d}{dt} u_{c2} \neq 0 \tag{8.2}$$

Thus:

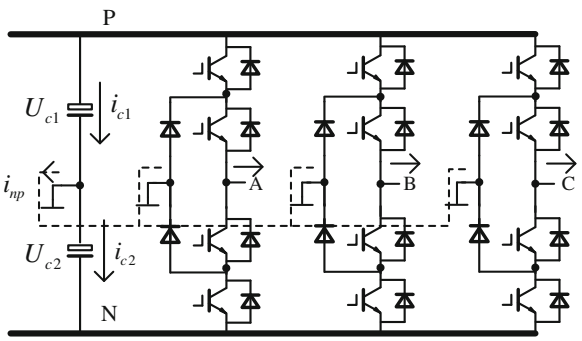


Fig. 8.1 The current composition in midpoint of NPC-TLC

Table 8.1 The equivalent current values outflowing the midpoint

No	S_a	S_b	S_c	i_{np}	No	S_a	S_b	S_c	i_{np}
Z_1	0	0	0	0	S_{12}	0	–1	0	$-i_b$
Z_2	1	1	1	0	M_1	1	0	–1	i_b
Z_3	–1	–1	–1	0	M_2	0	1	–1	i_a
S_1	1	0	0	$-i_a$	M_3	–1	1	0	i_c
S_2	0	–1	–1	i_a	M_4	–1	0	1	i_b
S_3	1	1	0	i_c	M_5	0	–1	1	i_a
S_4	0	0	–1	$-i_c$	M_6	1	–1	0	i_c
S_5	0	1	0	$-i_b$	L_1	1	–1	–1	0
S_6	–1	0	–1	i_b	L_2	1	1	–1	0
S_7	0	1	1	$-i_a$	L_3	–1	1	–1	0
S_8	–1	0	0	i_a	L_4	–1	1	1	0
S_9	0	0	1	$-i_c$	L_5	–1	–1	1	0
S_{10}	–1	–1	0	i_c	L_6	1	–1	1	0
S_{11}	1	0	1	i_b					

$$C_1 \cdot \frac{d}{dt}u_{c1} - C_2 \cdot \frac{d}{dt}u_{c2} = C(\frac{d}{dt}u_{c1} - \frac{d}{dt}u_{c2}) = C \cdot \frac{d}{dt}(2 \cdot \Delta u) \neq 0 \tag{8.3}$$

And then:

$$C \cdot \frac{d}{dt}\Delta u \neq 0 \tag{8.4}$$

where Δu is the fluctuation of midpoint voltage.

When the fundamental voltage vector is a large, Because $i_{np} = 0$, so we can get:

$$C \cdot \frac{d}{dt}\Delta u = 0 \tag{8.5}$$

It means that there is no point potential fluctuation.

From Table 8.1, i_{np} equal to zero or A phase alternating current instantaneous value or its opposite number. Meanwhile, i_{np} is equal size and opposite between any two mutually redundant small vector. (such as S_1 corresponded $i_{np} = -i_a$, S_2 corresponded $i_{np} = i_a$), Resulting in an opposite effect on midpoint potential.

8.3 Solution for Midpoint Potential Fluctuations Based on Redundant Small Vectors Compensation

The opposite effects of mutually redundant small vector can be used to offset the midpoint potential drift midpoint potential fluctuations [4, 5]. The diagram of the redundant small vectors compensation algorithm is shown as in Fig. 8.2.

In Fig. 8.2, u_{c1} is the voltage of upper capacitor; $t_{\Sigma s}$ is the active time of a small vector in vector space during a switching cycle, the switching of the two redundant

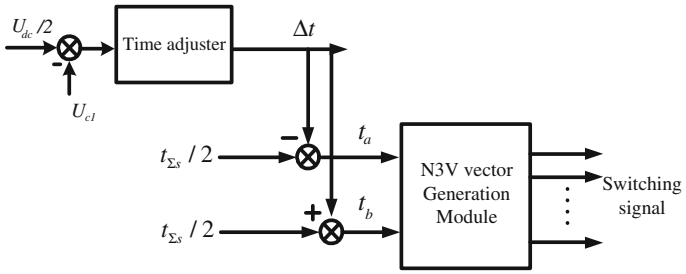


Fig. 8.2 The diagram of the redundant small vectors compensation algorithm

switch combination is realized by closed-loop control thereby inhibiting vector midpoint potential fluctuations. If the time controller of Fig. 8.2 is a simple proportional adjustment with the scale factor Kp, the adjustment mechanism can be expressed as follows:

$$\begin{cases} t_a = \dfrac{1}{2}t_{\Sigma s} - K_p(\dfrac{U_{dc}}{2} - U_{c1}) \\ t_b = \dfrac{1}{2}t_{\Sigma s} + K_p(\dfrac{U_{dc}}{2} - U_{c1}) \end{cases} \tag{8.6}$$

Actually, using this switching method modified the basic voltage vectors sets. For example, in Sector I the basic voltage vector set is modified as follows:

$$\begin{cases} S_{c1} & = \frac{t_a}{t} \cdot S_1 + \frac{t_b}{t_{\Sigma s}} \cdot S_2 \\ S_{c2} & = \frac{t_a}{t_{\Sigma s}} \cdot S_3 + \frac{t_b}{t_{\Sigma s}} \cdot S_4 \\ M_{c1} & = M_1 \\ L_{c1} & = L_1 \\ L_{c2} & = L_2 \\ 0_c & = 0 \end{cases} \tag{8.7}$$

Among them, vectors with the subscript c indicate the vectors modified by the compensation algorithms, the other vectors have the same meaning as before [6, 7]. The modified basic voltage vector sets shown in Fig. 8.3:

By using modified basic voltage vector sets as shown in Eq. (8.7), the midpoint potential can be effectively controlled. If basic voltage vector sets cannot be modified, in order to achieve redundant small vector compensation, the active time of each basic voltage vectors can be modified [8, 9], the amendment in Sector I is shown as follows:

In Table 8.2, the meaning of τ is shown in Eq. (8.8):

$$\tau = \frac{t_b}{t_{\Sigma s}} = \begin{cases} > \frac{1}{2}, & u_{c1} < \frac{U_{dc}}{2} \\ \leq \frac{1}{2}, & u_{c1} \geq \frac{U_{dc}}{2} (t_{\Sigma s} = t_1,\ t_2 \text{ or } t_3) \end{cases} \tag{8.8}$$

when the reference voltage vector is located inside triangle T1 in each sector, i.e., the modulation factor $MI \leq 0.4534$, only small vectors affect the midpoint potential during the entire active time; when the modulation factor $MI > 0.4534$,

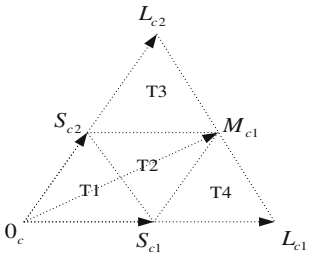


Fig. 8.3 The modified basic voltage vector sets

Table 8.2 Acting time correction in Sector I

Section	a	b	c	d	e	f	g
1	$\frac{t_3}{4}\cdot\tau$	$\frac{t_2}{2}$	$\frac{t_1}{2}$	$\frac{t_3}{2}\cdot(1-\tau)$	$\frac{t_1}{2}$	$\frac{t_2}{2}$	$\frac{t_3}{4}\cdot\tau$
2	$\frac{t_2}{4}\cdot\tau$	$\frac{t_1}{2}$	$\frac{t_3}{2}$	$\frac{t_2}{2}\cdot(1-\tau)$	$\frac{t_3}{2}$	$\frac{t_1}{2}$	$\frac{t_2}{4}\cdot\tau$
3	$\frac{t_3}{4}\cdot\tau$	$\frac{t_1}{2}$	$\frac{t_2}{2}$	$\frac{t_3}{2}\cdot(1-\tau)$	$\frac{t_2}{2}$	$\frac{t_1}{2}$	$\frac{t_3}{4}\cdot\tau$
4	$\frac{t_1}{4}\cdot\tau$	$\frac{t_2}{2}$	$\frac{t_3}{2}$	$\frac{t_1}{2}\cdot(1-\tau)$	$\frac{t_3}{2}$	$\frac{t_2}{2}$	$\frac{t_1}{4}\cdot\tau$
5	$\frac{t_1}{4}\cdot\tau$	$\frac{t_3}{2}$	$\frac{t_2}{2}$	$\frac{t_1}{2}\cdot(1-\tau)$	$\frac{t_2}{2}$	$\frac{t_3}{2}$	$\frac{t_1}{4}\cdot\tau$
6	$\frac{t_1}{4}\cdot\tau$	$\frac{t_3}{2}$	$\frac{t_2}{2}$	$\frac{t_1}{2}\cdot(1-\tau)$	$\frac{t_2}{2}$	$\frac{t_3}{2}$	$\frac{t_1}{4}\cdot\tau$

the reference voltage vector is combined by small and medium vectors, so the effects of medium vector can also be compensated by modified small vectors.

8.4 Simulation Results of Midpoint Potential Compensation Algorithms

In order to verify the compensation algorithms purposed above, a simulation module of NPC-TLC was built based on MATLAB/Simulink. The voltage of DC supply was 700 V, the instantaneous voltage and of the capacitors and under different modulation factor MI and different load conditions was shown as follows:

In Fig. 8.4, is shown u_{c1} and u_{c2} without the midpoint potential compensation algorithm while MI = 0.314, R = 2 Ω. The Fig. 8.5 shows u_{c1} and u_{c2} with the midpoint potential compensation algorithm while MI = 0.314, R = 2 Ω.

In Fig. 8.5, is shown u_{c1} and u_{c2} without the midpoint potential compensation algorithm while MI = 0.628, R = 2 Ω. The Fig. 8.6 shows u_{c1} and u_{c2} with the midpoint potential compensation algorithm while MI = 0.628, R = 2 Ω.

In Fig. 8.6, is shown u_{c1} and u_{c2} without the midpoint potential compensation algorithm while MI = 0.628, R = 1 Ω. The figure below shows u_{c1} and u_{c2} with the midpoint potential compensation algorithm while MI = 0.628, R = 1 Ω.

In Fig. 8.7, is shown u_{c1} and u_{c2} without the midpoint potential compensation algorithm while MI = 0.628, R = 4 Ω. The Fig. 8.8 shows u_{c1} and u_{c2} with the midpoint potential compensation algorithm while MI = 0.628, R = 4 Ω.

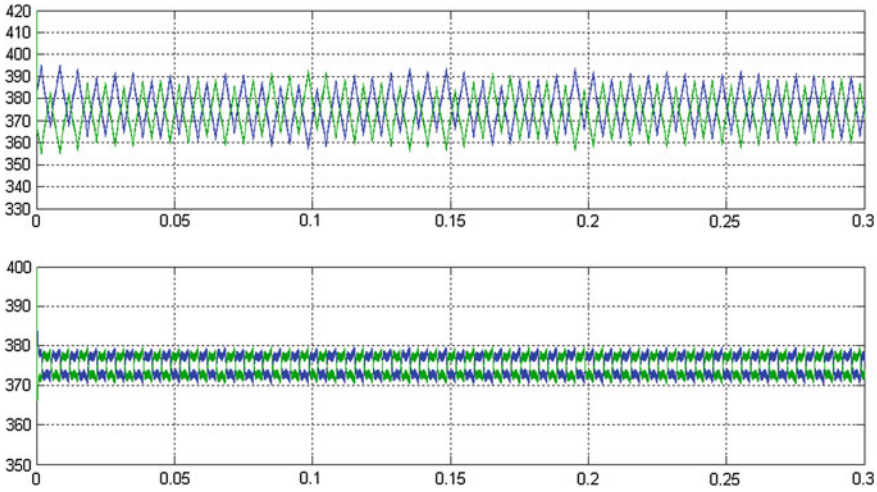


Fig. 8.4 The instantaneous voltage u_{c1} and u_{c2} of the capacitors C_1 and C_2

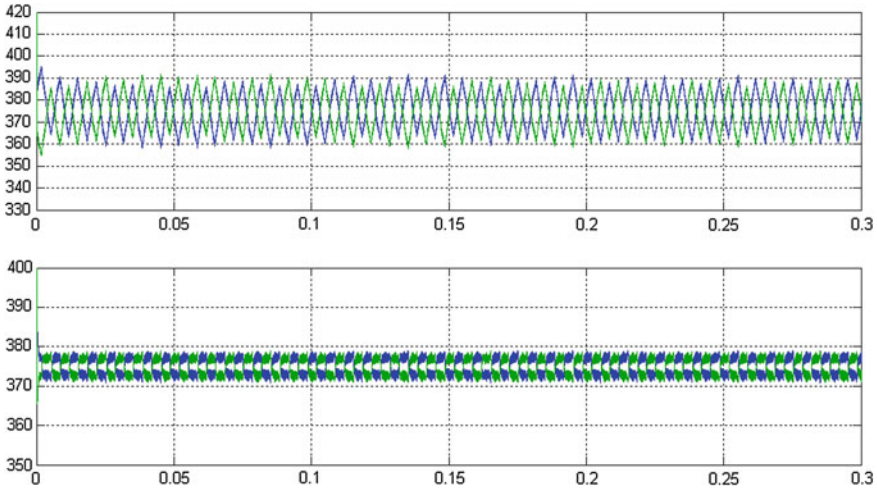


Fig. 8.5 The instantaneous voltage u_{c1} and u_{c2} of the capacitors C_1 and C_2

In Fig. 8.8, is shown the line voltage u_{ab} without the midpoint potential compensation algorithm while MI = 0.628, R = 1 Ω. The figure below shows u_{ab} with the midpoint potential compensation algorithm while MI = 0.628, R = 1 Ω.

The simulation results show that the midpoint potential under different modulation factor MI and different load conditions was successfully controlled by using midpoint potential compensation algorithm.

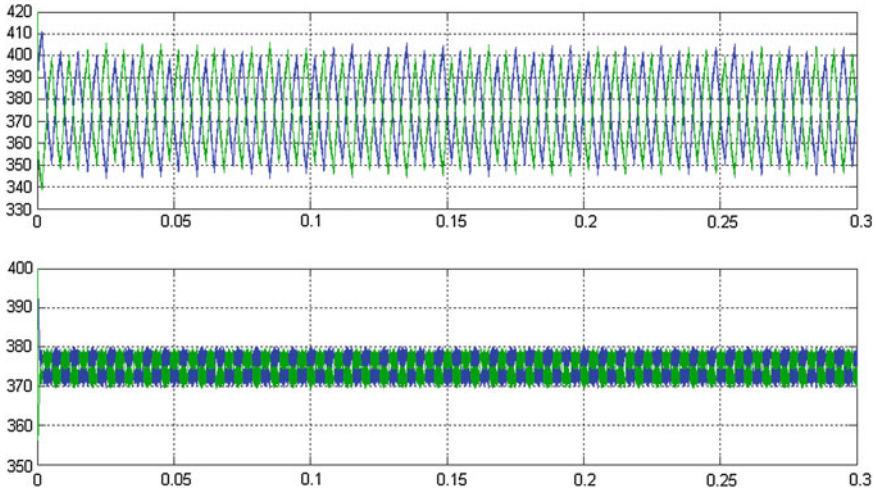


Fig. 8.6 The instantaneous voltage u_{c1} and u_{c2} of the capacitors C_1 and C_2

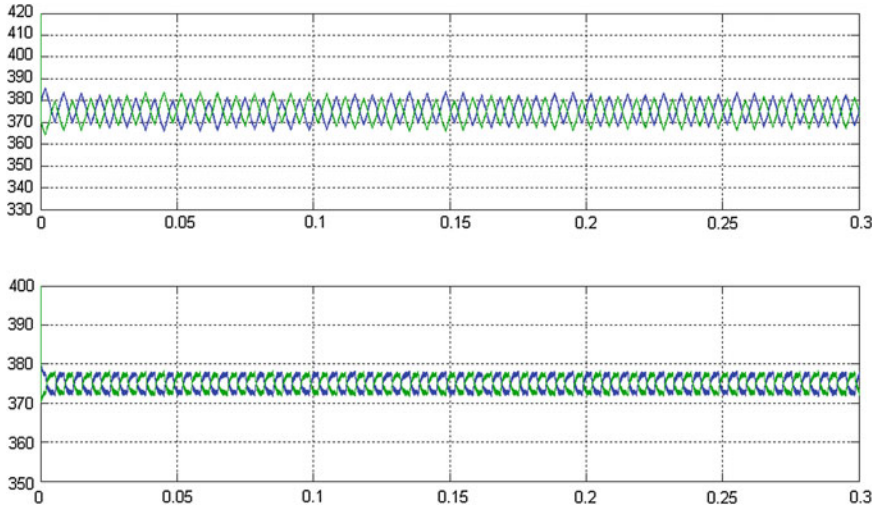


Fig. 8.7 The instantaneous voltage u_{c1} and u_{c2} of the capacitors C_1 and C_2

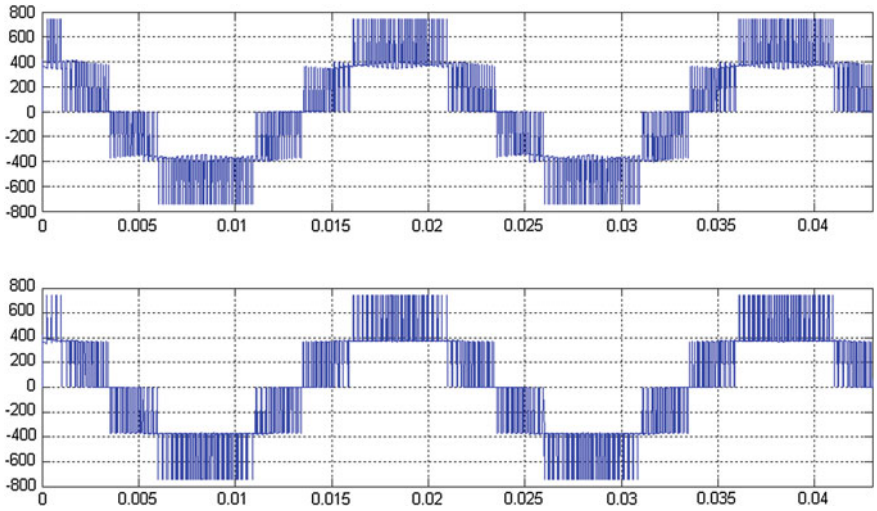


Fig. 8.8 Line voltage u_{ab}

8.5 Conclusions

This paper first analyzed the reason of midpoint potential fluctuations from the perspective of the current relationship of TLC inverter, proposed an algorithm based on redundant small vectors compensation with modified basic voltage vector sets or modified active time of original basic voltage sets. And then, a simulation module was built based on MATLAB/Simulink to verify the algorithm. The Analysis of simulation results proved that using the midpoint potential compensation algorithm can effectively control the midpoint potential fluctuations.

Acknowledgments The work was supported by Beijing Municipal Commission of Transport under grant E13M00020 (Research about High performance electric traction control and diagnostic techniques of Light rail vehicles). We would like to thank the anonymous reviewers for their comments and suggestions.

References

1. Zhang Y, Tang Y (2010) Research on SVPWM method and neutral-point voltage balance in three-level inverter. Electr Drive 40(2):33–36 (in Chinese)
2. Xie Z, Xiao F (2012) A novel neutral-point potential balancing technique based on fuzzy logic for three-level NPC SVPWM inverter. J Power Supply 2012(5):24–31 (in Chinese)
3. Pou J, Pingdado R, Boroyevich D et al (2005) Evaluation of the low-frequency neutral-point voltage oscillation in the three-level inverter. IEEE Trans Ind electron 52(12): 1582–1588

4. Jin H, Zou Y (2010) Research on the technology of the neutral-point voltage balance and dual-loop control scheme for three-level PWM rectifier. Proc Chin Soc Electr Eng 26(20):64–68 (in Chinese)
5. Song W, Chen G (2006) A novel SVPWM strategy and its implementation considering neutral-point potential balancing for three-level NPC inverter. Proc Chin Soc Electr Eng 26(12):95–100 (in Chinese)
6. Li P, Liu G (2012) Research on SVPWM control scheme for NPC three-level inverter and balance of neural-point voltage. Chin Meas Test Technol 38(5):75–79 (in Chinese)
7. Xiao F, Xie Z (2013) Integrate control strategy for three-level neutral-point clamped voltage source inverter with unbalanced load. High Voltage Eng 39(2):474–480 (in Chinese)
8. Dou Z, Zhang T (2008) SVPWM and neutral-point-potential balance control of three-level NPC rectifier. Electr Power Autom Equip 28(2):65–69 (in Chinese)
9. Liu Z, Wu M (2007) Research on space vector pulse width modulation method for neutral-point-clamped three-level inverter. Meas Control Technol 26(11):80–83 (in Chinese)

Chapter 9
Design and Research for the Low-Loss Passive Snubber Circuit Used in the Metro High Frequency Auxiliary Converter

Yao Fu, Lijun Diao, Huiqing Du, Haijie Jia and Lei Wang

Abstract In view of the secondary rectifier diode overvoltage damage problems in the DC–DC link during the operation process of the metro auxiliary converter, this paper analyzes the mechanism of overvoltage and a novel way of RCD snubber is proposed. Different from the traditional snubber circuits, it has the characteristics of less resistor heat, high energy utilization rate, good voltage spike suppression effect, etc. On the basis of metro auxiliary converter circuit topology, this paper has carried out detailed research on the energy flow during the snubber circuit voltage clamped process and put forward the snubber parameter calculation method. Then experimental verification is made by a simulation model based on SABER and 77 kVA prototype platform, which proves the rationality of the RCD snubber circuit design.

Keywords Auxiliary converter · Resonance · RCD snubber circuit · Voltage clamped

9.1 Introduction

In recent years, with the speeding up of urbanization, construction of the metro has developed rapidly, which makes power supply for the metro on-board equipment requirement gradually increased. As an important part of the auxiliary power supply system, auxiliary converter is also developing in the direction of high power and high frequency. In this direction, the problem of the rectifier diode used in the auxiliary converter under high voltage spike is becoming more and more serious, especially in high frequency working conditions. In the diode reverse

Y. Fu (✉) · L. Diao · H. Du · H. Jia · L. Wang
School of Electrical Engineering, Beijing Jiaotong University, No.3 Shang Yuan Cun, Beijing, Hai Dian District, China
e-mail: 12121479@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_9,
© Springer-Verlag Berlin Heidelberg 2014

recovery time, due to the resonance between its junction capacitor and circuit stray inductor, the reverse recovery overvoltage often causes diode breakdown [1], which seriously affects the reliability and service life of the auxiliary converter. For this reason, measures are required to be taken to suppress the rectifier diode voltage spike.

In this paper, RCD snubber method is presented through analyzing the generation mechanism of diode voltage spike. Moreover, different snubber circuit topologies and parameters lead to different effects on voltage spike suppression. Combined with the characteristics of metro auxiliary converter circuit, a novel way of RCD snubber is put forward and the calculation principle of snubber parameters is deduced. Finally, the rationality of the design is validated through the simulation and experiment.

9.2 Circuit Overvoltage and the Principle of RCD Snubber

The metro auxiliary converter mainly supplies alternating current (ac) for equipment such as air conditioners and fans. Figure 9.1 shows the auxiliary inverter circuit topology. Since the rated input dc voltage is 1,500 V, whose range is DC1,000–DC2,000 V, two DC/DC conversion unit series connection structure is adopted to reduce the switching devices stress. What is more, within each unit there are capacitors and equal resistors to keep the voltage balanced. LC output filter circuit can reduce the intermediate dc side harmonics and three-phase bridge converter converts the intermediate dc voltage to three-phase 380 V, 50 Hz alternating current. In order to prevent transformer secondary rectifier circuit voltage spike, RCD snubber circuit is added in the output side of each DC/DC conversion unit.

9.2.1 Circuit Overvoltage Mechanism

Shown in Fig. 9.2 is a simplified circuit of the conversion unit. The transformer turns ratio is 1.5:1, V_s is an equivalent pulse voltage source of the transformer secondary side output voltage and L_1 is the equivalent leakage inductor of the transformer secondary side. In a switching period, for example, prior to the t_0, $V_s = 0$. With the output filter inductor freewheeling, output current I_o flows through four rectifier diodes, which makes rectifier bridge output voltage V_{sn} clamped at zero; at t_0 moment, V_s steps from zero to V_d just when the front-end switching tubes complete current reversal process. At this time, due to point A electric potential is approximately zero, L_1 instantly withstands voltage of V_d that results in I_s increasing with a slope of V_d/L_1 [2]. When I_s reaches I_o, D_2, and D_3 stop conducting and current flows through D_1 and D_4. When I_s exceeds I_o, the excess current is given as $I_r = I_s - I_o$. Ir flows through the parasitic capacitor C_p

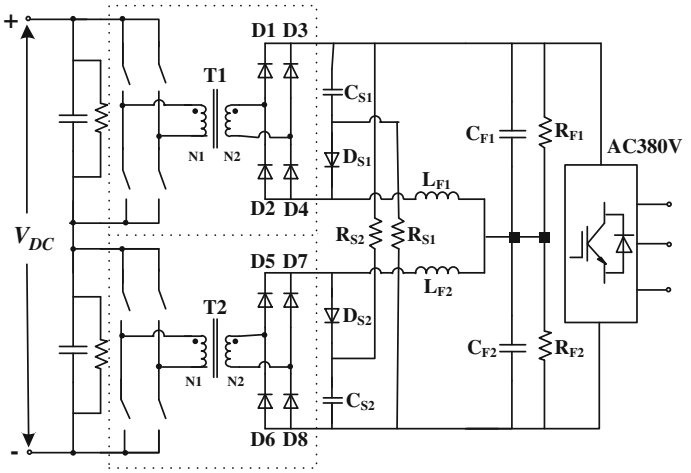


Fig. 9.1 Circuit topology of auxiliary converter for metro

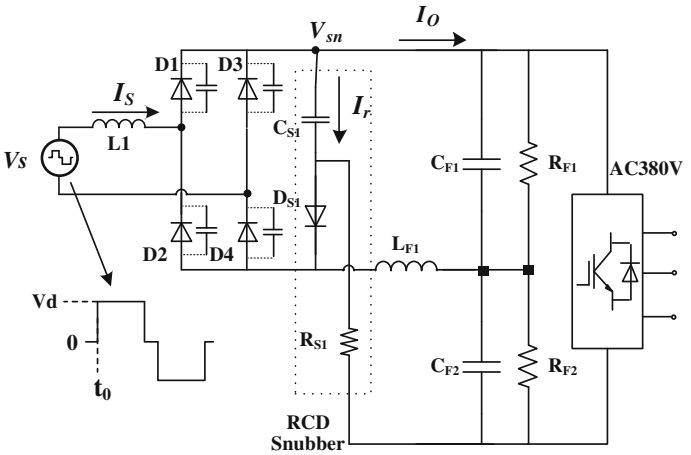


Fig. 9.2 Simplified circuit of conversion unit

(including the junction capacitor of diode D_2 and D_3 and isolation transformer stray capacitor), which at this time causes the resonance between C_p and L_1 [3]. Excessive harmonic current leads to reverse voltage drop on the inductor L_1, causing V_s to increase. Assuming that the filter inductor L_f is large enough to hold the output current I_o within a short time, dynamic equation of the resonant circuit is

$$L_1 \frac{dI_S}{dt} + V_{C_P} = V_d \tag{9.1}$$

$$C_p \frac{dV_{C_p}}{dt} + I_O = I_S \tag{9.2}$$

With $I_s = I_o$ as the initial time and initial conditions $V_{Cp0} = V_{s0} = 0$ and $I_{s0} = I_o$, solving the differential equations, we get:

$$V_{sn} = V_{C_P} = V_d \left[1 - \cos\left(\frac{t}{\sqrt{L_1 C_P}} \right) \right] \tag{9.3}$$

$$I_s = I_o + V_d \sin\left(\frac{t}{\sqrt{L_1 C_P}} \right) \Big/ (L_1 / C_P) \tag{9.4}$$

The maximum reverse voltage spike diode D_2 and D_3 bears is $2V_d$ [4]. Under the rated working conditions, $V_d = 500$ V while diode voltage spike reaches 1,000 V. For the commonly used 1,200 V fast recovery diode, if we take into account the voltage spike brought by front-end switching devices, V_s could reach diode reverse voltage limit value, which may reduce the reliability of the circuit. When the input voltage is up to 2,000 V, rectifier diode would even burn up unless the voltage spike suppression measure is taken. At present, RCD snubber circuit gets widespread adoption to suppress the diode voltage spike.

9.2.2 RCD Snubber Principle Analysis

Considering the different circuit structures, we can adopt different RCD snubber topologies. Figure 9.3 shows the two RCD topologies constantly used in the output of diode rectifier bridge at present.

In Fig. 9.3a, the snubber resistor R_{s1} is directly parallel to the snubber capacitor. This topology structure can make the snubber capacitor discharge through the parallel resistor rapidly, eliminating snubber capacitor charge accumulation due to extremely short switching period, which is suitable for high frequency and low power applications. However, in high power situations, too much instantaneous discharge energy flows through snubber resistor R_{s1}, leading to resistor over temperature and even directly burned in severe cases [5, 6].

As Fig. 9.3b shows, the snubber capacitor discharges through the snubber resistor R_{s1}, the resistor R_{f1}, and the filtering inductor L_{f1}. This topology, using resistor R_{f1} withstands voltage, reducing the discharge energy that snubber resistor bears. But because of the filter inductor, the capacitor discharges too slowly, which is likely to cause charge accumulation of the snubber capacitor.

On the basis of the auxiliary converter topology plus the snubber circuit structure shown in Fig. 9.3b, a new RCD snubber circuit is to be designed. As shown in Fig. 9.2, the RCD snubber circuit consist of R_{s1}, C_{s1}, and D_{s1} can effectively solve the problem of oscillation. C_{s1} capacitance is generally large enough to maintain the C_{s1} voltage V_{Cs}. When D_2 and D_3 turn off, L_1 and C_p start resonance. D_{s1} cannot conduct until V_S rise up to V_{Cs}, when C_{s1} begins to join in

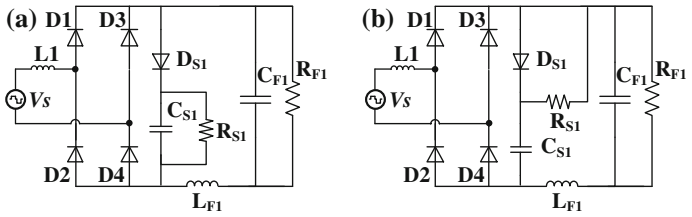


Fig. 9.3 Commonly used circuit topology of RCD snubber

the resonance. Because of large C_s capacitance, V_{Cs} will be constant and V_s is clamped at V_{Cs}, suppressing the voltage spike. When it comes to the end of the resonance, the charge of C_{s1} will flow through R_{f1}, R_{f2}, and R_s. The discharge route using resistor R_{f1} and R_{f2} to withstand voltage makes the snubber resistor R_s bear less discharge energy. What is more, in the discharge circuit there is no nonlinear device such as inductor, which may slow down the discharge speed [7, 8]. Resistor R_{f1} and R_{f2} can also give back the extra energy of snubber capacitor C_{s1} gained in the resonant process to output voltage, which can improve the circuit efficiency. In brief, this RCD snubber circuit takes advantage of the capacitor charging and discharging to release the resonance energy by resistors and output loop.

9.2.3 Snubber Energy Analysis

In order to analyze the circuit energy flow in the process of resonance, it is necessary to study voltage and current of the snubber circuit. As shown in Fig. 9.4, it is the curve using the new RCD snubber circuit, voltage on both ends of the snubber circuit V_{sn} and snubber circuit charging current I_r in two neighboring resonance cycles.

In Fig. 9.4, when the voltage spike V_{sn} reaches V_{Cs}, Ds1 starts conducting. At the same time, Ir declines to zero at a rate of $(V_{Cs}-V_d)/L1$. From then, the next resonance cycle begins. Its initial conditions are $V_{sn} = V_{Cs}$, $I_s = I_o$. Again solving the differential equations, you can find the next voltage spike will not reach V_{Cs} because of loss in the circuit. The clamping function of RCD circuit occurs only once and its duration is

$$\Delta t = L_1 \frac{I_r(t_1)}{V_{Cs} - V_d} \tag{9.5}$$

The energy correspondingly flown into the snubber capacitor is

$$E_{Cs} = \frac{1}{2} I_r(t_1) V_{Cs} \Delta t = \frac{1}{2} L_1 I_r^2(t_1) \frac{V_{Cs}}{V_{Cs} - V_d} \tag{9.6}$$

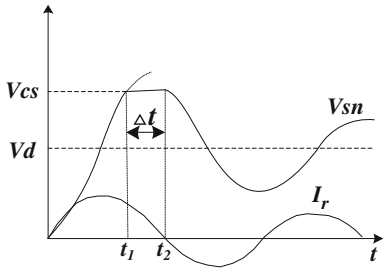


Fig. 9.4 *Curve* voltage and current of the snubber circuit

Substituting V_{sn} $(t_1) = V_{Cs}$ into Eqs. (9.1) and (9.2), we have

$$I_r^2(t_1) = \frac{C_p}{L_1} V_{Cs}(2V_d - V_{Cs}) \tag{9.7}$$

By substituting (9.7) into (9.6), we know the energy flown into the snubber capacitor is

$$E_{Cs} = \frac{1}{2} C_p V_{Cs}^2 \frac{2V_d - V_{Cs}}{V_{Cs} - V_d} \tag{9.8}$$

The analysis of the above is based on linear capacitor. As the clamped voltage of RCD snubber circuit is larger, the charging energy of the snubber capacitor is smaller. Nevertheless, the voltage stress rectifier diodes withstand also increases. Therefore, so as to make the RCD snubber circuit achieve the best clamping effect, the selection of snubber resistor and capacitor is particularly important.

9.3 Calculation Principle of the Snubber Parameters

9.3.1 Snubber Resistance

In steady state, in order to ensure that the snubber capacitor C_p maintains a stable voltage V_{Cs} in the process of charging and discharging, the average current of the snubber capacitor must be zero [9]. On account of this, we can calculate the snubber resistance.

During t_1 and t_2 in Fig. 9.4, I_r declines with a certain linear slope. The charging current of the snubber capacitor by integrating time is

$$\int_{t_1}^{t_2} I_r(t)\,\mathrm{d}t = \frac{I_r(t_1)}{2}\Delta t \tag{9.9}$$

Substituting Eqs. (9.5) and (9.7) into (9.9), we have

$$\int_{t_1}^{t_2} I_r(t)\,\mathrm{d}t = \frac{1}{2} C_p V_{Cs} \frac{2V_d - V_{Cs}}{V_{Cs} - V_d} \tag{9.10}$$

The current through resistor R_s by integrating time is

$$\int_{t_1}^{t_2} I_{R_S}(t)\mathrm{d}t = \left(\frac{V_{Cs} - V_o}{R_S}\right) \bullet \frac{T}{2} \tag{9.11}$$

Because the average current through the snubber capacitor is zero in steady state, the result of (9.10) is equal to that of (9.11). So the snubber resistance is

$$R_S = \frac{T(V_{CS} - V_o)(V_{CS} - V_d)}{C_p V_{CS}(2V_d - V_{CS})} \tag{9.12}$$

Accordingly, the power consumption of the discharging snubber resistor is

$$P_R = \frac{(V_{Cs} - V_o)^2}{R_{Cs} + R_{F1} + R_{F2}} \tag{9.13}$$

It can be seen that the smaller the snubber resistance, the lower the clamped voltage. This improves the circuit reliability but gives rise to the loss of snubber.

Similarly, it is found that the snubber circuits shown in Fig. 9.3a, b both produce more energy loss than this novel RCD snubber topology. In terms of the resistor energy consumption, the new RCD has more advantages.

For example, with the circuit conditions of $V_\mathrm{d} = 667$ V, $V_\mathrm{o} = 320$ V, $T = 0.2$ ms, $C_\mathrm{p} = 5$ nf, and the clamped voltage $V_\mathrm{Cs} = 700$ V, it's calculated that $R_\mathrm{Cs} = 1.13$ kΩ, chosen 1.2 kΩ.

9.3.2 Snubber Capacitance

In view of energy conservation, there is no difference in the resonance energy regardless of whether the RCD snubber circuit is added. Hence, the snubber capacitance can be calculated.

From Eq. (9.3), it is known that the spike of V_sn could reach $2V_\mathrm{d}$ without RCD snubber circuit. All the energy of the resonant circuit is stored in the parasitic capacitor C_p. According to the capacitor energy calculation method, the resonance energy is given by [10]

$$E_{res} = \frac{1}{2} C_p (2V_d)^2 \tag{9.14}$$

However, adding the RCD snubber circuit, the spike of V_{sn} is clamped at V_{Cs}, which makes resonance energy stored in the parasitic capacitor C_p and the snubber capacitor C_s partly, and the rest is consumed through the discharge resistor. The resonance energy is

$$E_{res} = \frac{1}{2}(C_p + C_s)V_{Cs}^2 + E_R \tag{9.15}$$

Due to the constant resonance energy, which is caused by the parasitic capacitor and the transformer's secondary leakage inductor, Eqs. (9.14) and (9.15) must be equal. Considering that the resonance cycle is much shorter than the switching cycle, E_r can be ignored. Under the same circuit condition used to calculate the snubber resistance, the snubber capacitor $C_s = 13.16$ nF. Based on the actual production technology, the selected capacitance is 15 nF.

9.4 Simulation and Experiment

9.4.1 Snubber Circuit Simulation

According to the above-calculated parameters, the RCD snubber resistance is 1.2 kΩ and the RCD snubber capacitance is 15 nf. The simulation is conducted by SABER on the basis of the metro auxiliary converter circuit topology. The selected circuit parameters are 1,500 V input dc voltage, 5 kHZ switching frequency, the transformer turns ratio of 1.5:1, and the power of 77 kVA. Figure 9.5a, b shows the simulation voltage waveforms of the rectifier diode.

Figure 9.5a is the voltage waveform of the transformer secondary rectifier diode without the RCD snubber circuit. As you see, the voltage spike could be up to 970 V due to the resonance between the transformer leakage inductor and the junction capacitor of the rectifier diode during the diode reverse recovery process [11].

The waveform in Fig. 9.5b is the voltage of transformer secondary rectifier diode after the new RCD snubber circuit is added. The voltage spike is clamped at 700 V due to the effect of voltage clamped of the RCD circuit.

Obviously, by contrasting the above simulation waveforms, the function of the RCD circuit is to use the capacitor to eliminate the voltage spike caused by the LC resonance in the form of energy storage. Consequently, the simulation results illustrate the clamped effect and the correct parameter calculation of the snubber circuit.

9.4.2 Experimental Verification

Adopting the circuit parameters SABER simulation applies, experiment is carried out on the prototype of metro auxiliary converter.

Figure 9.6a shows that without the RCD snubber circuit, the voltage spike on the rectifier diode of the transformer secondary winding is 950 V. The 1,200 V

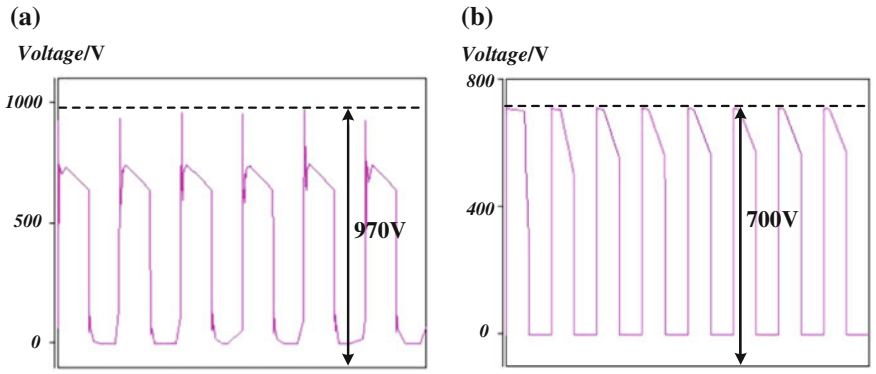
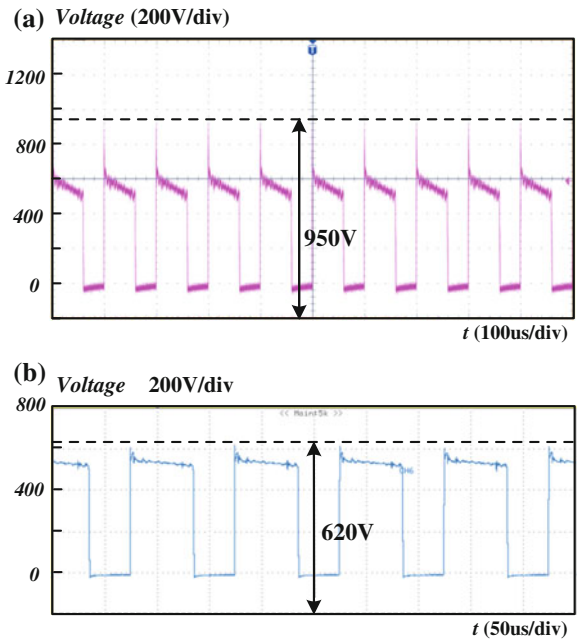


Fig. 9.5 Waveforms of the rectifier diode voltage

Fig. 9.6 Experimental waveforms



fast recovery diodes are used in the prototype. If considering the grid voltage jump in the process of actual operation, the diode will withstand much larger voltage spike, which could reduce the working reliability and endanger the system operation of equipment.

In Fig. 9.6b, the voltage spike is clamped at 620 V. Compared with no RCD snubber circuit, diode voltage spike is reduced 400 V, improving the reliability of the running system.

It can be seen that the prototype experiment results and the simulation results are consistent, verifying the effect of the snubber circuit. On the condition of correct snubber parameters, it is effective to make use of the novel RCD snubber circuit to lower the voltage spike surge that the rectifier diode bears.

9.5 Conclusion

Through the analysis, simulation, and experiment, the diode voltage spike suppression of the topology of metro auxiliary converter is verified. The new RCD snubber circuit making use of the capacitor charging and discharging achieves the goal of voltage clamped by releasing the voltage spike the diode reverse recovery process causes through the discharge circuit partly. Meanwhile, compared with other forms of RCD snubber topology relatively, this new type of RCD circuit sends out less heat. What is more, it gives back partial energy to the output load through the discharging loop, which improves the utilization rate of energy to make the auxiliary converter more secure, reliable, and energy saving.

Acknowledgments This work is supported in part by the National Natural Science Foundation of China under Grant U1134204, the China National Science and Technology Support Program under Grant 2013BAG24B01 and "the Fundamental Research Funds for the Central Universities" under Grant 2012JBM099.

References

1. Xu X (2009) Research on RCD snubber for rectifier circuit of 15 kW full-bridge converter. Power Electron 43(1):41–43 (in Chinese)
2. Cai C, Xue-jun P, Yu C et al. (2011) The loss calculation of RCD snubber with forward and reverse recovery effects considerations. In: Proceedings of the 8th international conference on power electronics - ECCE Asia, the Shilla Jeju, Korea, pp 3005–3012
3. Sun T, Wang H, Zhang X et al (2006) A novel ZVZCS converter using passive clamp circuit. Proceedings of the CSEE 26(17):72–76
4. Zhang W, Wang H, Yuan Ren et al (2010) Investigation on control of three-phase grid-connected inverters under unbalanced grid voltage conditions. Trans China Electrotechnical Soc 25(12):103–110 (in Chinese)
5. Xiao X, Ye X, Wang F (2004) New type lossless absorption circuit of voltage sharp peak in high frequency rectifying circuit. Power Supply Technol Appl 7(8):501–503 (in Chinese)
6. Songyi Lin, Chernlin Chen (1998) Analysis and design for RCD clamped snubber used in output rectifier of phase-shift full-bridge zvs converters. IEEE Trans Industr Electron 45(2):358–359
7. Huang CK, Nien HH, Changchien SK et al. (2008) An optimal designed RCD snubber for DC–DC converters. In: Proceedings of the 10th international conference on control, automation, robotics and vision. pp 2202–2207
8. Thong WK, Pollock C (2002) A novel low-cost RCD snubber for bifilar-wound motors. IEEE Trans Ind Appl 38(3):688–694

9. Meng P, Chen H, Zheng S et al. (2010) Optimal design for the damping resistor in RCD-R snubber to suppress common-mode noise. In: Proceedings of the IEEE 25th applied power electronics conference and exposition (APEC), pp 691–695
10. Chen J, Diao L, Ruan B et al (2011) Suppression and modeling analysis of auxiliary converter rectifying voltage output. Trans China Electrotechnical Soc 26(1):71–76 (in Chinese)
11. Wang P, Zheng J, You J (2009) Improved design and simulation of snubber circuits for IGBT inverters. Electr Meas Instrum 46(10):67–71 (in Chinese)

Chapter 10
Research on Parallel Characteristics of Lithium Iron Phosphate Batteries for Dual Electric Multiple Units

Yanru Zhang, Jiuchun Jiang, Weige Zhang, Wei Shi, Zeyu Ma and Fangdan Zheng

Abstract The performance of power lithium ion battery pack in parallel will be further degraded due to the inconsistency of the cells. Under different working conditions, battery pack in parallel reflects different charging and discharging characteristics. In this paper, based on the series-parallel simulation platform, the actual current of parallel battery pack was obtained from the power characteristic of dual electric multiple unit (DEMU). The charging and discharging characteristics of parallel connection for Lithium iron phosphate ($LiFePO_4$) battery batteries with constant current and the loop current phenomenon under different state of charge (SOC) were investigated combined with the practical charging and discharging tests in the laboratory, which are helpful to get the main causes of aging of battery pack in parallel. An optimized usage range of SOC is proposed which provides the basis for battery management and prolongs the battery life cycle.

Keywords Dual electric multiple unit · Lithium iron phosphate battery · Parallel connection · Disequilibrium currents

10.1 Introduction

The global energy problem makes more and more countries pay attention to the development of new energy industries [1]. In recent years, China's railway industry has made considerable progress, however, the non-electrified railways still call for a certain proportion, limiting the development of rail transport. Compared to the accelerated pace of commercialization of electric vehicles (EV), the power battery for DEMU is limited to the auxiliary power only. In order to

Y. Zhang (✉) · J. Jiang · W. Zhang · W. Shi · Z. Ma · F. Zheng
National Active Distribution Network Technology Research Center,
Beijing Jiaotong University, 100044 Beijing, China
e-mail: 12121583@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_10,
© Springer-Verlag Berlin Heidelberg 2014

solve the above problems, the research related to DEMU has received more attention. DEMU uses two power modes: In the electrified areas DEMU is powered by the catenary, at the same time the battery pack is also charged; in the non-electrified areas DEMU is powered by the power lithium-ion battery pack. This power supply mode realizes cross-line operation between electrified and non-electrified railway.

However, the mode of DEMU powered by batteries in the non-electrified areas means high demands on the battery pack. First of all, high energy density is useful for the weight diminution of bodywork. Secondly, the complex operating environment requires better rate characteristics. Thirdly, higher voltage of single cell can reduce the number of series battery and reduce the requirements on the consistency of cells. Finally, long cycle life reduces the replacement cycle as well as the cost. Lithium-ion battery has the advantages of being light weight, small volume, high voltage, long life, high charge and discharge efficiency, and no pollution [2], therefore it is one of the most suitable power sources which can be used in DEMU.

Usually, batteries parallel into packs to provide a higher capacity, but the presence of inconsistency and short-board effect causes performance deviation [3]. Furthermore, over-charged and over-discharged are common for lithium-ion battery, so irrational use will lead to rapid decline in capacity [4]. This paper first analyzes the lithium-ion battery working current condition on the basis of DEMU's operational characteristics. Then the charging and discharging characteristics of parallel connection for $LiFePO_4$ batteries under previous condition, the loop current phenomenon, and the difference of capacity under different SOC are investigated combined with the practical charging and discharging tests in the laboratory. Finally, this paper has proposed an optimal working interval for battery packs.

10.2 Working Condition of $LiFePO_4$ Battery for DEMU

Other than electric vehicle (EV), DEMU generally moves at a constant speed except during starting and braking, therefore, the working condition of batteries in DEMU is simpler than batteries in EVs. This paper takes an EMU whose formation is two motors and one trailer (2M1T) as the research objective, which accesses to electricity from AC25 kV pantograph and power lithium-ion battery packs. The maximum speed is 120 km/h when the motor is supplied by the battery, while speeding up to 160 km/h under overhead line system. The maximum driving range is 200 km and the axle load is less than 17 tons. Maximum rated load (AW2) is 186 tons and overload (AW3) is 204 tons [5]. Adequately considering the power and energy output requirements, every motor is equipped with two battery packs which contain four cells in parallel and 300 parallel batteries in series. So every formation has four battery packs, and the rated voltage is 960 V. This section analyzes operation characteristics of the battery from acceleration, uniform motion coasting, and braking three phases.

10.2.1 Acceleration

The tractive characteristic curve of electric multiple unit (EMU) includes three stages: constant torque, constant power, and nature characteristic. During the constant torque stage, the output power increases linearly with velocity, and the acceleration time is 24.5 s; during the second stage, the output power is the same until the speed reaches 120 km/h, requiring 86.3 s; during the nature characteristic, tractive force has the inverse ratio with the square of velocity, but because the maximum speed is 120 km/h when battery supplies, the third stage does not exist. Figure 10.1a shows the tractive power characteristic curve of DEMU.

In a common scene, the battery's power is mostly enough when the DEMU starts. The parallel battery current during accelerated stage is shown in Fig. 10.1b, according to the battery simulation platform. The initial SOC has been set to 95 %. The current is near-linearly increasing during the constant torque stage, but because of the short period of time, the difference between cells are not reflected. During the constant power stage, the voltage of battery pack in series slightly decreases from 942.1 to 924.5 V due to discharging, which causes the tiny increase of current, from 332.2 to 333.84 A. This incredibly small variation is negligible when shared by four cells. Therefore, the current can be regarded as constant maintaining at 1.1 C.

10.2.2 Uniform Motion and Coasting

DEMU moves at a constant speed most of the time because acceleration and braking is seldom. Uniform motion is the most common condition. No doubt the voltage of battery pack will decrease with the constant power discharging, however, the voltage has tiny change during 90 to 10 % SOC for $LiFePO_4$ battery, and the current can be considered as constant. The output power is 330 kW [5] during uniform motion, and adding to auxiliary power supply (APS) demand, the discharging rate is 0.35 C.

When the DEMU is idle running relying on the inertia, the battery rests without charging and discharging. During this time, the loop current will balance the difference between cells because of the inconsistency, and the size and continuous time of current is related to SOC of battery.

10.2.3 Braking

The DEMU motor converts kinetic energy into electrical stored in batteries during braking. The braking characteristic curve has two stages: the braking power is constant until the first turning point which continues till 106.3 s. Then it turns to

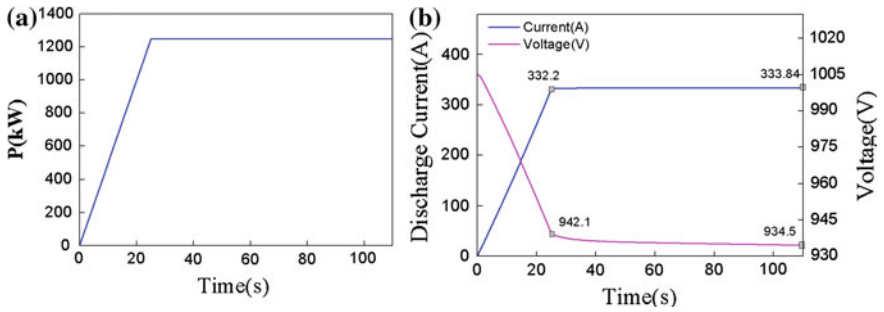


Fig. 10.1 **a** Tractive power characteristic curve of DEMU. **b** The battery current and voltage changes during acceleration

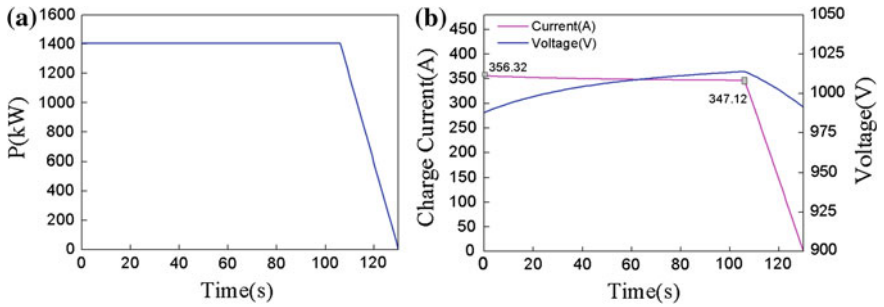


Fig. 10.2 **a** Braking power characteristic curve of DEMU. **b** The battery current and voltage changes during braking

constant torque stage, and the power linearly decreases within 24.3 s. Figure 10.2a shows the braking power characteristic curve.

In a common scene, the battery power is mostly lower when the DEMU brakes. The parallel battery current during braking stage is shown in Fig. 10.2b. The initial SOC has been set to 10 %. During the constant power stage, the voltage of battery pack in series increases due to charging, and the current slightly decreases from 356.32 to 347.12 A. So the current maintains the same on the whole and the charging rate is 1.2 C. The current is near-linearly decreasing during the constant torque stage, but because the batteries are scarcely possible to be full, the lesser current does not give rise to over-charge.

10.3 Parallel Characteristics of $LiFePO_4$ Batteries

Through the condition analysis during acceleration, uniform and idle running, and braking stages, it can be seen that, differing from the current crossfade for batteries in EVs, the current of battery pack for DEMU is mostly constant. In this paper, two

$LiFePO_4$ cells in parallel are tested to be easy to analyze the change in disequilibrium currents and the causes. As mentioned above, the charging and discharging rate is 0.35–1.2 C, so in this section, the defined rate is 0.5 C and the selected battery is ATL 60 Ah battery.

10.3.1 Constant Current–Constant Voltage Charging

The batteries for DEMU are constant current charged within a short time during braking and it will be fully charged in constant current–constant voltage method after running.

Figure 10.3 shows the change of charging disequilibrium currents for two $LiFePO_4$ cells numbered 1 and 2. The record of disequilibrium currents includes three processes: constant current, constant voltage, and rest. At the beginning of charging, the current of battery one (B1) was larger. As the batteries were charged, the difference became reductive and then increscent, which resulted from the OCV difference between the two cells due to incomplete rest. B1 was fully discharged whose OCV is lower resulting in larger charging current. The OCV varies greatly between 0 and 10 % SOC, so the OCV of B1 increased quickly, and then the difference of current became smaller. When the two cells reached plateau region, the difference of impedance began to emerge and the current differential continued to increase. The two currents just coincided at cut-off voltage, reflecting B1 is closer to full with OCV rising faster and the ability of accepting current dropping. What is worse, if the disequilibrium currents coincide before cut-off voltage, the OCV of B1 will continue rising and the current will decrease rapidly, leading to large current of battery 2 (B2) which may be beyond the normal tolerant range. The two currents decrease at constant voltage period and the current of B2 is 10.7 A at cut-off current, while the current of B1 is 1.3 A which is far less than set cut-off current. So B1 is over-charged. The large loop current states that the OCV difference was very large at the end of charging, and B2 is still in plateau region. So B1 charged B2. According to the data, changing capacity of B1 is 4.7 Ah more than that of B2, and the lager the changing rate, the more the difference will be. Therefore, this charging method can chronically result in over-charge of one cell and over-current of the other for many times.

10.3.2 Constant Current Discharging

The batteries for DEMU work under constant current when discharging except for current changes in a short time during constant torque acceleration.

Figure 10.4 shows the change of discharging disequilibrium currents for the two $LiFePO_4$ cells. First, the difference in currents decreased which still resulted from OCV difference because of different termination SOC. When the two cells

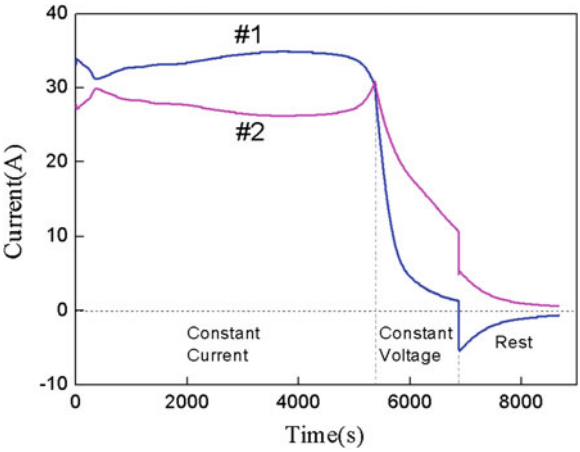


Fig. 10.3 Charging disequilibrium currents for two cells

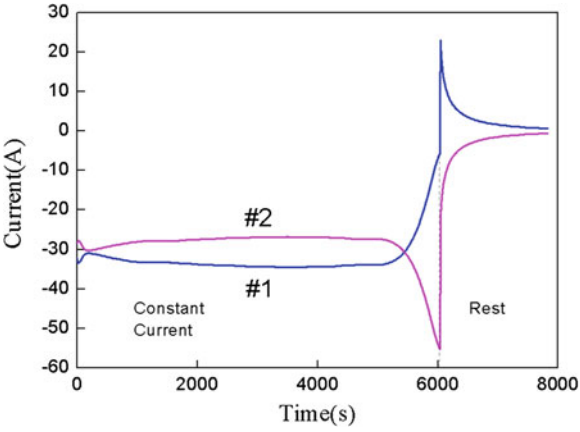


Fig. 10.4 Discharging disequilibrium currents for two cells

reached plateau region, the discharging current of B1 is larger as well as the charging current because of the smaller impedance. The two currents coincided after discharging for 1.5 h. The discharging ability of B1 dropped and the discharging current decreased rapidly until the voltage reached the cut-off voltage. This moment, B1 was over-discharged as it reached cut-off voltage with a small current, while B2 whose current increased to two times was over-current due to the small current of B1. If the discharging rate is slightly larger, the over-current may damage B2. Finally, B1 discharged 54 Ah while B2 discharged 48.6 Ah. The large loop current still states that B1 was in low-end period while B2 was still in plateau region. Therefore, this discharging method can chronically result in over-discharge of one cell and over-current of the other one for many times.

10.3.3 Loop Current Under Different SOC

During coasting period, after running or after full charging, the batteries rest. At these moments, loop current will exist resulting from different OCV. The loop current will decrease the difference of SOC and its size depends on the difference of OCV. In this section, a series of charging and discharging tests was performed with different usage range of SOC, and then the battery will adequately rest to observe the size of loop current with battery management system (BMS). The computational capacity of battery pack is the actual capacity.

Figure 10.5 and Table 10.1 show the loop current with different usage range of SOC (the end of loop current refers that the BMS cannot detect current due to precision problem, and in fact the loop current still exists). At the emptying state after discharging, the peak loop current was about 0.3 C because of the large OCV difference. The difference in discharging capacity was 5.4 Ah before rest and loop current balanced it to 2.3 Ah. When the battery pack was discharged to 5 % SOC, the loop current decreased to one third, and the difference in discharging capacity was from 9.6 to 7.1 Ah. When the batteries discharged to 10 %SOC, BMS cannot detect loop current, explaining the small size of current. From Table 10.1, it can be seen that B1 discharged 10 Ah more than B2, showing large difference of SOC. But because the OCV of $LiFePO_4$ battery varies little in plateau region, it results in little loop current, leading to difficulty to balance the difference of SOC.

The charging result is the same as discharging, which are shown in Fig. 10.6 and Table 10.2. The loop current could be detected at 100 and 95 % SOC while it was not detected when charging to 90 % SOC. The large loop current balanced more difference capacity at 100 % SOC.

The battery pack needs a long time to balance difference of SOC on account of little loop current. Therefore, every cell works under difference range of SOC indeed. According to the analysis of parallel characteristics of constant charging and discharging and the loop current, it can be seen that, though the working current condition for DEMU is simple without current variation, the charging and discharging method is extremely easy to cause the accumulation of disequilibrium current because of the parameters of cells, resulting in large difference of SOC, which is difficult to balance because of little difference of OCV.

10.4 Optimized Usage Range of SOC

On account of the inconsistency of cells and the special working condition of battery for DEMU, the full charging and discharging will result in some cells over-charged and over-discharged, and some over-current, which accelerates aging of batteries. What is more, if the rate is larger, the over-current may cause danger. Therefore, reasonable control on the parallel battery has an important significance to prolong the life of the battery.

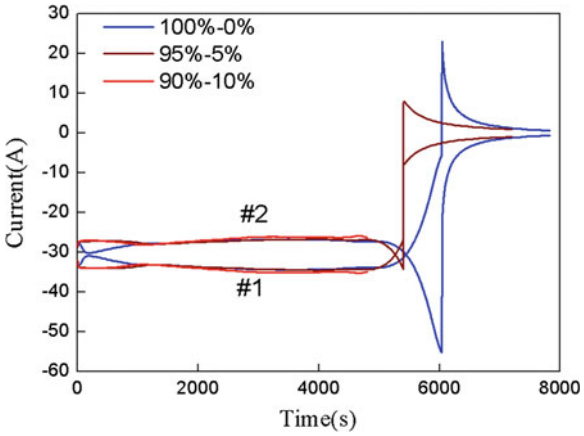


Fig. 10.5 The loop current of different usage range of SOC after discharging

Table 10.1 Loop current and difference of discharging capacity before and after rest

Usage interval (%)	Peak of loop current (A)	Before rest (Ah)	After rest (Ah)
100–0	22.9	5.4	2.3
95–5	7.5	9.6	7.1
90–10	0	10	10

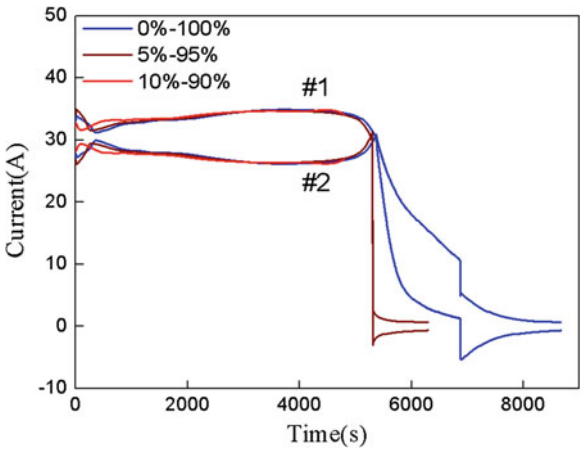


Fig. 10.6 The loop current of different usage range of SOC after charging

Table 10.2 Loop current and difference of charging capacity before and after rest

Usage interval (%)	Peak of loop current (A)	Before rest (Ah)	After rest (Ah)
0–100	5	4.7	2.8
5–95	7.5	9.3	8.7
10–90	0	9.5	9.5

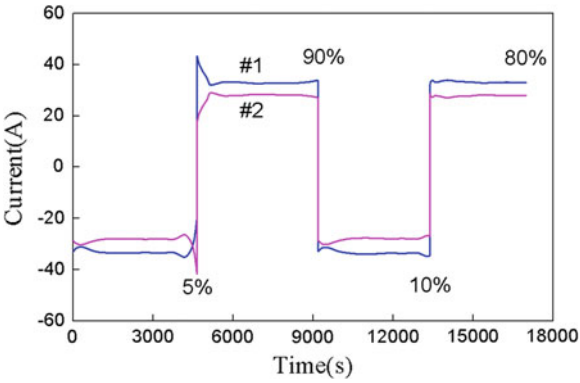


Fig. 10.7 Consecutive charging and discharging disequilibrium currents

Figure 10.5 shows that when the parallel pack is discharged to 0 or 5 % SOC, the disequilibrium currents will cross which states that one cell's OCV decrease rapidly and the other has the danger of over-current. In Fig. 10.6, the currents coincide at 95 % SOC and the OCV of one cell will increase rapidly. Meanwhile, many studies have shown that when the battery comes through the platform to a high voltage, oxidizing reaction happens between the electrolyte and the anode, while the electrolyte loses electrons, and reduction reaction happens between electrolyte and the cathode or SEI, while the electrolyte gets electrons, resulting in the loss of active materials and electrolyte [6], making battery decline rate accelerate. The best way is to control the battery usage range of SOC between two intersections. Although it slightly lowers battery capacity utilization, however, on one hand, it avoids over-charged, over-discharged, or over-current phenomenon, and on the other hand, it also avoids the battery working in high-end area, effectively prolonging the life of battery.

Figure 10.7 shows the consecutive charging and discharging disequilibrium currents. If the battery pack is discharged to 5 %, the charging current of B1 is larger instantaneously without rest, while this phenomenon does not appear when discharged to 10 % SOC without intersection. From this figure, it can be seen that the larger charging one will have larger discharging current within the optimized usage range, which ensures the usage range will not come up excursion, and the problem of one cell fully charged first and the other fully discharged first can be avoided.

10.5 Conclusion

This paper analyzes the power characteristics of batteries for DEMU under three conditions which are locomotive acceleration, constant speed and coasting, and braking, using the battery pack simulation platform to get parallel batteries' current conditions. Then according to operation characteristics of battery pack, the

disequilibrium currents of constant current-constant voltage charging and constant-current discharging and loop current of different SOC usage interval are shown. Finally, according to the analysis of the parallel characteristics this paper proposes an optimal working interval of battery pack.

Studies show that: (1) The batteries for DEMU generally work under constant current charge and discharge condition; (2) This condition will cause current difference accumulation among different cells. In the process of charging, some batteries are over-charged since the cut-off current is very small, while some batteries are over-current because of smaller OCV later period of charging process. In the discharging process, some batteries are over-discharged because the cut-off voltage is reached with minimal current, and some batteries are over-current because of larger OCV at the end of discharging process. At the same time, the loop current in platform area is too small to balance the difference in SOC among different cells; (3) Over-charged and over-current can be effectively avoided if the working interval of SOC is controlled between two current intersections. Furthermore, although it is difficult to balance the difference in SOC, this difference is essentially fixed and will not shift with changing of cycle index. This paper studies the major factors which influence battery aging, accordingly proposing the optimal usage range of SOC and control methods of battery packs, which is of great significance to extend the battery life.

Acknowledgments The work is supported by National Key Technology R&D Program (2013BAGL1B00).

References

1. Dubarry M, Vuillaume N, Liaw BY (2009) From single cell model to battery pack simulation for Li-ion batteries. J Power Sources 186(2):500–507
2. Zhang W, Shi W, Jiang J et al (2012) Numerical simulation technique of series-parallel power Lithium-ion battery. Power Syst Technol 36(10):70–75 (in Chinese)
3. Wei F, Mi C (2008) Study on parallel charge of Cd-Ni batteries. Telecom Power Technol 25(5):49–53 (in Chinese)
4. Li X, Hsu S-p, Lin C-L, Wu J-Y (2012) Balancing charge/discharge management for series/parallel battery packs. In: Proceedings of the 2012 7th IEEE conference on industrial electronics and applications, pp 613–618
5. Wang Z, Ma Z (2013) Battery system design of lithium ion battery for dual electric multiple units. Chinese Railw (in Chinese)
6. Gong M, Shi W, Zhang Y et al (2013) Operating conditions control of large format $LiMn_2O_4$ battery for electric bus. J Jilin Univ (Engineering and Technology Edition) (in Chinese)

Chapter 11
Discussion on the Application of Energy-Saving Traction Power Supply Device

Zhiwei Han, Gang Zhang, Zhigang Liu and Fuqiang Mu

Abstract The energy of traditional traction power supply rectifier unit, which is based on the diode rectifiers, can only be transmitted in one-way, so the regenerative braking energy generated by train has to be consumed with resistance. It not only affects the regenerative braking capacity of the train, but also brings about a huge waste of energy, and leads to some problems like the rise of tunnel temperature, increase of the burden of environmental control system and corresponding energy consumption. A new energy-saving traction power supply device, which is installed in Beijing metro line 10 (two engineering), cannot only make regenerative braking energy of the train back to the grid, but also has the function of traction power supply, can reduce the fluctuation of DC voltage, as well as can achieve reactive power compensation to the medium-voltage ring network. Its application prospect is pretty good. The features, working principle, key technology, and engineering applications of the energy-saving traction power supply device are introduced in this paper and field test data are given.

Keywords Urban rail transit · Traction power supply · Braking energy feedback · Rectifier

Z. Han (✉)
Beijing Railway Construction and Management Co., Ltd., 100034 Beijing, China
e-mail: gzhang@bjtu.edu.cn

G. Zhang · Z. Liu
School of Electrical Engineering, Beijing Jiaotong University, 100044 Beijing, China

F. Mu
Beijing Qiansiyu Electrical Electric Co. Ltd., 100082 Beijing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_11,
© Springer-Verlag Berlin Heidelberg 2014

11.1 Introduction

Urban rail transit, which has many advantages, like high speed, large capacity, low noise, comfort and high safety, energy-saving and less pollution, becomes the first choice to solve the expanding problem of urban traffic in many nations.

At present, diode rectifier is still widely used in the traditional traction power supply rectifier unit of urban rail transit traction system. Because its energy can only be transmitted in one-way, the regenerative braking energy generated by train has to be consumed with resistance. It not only affects the regenerative braking capacity of the train, but also brings about a huge waste of energy, and leads to some problems like the rise of tunnel temperature, increase of the burden of environmental control system and corresponding energy consumption. How to ensure the regenerative braking capacity of the train being exerted effectively and achieve effective use of the braking energy of the train is the problem concerned by the industry for a long term [1–3].

The energy-saving traction power supply device, whose distinguishing feature is that the device can work in four quadrants, is introduced in the paper. When it works in inverter mode, the energy feedback traction power supply device could feed train regenerative braking energy back to AC grid to avoid energy consumption on resistance and conserve energy. And when working in rectifier mode, the device can provide traction power, which could maintain the DC voltage in some extent and improve the quality of DC power supply. Moreover, when it works on reactive compensation mode, the device is able to accomplish the reactive power compensation of the AC grid and improve system's power factor.

The features, working principle, key technology, and engineering applications of the energy-saving traction power supply device are introduced specifically in this paper [4–7]. In addition, part of field test data are given.

11.2 The Overview of Traction Power Supply System of the Urban Rail Transit

The simplified schematic diagram of traditional traction power supply system is shown in Fig. 11.1. SVG represents static reactive power generator. And Rd represents consuming device braking energy, namely braking resistor.

The rectifier unit shown in Fig. 11.1 is based on a 12-pulse or 24-pulse diode rectifier. Low cost and mature application are its advantages, but there are also some significant shortcomings.

a) Serious waste of regenerative braking energy.
b) Huge fluctuation range of DC output voltage.
c) Unable to regulate power factor of the AC system, and additional SVG is needed.
d) Complex system and taking up a large space.

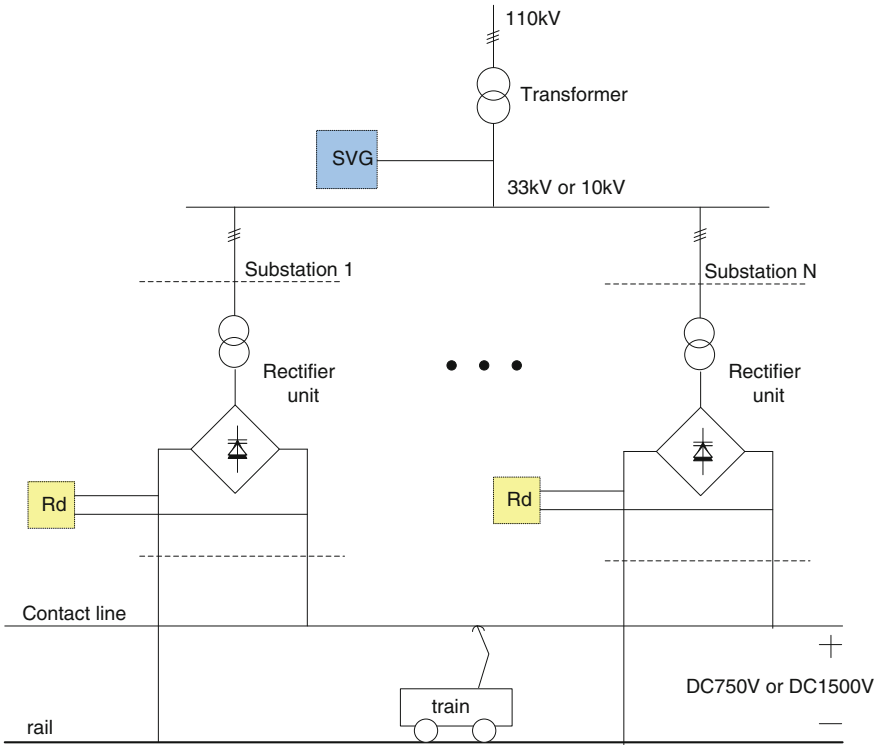


Fig. 11.1 Simplified schematic diagram of traditional traction power supply system

11.3 Features of the Energy-Saving Traction Power Supply Device

A kind of energy-saving traction power supply device, which is based on high-power PWM Rectifier, is discussed in this paper. The device owns many fine features as follows:

a) Bidirectional flow of energy

When the train braking, excess braking energy can directly feedback to medium-voltage network through the traction power supply device and energy can be saved. When the train is driving, diode rectifier units and energy-saving traction power supply device supply power to the train, simultaneously. It could help stabilize DC voltage and increase power supply distance.

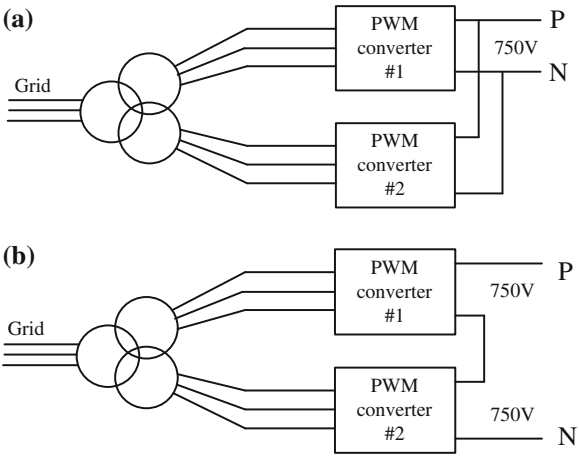


Fig. 11.2 Topology of traction power supply system with energy feedback. **a** Used for 750 V system, **b** 1500 V system

b) Stabilizing DC output voltage
By detecting the size of DC output voltage and controlling the size and direction of energy, which is converted by energy-saving traction power supply device, the adjustment of DC output voltage can be realized, and thereby DC output voltage could be stabilized.

c) Realizing reactive power compensation
The power factor of the energy-saving traction power supply device can be adjusted easily, so it can be used to compensate for the low power factor of the off-peak periods of the medium-voltage ring network.

11.4 Configurations, Working Principle, and Control Strategy of the Proposed Device

11.4.1 Configuration and Working Principle

The main topology of traction power supply system with energy-saving is shown in Fig. 11.2.

The main circuit of PWM rectifier and its single-phase equivalent circuit are shown in Fig. 11.3.

When the grid voltage e_a and the AC inductance L can be determined, by controlling the magnitude and phase of U_a, it is possible to control the magnitude and phase of the current i_a, and thereby to control the active and reactive power conducted by the converter.

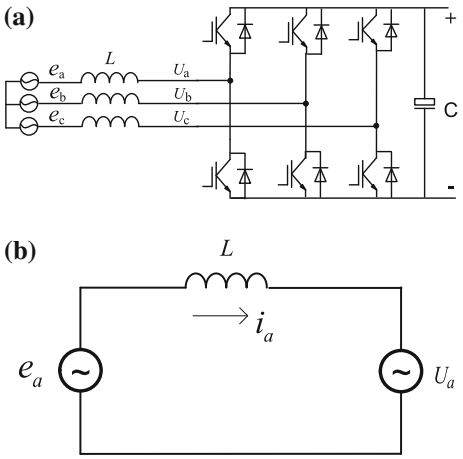


Fig. 11.3 Main circuit and its single-phase equivalent circuit for PWM converter. **a** Main circuit, **b** The single-phase equivalent circuit

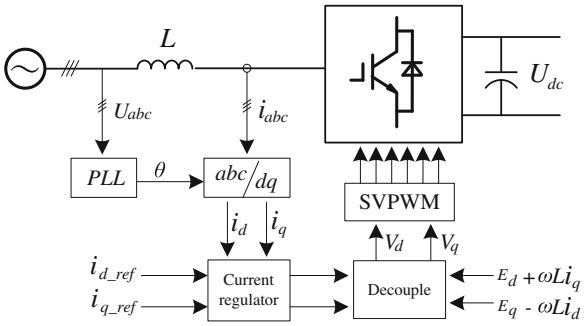


Fig. 11.4 Current closed-loop control scheme based on synchronous rotating coordinate system

11.4.2 Control Technology of Current

The mathematical model of PWM rectifier could be transformed to the synchronous rotating *d-q* coordinate system, and D-axis is coaxial with the grid voltage vector. And then, the control of active and reactive power can be realized by controlling i_d and i_q, respectively.

As shown in Fig. 11.4, current closed-loop control scheme [8], which is based on the synchronous rotating coordinate system, is the basis of the control of the PWM converter.

11.4.3 Interleaved SVPWM Technology

In order to reduce the current harmonics injected into the grid, interleaved SVPWM technology is able to be adopted in the multiple converter system

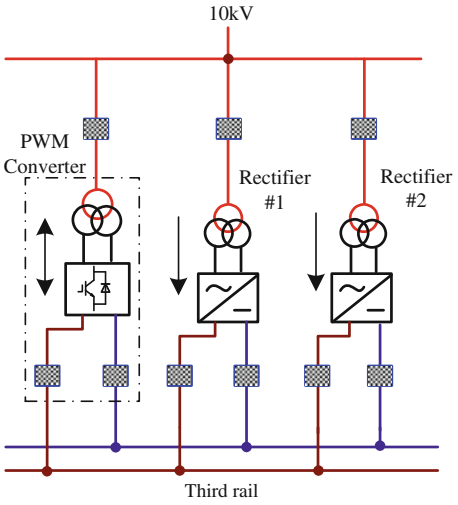


Fig. 11.5 Scheme of application technology in filed

constituted by a plurality of inverter [9]. By shifting the sampling time of SVPWM modulation of the respective converter by 1/*N* switching cycle from each other, it can offset harmonics in the vicinity of the switching frequency. And equivalent switching frequency of system upgrades to Nf_s.

11.5 The Scheme of Application Technology and Field Experiments

In order to verify the performance of energy-saving traction power supply device, demonstration applications of the device is conducted in Beijing Metro Line 10 (two engineering).

Specific applications is shown in Fig. 11.5, the energy-saving traction power supply device, which could transmit energy bi-directionally as indicated by arrows, is connected with the AC and DC bus of diode rectifier unit (rectifier #1, rectifier #2).

The rated capacity of diode rectifier unit is 2200 kW, and its load capacity is Class VI load. The rated capacity of energy-saving traction power supply device is 2 × 1 MW. The rated voltage of AC grid is 10 kV and the range of DC voltage is 750–1000 V.

Figure 11.6 shows the AC grid voltage and AC current waveform of the device when it works at inverter mode at rated power. CH1 is grid voltage waveform and CH2 is AC current waveform. By using ISVPWM technology, the harmonic of the AC current can be controlled at a lower level.

Figure 11.7 shows the DC-side voltage waveform when energy-saving traction power supply device and the diode unit (12-pulse) works in parallel. CH1 is the

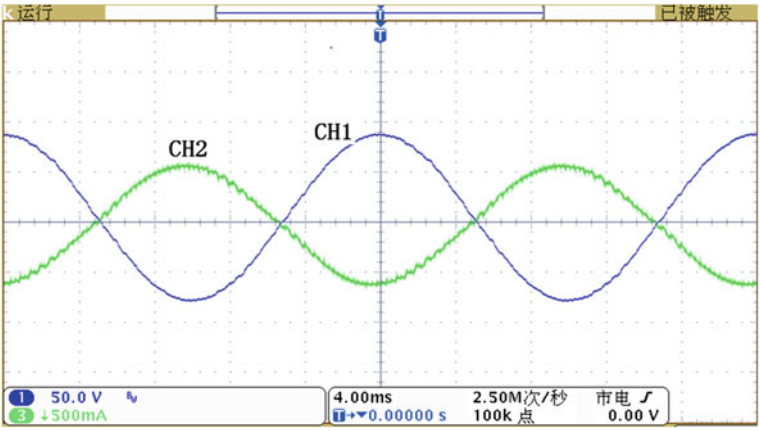


Fig. 11.6 Grid voltage and AC current waveform

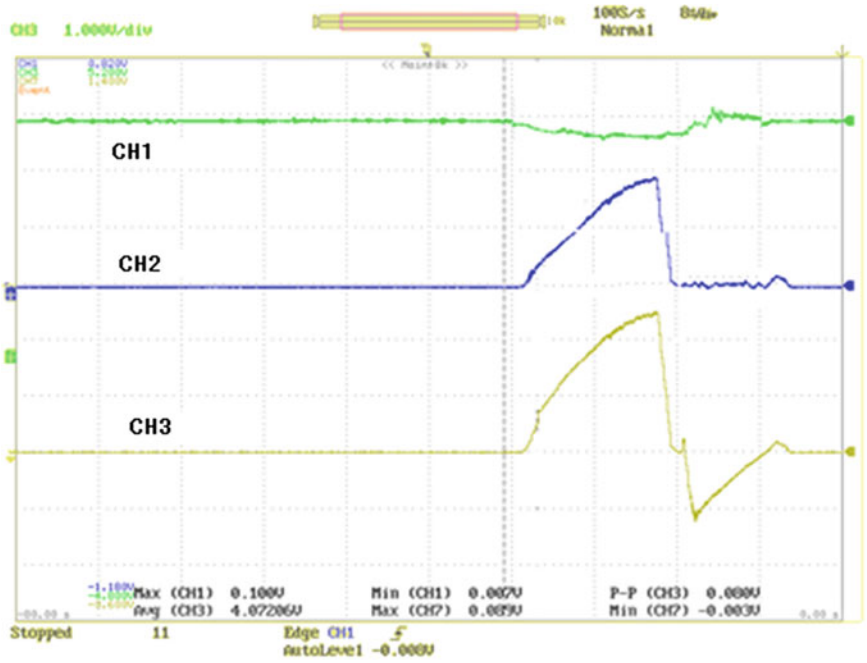


Fig. 11.7 DC current waveforms of the two sets

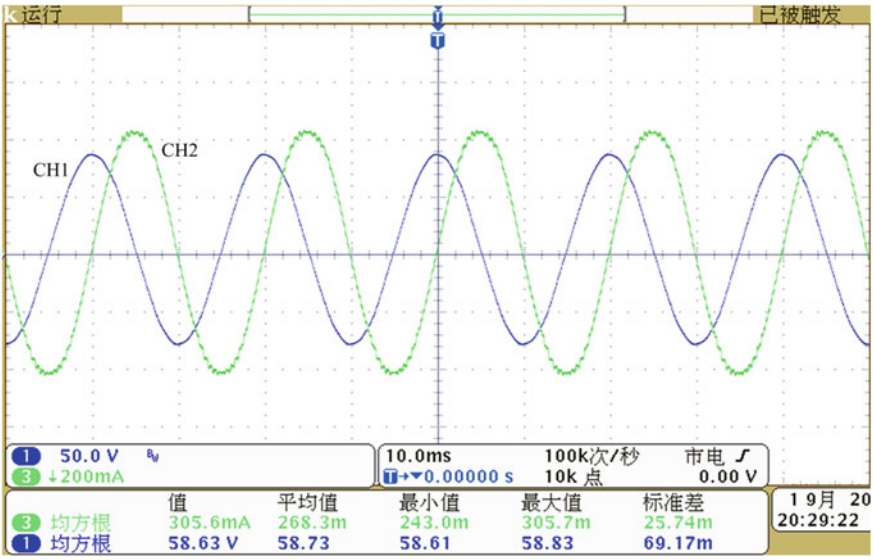


Fig. 11.8 Reactive power compensation test waveform

DC bus voltage in traction substation. CH2 is the DC current of the diode rectifier unit. And CH3 is the DC current of energy-saving traction power supply device.

It can be seen from Fig. 11.7, when the train works in traction, the diode rectifier unit and energy-saving traction power supply device can provide traction energy together to the train. And when the train brakes, the device will let renewable braking energy feedback into the grid. Thus, coordinating control can be realized.

Figure 11.8 shows the waveforms of the reactive power compensation test. And CH1 is the voltage of medium-voltage grid. CH2 is reactive current for supply device.

Table 11.1 shows everyday's running data of substation after energy-saving traction power supply device starts to operate. And total traction power, which is provided by the substation (rectifier #1, rectifier #2, PWM Converter), is on behalf of total energy and is used in train traction. While total feedback energy is on behalf of the regenerative braking energy of train, which is back to AC grid though energy feedback traction power supply device everyday. Energy-saving rate is defined as total feedback energy/total traction power in percentage.

As shown in Table 11.1, the average saving rate of energy-saving traction power supply device in each day of a week is 17.52 %, and the highest power saving of single-day is 1898 kWh. Thus, the effect of energy saving is very significant.

Table 11.1 Substation data in a day

Time	Traction power/kwh	Feedback power/kwh	Energy-saving rate (%)
0:00	8	0	0.00
1:00	6	0	0.00
2:00	7	0	0.00
3:00	7	0	0.00
4:00	27	11	40.74
5:00	246	105	42.68
6:00	666	60	9.01
7:00	796	54	6.78
8:00	804	122	15.17
9:00	766	92	12.01
10:00	520	109	20.96
11:00	482	95	19.71
12:00	486	104	21.40
13:00	453	98	21.63
14:00	458	127	27.73
15:00	433	79	18.24
16:00	573	90	15.71
17:00	759	59	7.77
18:00	826	73	8.84
19:00	751	59	7.86
20:00	467	73	15.63
21:00	413	98	23.73
22:00	345	140	40.58
23:00	194	159	81.96
Total	10493	1807	17.22

11.6 Conclusions

Current situation and existing problems of traction power supply system for urban rail transit are analyzed and a new type of energy-saving traction power supply device is introduced in the paper. And then its features, working principle and engineering application plan are introduced. In addition, field test has been carried on. The result shows that the device has the ability of rectifier, inverter, and reactive power compensation and the effect of energy-saving is significantly when the device is applied into the urban rail transit traction power supply system.

Acknowledgments This work was supported in part by Beijing Municipal Traffic Commission under Grant 2012kj-029x and Beijing Municipal Education Commission E11H00011.

References

1. Koseki T, Okada Y, Yonehata Y, Sone S (2004) Innovative power supply system for regenerative trains
2. Chuang H-J, Chen C-S, Lin C-H et al (2005) Optimization of inverter placement for mass rapid transit systems using genetic algorithm. In: Transmission and distribution conference and exhibition: Asia and Pacific, 2005 IEEE/PES, pp 1–6
3. Suzuki T (1982) DC power-supply system with inverting substations for traction systems using regenerative brakes. IEEE Proc Electr Power Appl 129(1):18–26
4. Zhang G, Liu Z, Chen D et al (2007) Control and implementation of reversible multi-modular converter with interleaved space vector modulation. In: International conference on mechatronics and automation, ICMA 2007, pp 3463–3468 (in Chinese)
5. Zhang G, Liu Z, Wang L et al (2008) A DC traction power supply converter system with very low harmonics for railway vehicles. In: Vehicle power and propulsion conference, VPPC '08. IEEE, pp 1–6 (in Chinese)
6. Wang L, Zhang G, Shen M et al (2009) A novel traction supply system for urban rail transportation with bidirectional power flow and based on PWM rectifier. In: Energy and environment technology, ICEET '09. International Conference, pp 40–43 (in Chinese)
7. Wang L, Liu Z, Zhang G et al (2008) A novel traction power supply system for urban rail transportation. In: Vehicle power and propulsion conference, VPPC '08. IEEE, pp 1–5 (in Chinese)
8. Dai K, Liu P, Xiong J et al (2003) Comparative study on current control for three-phase SVPWM voltage-source converter in synchronous rotating frame using complex vector method. In: Power electronics specialist conference, PESC '03. 2003 IEEE 34th Annual, pp 695–7002 (in Chinese)
9. Keliang Z, Danwei W (2002) Relationship between space-vector modulation and three-phase carrier-based PWM: a comprehensive analysis [three-phase inverters]. IEEE Trans Industr Electron 49:186–196

Chapter 12
Research on Applicability of Lithium Titanate Battery for Low-Floor Vehicles

Sijia Liu, Jiuchun Jiang, Zhanguo Wang, Zeyu Ma, Wen Chen and Ting Zhao

Abstract Based on the application characteristics of lithium battery for low-floor vehicles, lithium titanate battery has advantages of application over other types of lithium batteries. The reasonable test scheme of lithium titanate battery can be formulated on the basis of existing test platform of single battery. Rate characteristics and low temperature characteristics of lithium titanate battery can be obtained through charge and discharge tests of two lithium titanate batteries at different temperatures and rates. Combined with characteristics of lithium titanate battery and operation characteristics of low-floor vehicles, the applicability of lithium titanate battery for low-floor vehicles is analyzed.

Keywords Lithium titanate battery · Low-floor vehicles · Rate characteristics · Low temperature characteristics

12.1 Introduction

With the rapid development of social economy and city urbanization, urban rail transit in China has developed rapidly. Among which, the light rail transit has gradually become the focus of various countries due to its characteristics of unique structure, humanization, and short construction period, low cost, and low noise [1–3]. With the support of "Research on 100 % low-floor vehicles" which belongs to China's 11th "five-year" science and technology key project, CNR Changchun Railway Vehicles Corporation Limited has successfully developed China's first 100 % low-floor light rail vehicle with independent intellectual

S. Liu (✉) · J. Jiang · Z. Wang · Z. Ma · W. Chen · T. Zhao
National Active Distribution Network Technology Research Center (NANTEC),
No. 3 Shang Yuan Cun, Beijing, China
e-mail: 13121433@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_12,
© Springer-Verlag Berlin Heidelberg 2014

Table 12.1 The performance parameters of lithium-ion batteries and super capacitor

Performance	Lithium iron phosphate battery	Lithium manganese battery	Lithium titanate battery
Rated voltage (V)	3.2	3.7	2.3
The maximum continuous charge current (A)	10C	5C	10C
The maximum continuous discharge current (A)	25C	20C	10C
Working environment temperature-charge	0 °C–45 °C	0 °C–40 °C	−30 °C–55 °C
Working environment temperature-discharge	−20 °C–60 °C	−20 °C–55 °C	−30 °C–55 °C
Safety	Ordinary	Ordinary	Good
The cycle life (room temperature)	≥1500 times	≥2000 times	≥6000 times

property rights in cooperation with Beijing Jiaotong University and other organizations, which breaks the situation of dependence on imports [4].

Since lithium battery has advantages of being light weight, small volume, long service-life, high voltage, high charge and discharge efficiency, and no pollution, it is widely used in new energy vehicles and energy storage of grid [5]. According to the requirements of the ministry of science and technology, low-floor light rail vehicles need to increase the power battery and be adaptively reconstructed, using lithium battery as an energy storage element. It can not only supply power to the traction and auxiliary system of vehicles, but also supply power when vehicles enter the section with no overhead contact line in operation.

According to the energy and power demands of low-floor vehicles, the lithium batteries available mainly include power type lithium iron phosphate battery, power type lithium manganese battery, and energy type lithium titanate battery, the performance parameters of which are given in Table 12.1 [6].

It can be seen from the table that lithium titanate battery has the best temperature characteristics, as its charge temperature range is the same as discharge range. Besides, energy type titanate lithium battery is safer than that of these two kinds of power type batteries. Furthermore, the cycle life of lithium titanate battery is the longest. Therefore, energy type lithium titanate battery has advantages of application for low-floor vehicles.

The application of lithium titanate battery in the low-floor vehicles not only greatly reduces the construction cost of power grid and pollution to the environment, but also develops the application of low-floor vehicles.

12.2 Test Schemes of Lithium Titanate Battery

As an effective tool to understand the battery characteristics, experiment is of great significance for battery analysis.

Table 12.2 The basic parameters of measured batteries

Item	Parameter	Parameter
Battery manufacturer	A	B
Nominal capacity (Ah)	20	10
Rated voltage (V)	2.3	2.3
Upper limit voltage (V)	2.7	2.8
Lower limit voltage (V)	1.5	1.5
Lower limit current (A)	0.05C	0.1C
Packed form	Steel shell	The plastic film

The experimental platform is mainly composed of measured batteries, high-low temperature chamber, charge and discharge device, and monitoring computer. The high-low temperature chamber provides the test environment for measured batteries. The monitoring computer communicates with the charge and discharge device through the bus, which not only controls the charge and discharge device for the corresponding action on the tested batteries according to the programs, but also records the value of voltage, current, capacity, and other information at the same time. Besides, the computer provides the monitoring interface of battery information. The measured batteries are A-20 Ah and B-10 Ah energy type titanate lithium battery, whose basic parameters are shown in Table 12.2.

Test schemes mainly include test scheme of basic parameters and test scheme of basic performance. The test scheme of battery capacity is shown as follows:

(1) Stay 10 s, which ensures the synchronization of monitoring computer and the charge and discharge device;
(2) Charge the battery with 1C current until the voltage reaches upper limit voltage;
(3) Charge the battery with constant voltage until the current reaches lower limit current;
(4) Stay 1 h;
(5) Discharge the battery with 1C current until the voltage reaches lower limit voltage;
(6) Stay 1 h;
(7) (1)–(6) is a cycle, the test needs 4–5 cycles;
(8) (1)–(7) is carried out at each temperature.

Capacity test requires 4–5 charge and discharge cycles of the battery which ensures that the battery reaches a steady state. Then the accurate value of the battery capacity can be obtained.

The test scheme of OCV-SOC curve and ohmic resistance is shown as follows:

(1) Stay 10 s, which ensures the synchronization of monitoring computer and the charge and discharge device;
(2) Charge the battery with 1C current until the voltage reaches upper limit voltage or the charging capacity reaches 5 % of the capacity;

(3) Charge the battery with constant voltage until the current reaches lower limit current or the charging capacity reaches 5 % of the capacity;
(4) Stay 1 h;
(5) Discharge the battery with 8C current for 20 s;
(6) Stay 20 s;
(7) Charge the battery with 8C current for 20 s;
(8) (1)–(7) is a cycle, the test needs 20 cycles;
(9) (1)–(8) is the test of charging OCV-SOC curve and the charging ohmic resistance.

It takes 20 cycles for battery to reach 100 % of the capacity.

The test scheme of discharging process is shown as follows:

(1) Stay 10 s, which ensures the synchronization of monitoring computer and the charge and discharge device;
(2) Discharge the battery with 1C current until the voltage reaches lower limit voltage or the discharging capacity reaches 5 % of the capacity;
(3) Stay 1 h;
(4) Charge the battery with 8C current for 20 s;
(5) Stay 20 s;
(6) Discharge the battery with 8C current for 20 s;
(7) (1)–(6) is a cycle, the test needs 20 cycles;
(8) (1)–(7) is the test of discharging OCV-SOC curve and the discharging ohmic resistance.

It takes 20 cycles for discharging the battery empty.

The group of pulses is used to measure the ohmic resistance. Assuming the voltage of battery is U_1 when current is I_1 and the voltage of battery is U_2 when current is I_2,

$$R_\Omega = \Delta U/\Delta I = (U_2 - U_1)/(I_2 - I_1) \tag{12.1}$$

Since the purpose of the test of basic performance is to get the low temperature characteristics and rate characteristics of battery, temperatures and rates should be set as test variables, and other variables in the test should be set to a constant. According to the power limit of device and the actual environment of battery, the test temperatures and rates can be determined, which are shown in Table 12.3.

The test scheme of charging characteristics is shown as follows:

(1) Stay 10 s, which ensures the synchronization of monitoring computer and the charge and discharge device;
(2) Charge the battery with rate current until the voltage reaches upper limit voltage;
(3) Charge the battery with constant voltage until the current reaches lower limit current;
(4) Stay 2 h;

Table 12.3 The test temperatures and rates of batteries

Name	Temperature (°C)	Charging rates (C)	Discharging rates (C)
A	55, 25, 0, −10, −20, −30	1, 2, 3, 5, 6, 8	1, 2, 3, 5, 6
B	25	0.5, 1, 2, 4, 6, 8, 10	0.5, 1, 2, 4, 6, 8, 10
B	0	0.5, 1, 2, 4, 6, 8	0.5, 1, 2, 4, 6, 8, 10
B	−20	0.5, 1, 2, 4, 6	0.5, 1, 2, 4, 6, 8

(5) Discharge the battery with 1C current until the voltage reaches lower limit voltage;
(6) Stay 2 h;
(7) Discharge the battery with 0.2C current until the voltage reaches lower limit voltage;
(8) Stay 1 h;
(9) (1)–(8) is a cycle, which is carried out at each rate;
(10) (1)–(9) is carried out at each temperature.

The test scheme of discharging characteristics is shown as follows:

(1) Stay 10 s, which ensures the synchronization of monitoring computer and the charge and discharge device;
(2) Charge the battery with 1C current until the voltage reaches upper limit voltage;
(3) Charge the battery with constant voltage until the current reaches lower limit current;
(4) Stay 2 h;
(5) Charge the battery with 0.2C current until the voltage reaches upper limit voltage;
(6) Stay 1 h;
(7) Discharge the battery with rate current until the voltage reaches lower limit voltage;
(8) Stay 2 h;
(9) (1)–(8) is a cycle, which is carried out at each rate.
(10) (1)–(9) is carried out at each temperature.

12.3 Test Results and Analysis of Lithium Titanate Battery

With the analysis of test results, effects of temperatures and rates on the charging and discharging process can be obtained, namely low temperature characteristics and rate characteristics of the battery.

Table 12.4 The results of capacity test at different temperatures

Temperatures (°C)	Capacity (Ah)	
	A-20 Ah	B-10 Ah
55	21.3	11
25	20	10.5
0	16.6	9.1
−20	8.5	5.8

12.3.1 The Capacity Test at Different Temperatures

The results of capacity test at different temperatures are shown in Table 12.4. It is easy to be seen from the test results that the value of battery capacity is in reduction with the increase in temperature. For A-20 Ah battery, the capacity retention rate is only 43 % at −20 °C, while the capacity retention rate of B-10 Ah battery is 55 % at the same temperature. High capacity retention rate not only can provide more margins for the design of the battery system configuration of low-floor vehicles, but also reduce the cost and weight. Considering the capacity retention rate, B battery is better than A battery.

12.3.2 OCV–SOC Curve

The OCV–SOC curve of A battery is shown in Fig. 12.1, while that of B battery is shown in Fig. 12.2. The method for estimation of SOC in engineering application is current integral, which is modified by OCV-SOC curve [7]. It can be seen from the figures that the slope of OCV–SOC curve is large, and the charging OCV–SOC curve almost coincides with discharging OCV–SOC curves for B battery. Therefore, SOC estimation of lithium titanate battery is more easy and accurate. Then operation state of low-floor vehicles can be obtained. Considering the OCV–SOC curve, B battery is better than A battery.

12.3.3 The Charging Characteristics at Different Temperatures

The charging characteristics of A battery is basically the same as that of B battery. A battery still can charge with 3C current at −20 °C, and B battery still can charge with 2C current at −20 °C. Considering the wide range of working temperature for low-floor vehicles and the need of battery to absorb braking energy at low temperature as much as possible, the battery must have the ability to charge with high rate at low temperature. But the maximum charging rate at low temperature is

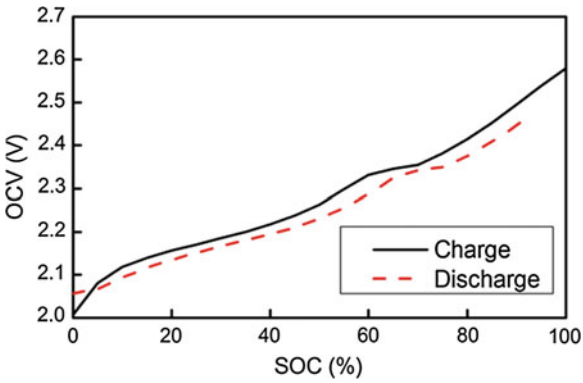


Fig. 12.1 OCV–SOC curve of A battery

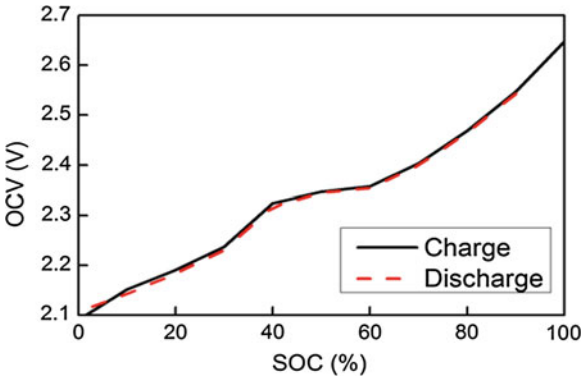


Fig. 12.2 OCV–SOC curve of B battery

obviously lower than that at room temperature. Therefore, heating of battery must be taken into account when recycling energy at low temperature.

12.3.4 The Discharging Characteristics at Different Temperatures

Taking discharging characteristics at 25 and 0 °C as examples, the discharging characteristics of A battery at 25 °C are shown in Fig. 12.3, while that of A battery at 0 °C are shown in Fig. 12.4. A battery still can discharge with 3C current at −30 °C, and B battery still can discharge with 4C current at −20 °C. Considering the wide range of working temperature for low-floor vehicles, the battery must have the ability to discharge with high rate at low temperature. When A battery discharges with high rate at low temperature, the discharging capacity of high rate is more than that of low rate. Therefore, the insulation and heating of battery must be taken into account to avoid the happening of this phenomenon, so that the

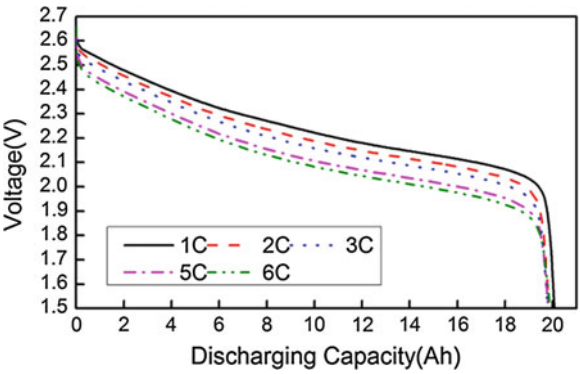


Fig. 12.3 Discharging characteristics of A at 25 °C

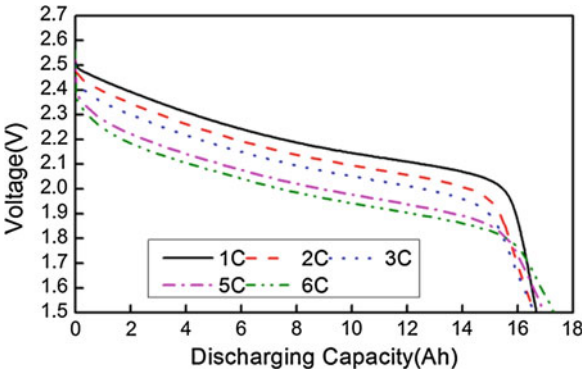


Fig. 12.4 Discharging characteristics of A at 0 °C

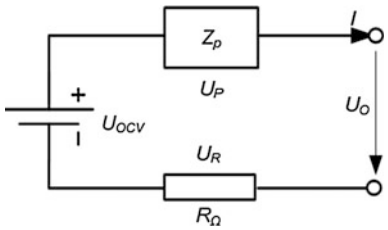


Fig. 12.5 Basic equivalent model of lithium battery

battery is in appropriate temperature. Because the discharge process of battery at high temperature and high rate for a long time may result in a lower battery life, the cooling of battery must be designed to prevent the reduction of battery life. Considering the discharging characteristics at different temperatures, A battery is basically the same as B battery.

Basic equivalent model of lithium battery is shown in Fig. 12.5 [8]. From the model, we can get the following equations:

Table 12.5 The R_Ω, U_R, U_P of 3C and 5C at 0 °C

SOC (%)	R_Ω			U_R			U_P		
	R_{3C}	R_{5C}	ΔR	U_{R3C}	U_{R5C}	ΔU_R	U_{P3C}	U_{P5C}	ΔU_P
90	1.21E-03	1.17E-03	4.0E-05	7.28E-02	1.17E-01	−4.4E-02	4.63E-02	8.23E-02	−3.6E-02
80	1.12E-03	1.08E-03	4.0E-05	6.75E-02	1.08E-01	−4.0E-02	6.25E-02	1.06E-01	−4.3E-02
70	1.06E-03	1.00E-03	6.0E-05	6.38E-02	1.00E-01	−3.6E-02	9.63E-02	1.38E-01	−4.2E-02
60	1.05E-03	9.67E-04	8.3E-05	6.30E-02	9.67E-02	−3.4E-02	1.26E-01	1.68E-01	−4.2E-02
50	1.04E-03	9.75E-04	6.5E-05	6.23E-02	9.75E-02	−3.5E-02	1.15E-01	1.55E-01	−4.0E-02
40	1.05E-03	9.75E-04	7.5E-05	6.30E-02	9.75E-02	−3.5E-02	1.06E-01	1.50E-01	−4.4E-02
30	1.09E-03	1.01E-03	8.0E-05	6.53E-02	1.01E-01	−3.6E-02	1.08E-01	1.52E-01	−4.4E-02
20	1.16E-03	1.08E-03	8.0E-05	6.98E-02	1.08E-01	−3.8E-02	1.16E-01	1.60E-01	−4.3E-02
10	1.53E-03	1.28E-03	2.5E-04	9.15E-02	1.28E-01	−3.6E-02	1.42E-01	1.83E-01	−4.1E-02
0	3.14E-03	1.56E-03	1.6E-03	1.88E-01	1.56E-01	3.2E-02	4.22E-01	4.54E-01	−3.2E-02

$$\begin{cases} U_O = U_{OCV} - U_R - U_P \\ U_R = I \times R_\Omega \\ U_{OCV} = f(SOC) \end{cases} \tag{12.2}$$

and

$$\begin{cases} R_\Omega = f(T, Ch, Dch, Aging) \\ U_P = f(SOC, T, I, Aging, Ch, Dch) \end{cases} \tag{12.3}$$

U_{OCV} is the open circuit voltage, R_Ω is ohmic resistance, U_R is ohmic voltage drop, Z_P is polarization resistance, U_P is polarization voltage drop, $Aging$ is the degree of aging, T is the temperature of environment, I is current, Ch, Dch is the state of charge and discharge.

Based on the basic equivalent model of lithium battery, the test results can be analyzed under the condition of 50 % SOC test point, which is in the charging process of battery. Because the batteries are new batteries and the value of SOC is fixed, the influence of aging and SOC can be ignored.

As can be seen from the test results, the discharging capacity of high rate is more than that of low rate when A battery discharges with high rate at low temperature, such as discharging capacity of 5C current is more than that of 3C current at 0 °C. When A battery discharges at low temperature with high rate, the internal heating of battery is serious. Therefore, the temperature is high, which results in the decrease of ohmic resistance. After the test of battery, the ohmic resistance, ohmic voltage drop, polarization voltage drop of 3C and 5C at 0 °C are shown in Table 12.5.

As can be seen, ΔR in the $SOC \in (0, 20)$ is bigger than ΔR in the $SOC \in (20, 90)$. That is to say, ohmic resistance of 5C is smaller relative to the ohmic resistance of 3C in this stage. As a result, the ohmic voltage drop of 5C is smaller relative to the ohmic voltage drop of 3C when near the end of the discharge process. Besides, the difference between polarization voltage drop of 5C and that of 3C is almost the same. Because the initial conditions and limits of discharging process with two rates are the same, the U_{OCV} of 5C is smaller than that of 3C. Considering the monotonic and nonlinear relationship between U_{OCV} and SOC, the SOC of 5C is smaller than that of 3C, which means the remaining capacity of 5C is smaller than that of 3C. So the discharging capacity of 5C current is more than that of 3C current at 0 °C.

12.4 Conclusion

With the rapid development of low-floor vehicles, lithium titanate battery has advantages of application over other types of lithium batteries. The rate characteristics and low temperature characteristics are obtained from the test results. Considering the operation characteristics of low-floor vehicles, the applicability of

lithium titanate battery for low-floor vehicles is analyzed. Although the rate characteristics and low temperature characteristics of lithium titanate battery can meet the operation demands of low-floor vehicles, there are limits for choosing rates when the battery charges or discharges at low temperature. The ohmic resistance decreases with the increase of temperature when the battery discharges at low temperature with high rate, which may result in special phenomena. Therefore, heating of battery must be taken into account when recycling energy at low temperature. Furthermore, the insulation and heating of battery must be taken into account to avoid the happening of special phenomena, so that the battery is in appropriate temperature. Because the discharge process of battery at high temperature and high rate for a long time may result in a lower battery life, the cooling of battery must be designed to prevent the reduction of battery life.

Acknowledgments The work is supported by National Key Technology R&D Program (2013 BAGL1B00).

References

1. Chen DW, Dong K, Diao LJ (2012) Control technologies of traction converter for 100 % low-floor vehicles. Urban Rapid Rail Transit 25(4):111–115 (in Chinese)
2. Liu ZG (2010) Research on traction drive system for 100 % low-floor vehicles and relevant equipment. Urban Rapid Rail Transit 23(5):17–21 (in Chinese)
3. Diao LJ, Liu ZG (2011) Traction and auxiliary converters of 100 % low-floor vehicles. Urban Rapid Rail Transit 24(6):14–17 (in Chinese)
4. Zhao MH (2010) Development and manufacture of low-floor vehicles. Urban Rapid Rail Transit 23(5):13–16 (in Chinese)
5. Wang HY, Wu F, Fu Y, Li R, Zhang Q (2011) Study on key technologies of lithium battery for electric vehicle. In: The 6th international forum on strategic technology, pp 291–294
6. Mao GL (2009) The development and application of Lithium ion power battery. China Electr Mark 8:14–20 (in Chinese)
7. Lao L (2007) Research on the SOC algorithm of battery management system. Beijing Jiaotong University, Beijing (in Chinese)
8. Wen F (2009) Study on basic problems of lithium-ion batteries management system for electric vehicles. Beijing Jiaotong University, Beijing (in Chinese)

Chapter 13
Research of the Lithium Battery-Based Energy Storage System for Light Rail Vehicle

Xiaoguang Jia, Zhigang Liu, Wenhui He, Zheming Jin and Lijun Diao

Abstract Light rail is developing very fast in China. With the help of battery-based energy storage system, the light rail vehicle will perform much better than before. Light rail vehicle energy storage system links lithium battery and DC-link together, and energy flow two-way through it. The topology of the energy storage system is studied and its working principle is analyzed. The lithium battery is a very suitable energy storage device for the energy storage system for its good charging and discharging characteristics. A double closed-loop including a voltage loop and a current loop is developed to control the energy storage system. The simulation in the MATLAB shows that the energy storage system coordinates very well with the other subsystems in the light rail vehicle in all working conditions.

Keywords Light rail vehicle · Energy Storage System · Lithium Battery · Double-loop control · MATLAB Simulation

13.1 Introduction

The urban rail transportation has got great development in China recently. As a critical part of the urban rail transportation, the light rail is getting more and more concern. Compared with the subway, the light rail has more advantages such as instant gratification, less investment, and wide range of application [1, 2].

There are two main power supply modes for the light rail vehicle, power from the catenary or the third rail [3]. All the two modes need a large number of supporting facilities such as catenary lines or the third rails. But in some special terrain condition such as mountain area, these complex facilities are difficult to

X. Jia (✉) · Z. Liu · W. He · Z. Jin · L. Diao
School of Electrical Engineering, Beijing Jiao-tong University, Beijing, China
e-mail: 12121495@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_13,
© Springer-Verlag Berlin Heidelberg 2014

establish. Even in the area where the facilities are available, the train can possibly lose power supply for there are gaps between the traction substations. Moreover, the power that feeds back when the vehicle is braking is wasted. If the power is fed back in a large amount in a very short time, the voltage of the catenary may be very high which can cause damages to the other facilities. If equipped with the energy storage system, the light rail vehicle is able to deal with these problems. The energy storage system for the light rail vehicle, which is among the DC-link and the traction system, is in charge of the power supply for the train when the catenary or the third rail is not available and transporting the energy that feeds back when the train is braking to the energy storage device.

At present, the technology of energy storage has been very mature in the field of hybrid electric vehicle [4], but is still very limited in the field of light rail vehicle. The key point of the research technology is how to regulate the DC-link voltage and coordinate with the other subsystems of the light rail vehicle.

Based on the research of the energy storage system for the 100 % low-floor light rail vehicle, the circuit topology and its working principle are studied in this article. And then the characteristic of the battery, the performance and the control strategy of the charging and discharging are analyzed. A MATLAB mathematical model of the energy storage system is built to validate the feasibility of the control strategy.

13.2 The Circuit Topology for the Energy Storage System and the Charging and Discharging Characteristics of the Lithium Battery

13.2.1 The Circuit Topology for the Energy Storage System and its Working Principle

13.2.1.1 The Circuit Topology for the Energy Storage System

Here are some main parameters of the 100 % low-floor light rail vehicle [5]. The DC-link voltage ranges from 750 to 930 V, the voltage of the lithium battery ranges from 500 to 700 V. The power that a single energy storage system has to deal reaches over 190 kW (including the power of the auxiliary system and the traction system). The voltage ratio of the DC-converter of the energy storage system is less than two and there is no need to set electrical isolation for the converter. The energy storage system should be capable of supplying the maximum power when the train is in traction mode and absorbing the maximum power that feeds back when the train is braking. So the energy flow of the converter should be bi-directional. Considering all the matters above, the bi-directional Buck-Boost DC converter without isolation which is shown in Fig. 13.1 is chosen as the main circuit topology of the energy storage system converter.

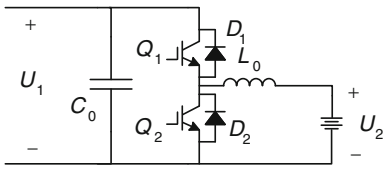


Fig. 13.1 Circuit topology of energy storage system

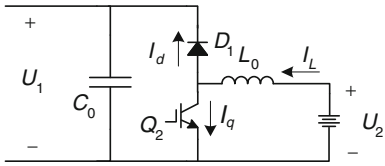


Fig. 13.2 Buck mode of the converter

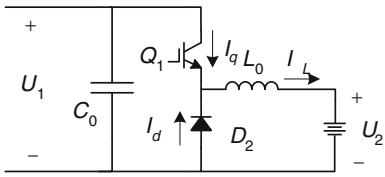


Fig. 13.3 Boost mode of the converter

13.2.1.2 Working Principle of the Circuit

There are two main working modes for the bi-directional Buck-Boost converter, the Buck mode and the Boost mode.

The Fig. 13.2 shows the Buck mode of the converter. The Q1 is working as the main switching tube and the D2 is working as the freewheeling diode. The Q2 and D1 are shut down in this mode. The voltage ratio ranges from 0 to 1 when the width of the control pulse for the Q1 varies from 0 to 100 %. The energy flows from the DC-link to the lithium batteries and the batteries are being charged in this mode. The energy storage system will work in this mode when the light rail vehicle is in braking mode or the SOC of the lithium batteries is too low.

The Fig.13.3 shows the Boost mode of the converter. The Q2 is working as the main switching tube and the D1 is working as the freewheeling diode. The Q1 and D2 are shut down in this mode. The voltage ratio ranges from 1 to $+\infty$ when the width of the control pulse for the Q2 varies from 0 to 100 %. The energy flows from the lithium batteries to the DC-link and the batteries are being discharged in this mode. The energy storage system will switch to this mode when the light rail vehicle is in traction mode without the catenary or the SOC of the lithium batteries is too low.

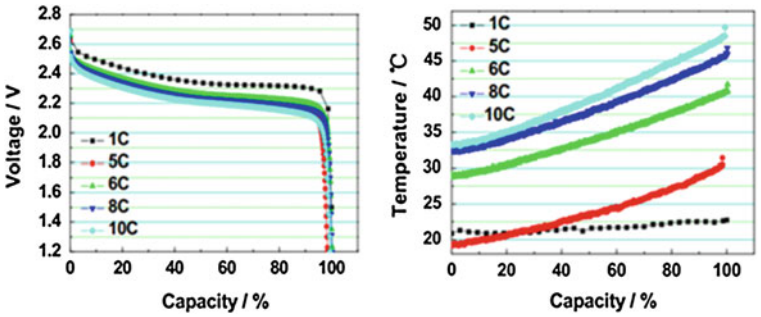


Fig. 13.4 Discharging characteristics of the LTO battery

13.2.2 Charging and Discharging Characteristics of the Lithium Battery

With some advanced features such as high power efficiency, high energy intensity, and brilliant energy storing performance, the lithium battery is quite suitable to be the energy storage device for the light rail vehicle energy storage system. The LTO battery is the latest kind of lithium battery [6]. With little compromise in its energy intensity, it gets great improvement in its life span, safety, and temperature performance and become the most advanced lithium batteries.

The voltage and temperature characteristics of the LTO battery when it is being discharged or charged in different current rate are presented in Figs. 13.3 and 13.4. We can see that the maximum discharge or charge current may reach over 10C, which is far higher than that of the conventional lithium battery. The voltage decrease or increases a little when the SOC of the battery is dropping from 80 to 10 %, which is quite better than the conventional battery [7]. In all course of discharge and charge, the temperature also increases far more little than the conventional one, which could provide a good cooling condition of the system.

13.3 Control of the Battery-Based Energy Storage System

A double closed-loop [8], as shown in Fig. 13.5, including a voltage outer loop and a current inner loop is applied to the control of the bi-directional DC converter of the lithium battery energy storage system when the vehicle is working. A voltage sensor and a current sensor is set in the DC-link to get the information of the DC-link voltage and load current which are the feedback of the control system. The compare value of the actual DC-link voltage and the given DC-link voltage is calculated in the voltage PI controller and the result is used as the given value of the inductance current. The compare value of the actual inductance current and the given inductance current is proceed in the current PI controller and the result is

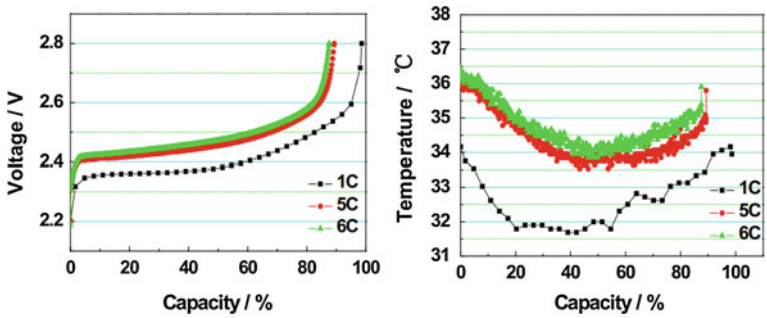


Fig. 13.5 Charging characteristics of the LTO battery

used as the input value of the PWM signal producer, and the consistent driving signal of the IGBT is generated to control the bi-directional DC converter. With the control diagram consisting of a voltage outer loop and a current inner loop, the stability of the DC-link voltage and the accuracy of the inductance current is efficiently ensured.

The digital signal processor is chosen as the core of the hardware part of the control system. With the help of the power module, input and output system, communication interface and all these sensors set in the vehicle, the energy storage system can coordinate with the other subsystems in all working conditions of the train efficiently.

13.4 Simulation of the Energy Storage System

A model of the system is built with the help of MATLAB/Simulink module to simulate the light rail vehicle energy storage system [9]. The parameters of the model are consistent to the actual system. The voltage of the catenary ranges from 750 to 850 V, and the maximum output current of the catenary is 800 A. Because there are two energy storage systems in one train, so each one got 400 A in maximum. The maximum voltage when the train is braking cannot be over 930 V. The rated voltage of the lithium battery group is 607 V and may change from 528 to 712 V. The battery group consist of 264 cells with an internal resistance of 0.04 Ω. The rated capacity of the battery is 40 Ah and the initial SOC is set to 50 %. The inductance is 1 mH and the DC capacitor is 3400 uF. The rated power of the traction inverter is 160 kW and the apparent power is 200 kW. The switch frequency is 1 kHz and the sample frequency is 4 kHz.

First, the situation of charging of the battery when the vehicle is not running with catenary is simulated. In this condition, the energy storage system is charging the battery in a big and constant value of current. The voltage of the catenary is set to 750 and the constant charging value of current is 8 C (320 A). The SOC of the

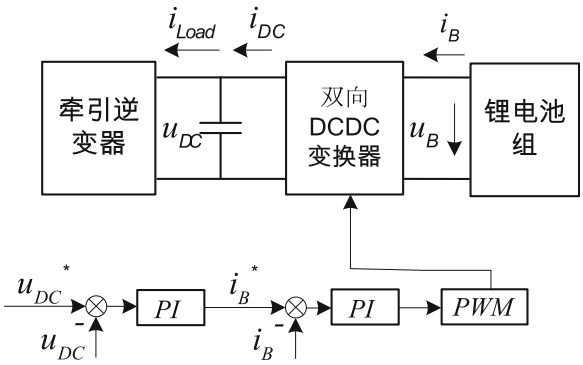


Fig. 13.6 Control diagrams of the system

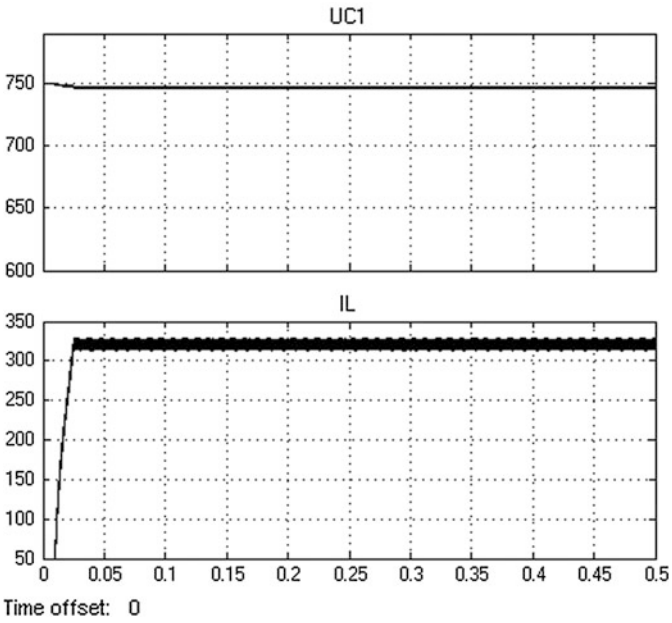


Fig. 13.7 Waveform of the DC-link voltage and the inductance current

battery cannot reach 100 % because there need to be some margin for the battery to absorb the energy that feeds back when the train is braking. The maximum value of SOC is set to 80 % according to the maximum value of the braking power. The simulation time is set to 0.5 s to evaluate the working performance of the system when it is charging in large current. The result is presented in Figs. 13.6 and 13.7. The charging current grows quickly to the pre-set value 320 A with the ripple of less than 20 % and a steady DC-link voltage of 750 V. The good working performance of the system when charging in large current is demonstrated.

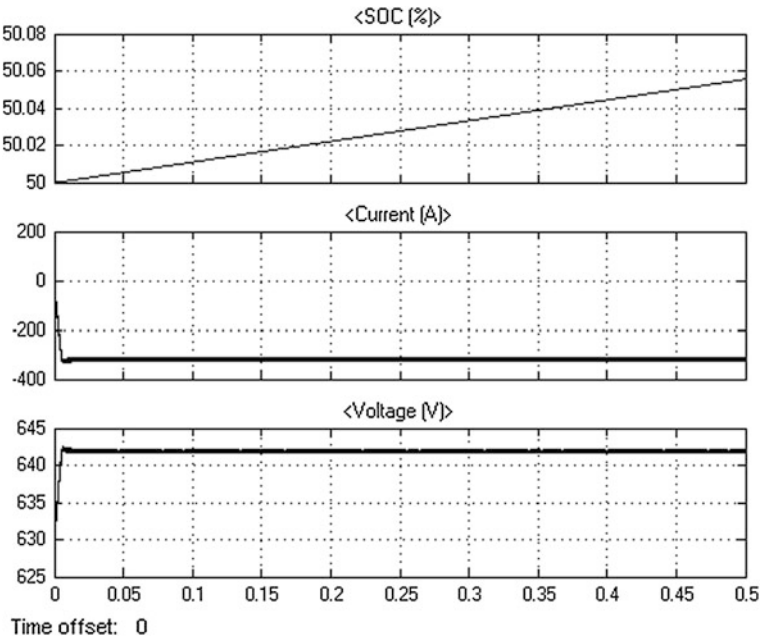


Fig. 13.8 Waveform of the battery voltage, current and SOC

Next the situation of traction and braking without catenary is simulated. The simulation time is set to 3 s. The traction power increases from 0 to maximum from 0 to 0.5 s and remain from 0.5 to 1 s. In the first second the train is in working condition of acceleration. The traction power is changed to the maximum braking power from 1 to 2 s and keeps up until the 2.5 s. In this 1.5 s, the train is in working condition of deceleration. The traction power decreases to 0 from 2.5 to 3 s to simulate the condition that the vehicle speeds down to 0. The simulation waveform is shown in Figs. 13.8, 13.9 and 13.10. From Figs. 13.8 and 13.9, we can see that when the train is in traction mode, the battery is being discharged and the current increases when the traction power increases, and when the train switches to braking mode, the battery is being discharged, the charge current increases when the braking power increases. In the entire course, the DC-link voltage stay steady and the energy storage system reacts quickly to the changes of the working condition of the vehicle. Figure 13.10 shows that the consistence of the given value and the actual value of the inductance current demonstrating the accuracy and efficiency of the control system (Fig. 13.11).

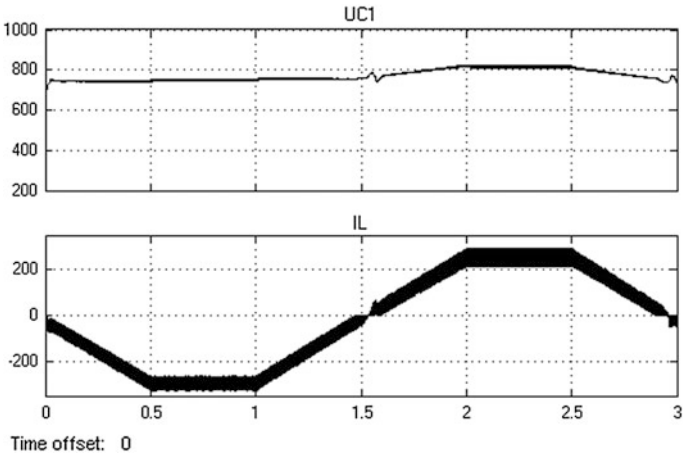


Fig. 13.9 Waveform of the DC-link voltage and the inductance current

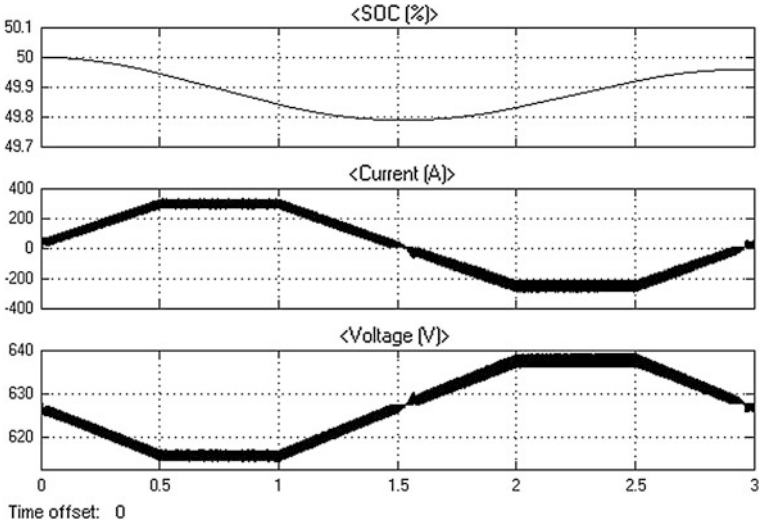


Fig. 13.10 Waveform of the battery voltage, current and SOC

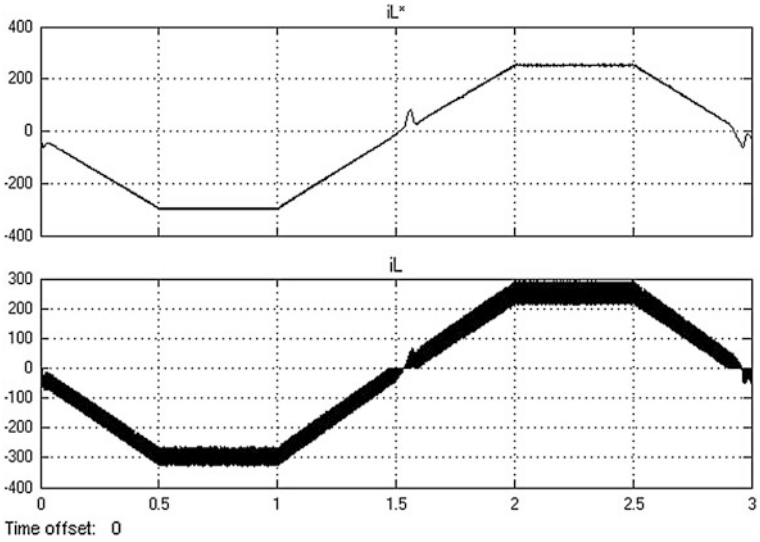


Fig. 13.11 Waveform of the given and the actual inductance current

13.5 Conclusion

Bi-directional Buck-Boost converter is chosen as the main topology of the energy storage system. The control strategy is studied and The MATLAB simulation results demonstrate that the energy storage system can coordinate with the vehicle in all working condition. The energy storage system is suitable to be applied to the light rail vehicle.

References

1. Guo L, Yao Y (2000) The technical development of light rail vehicles. Electr Drive Locomot 2:22–23 (in Chinese)
2. Zhang D, Fan P (1998) Light rail rapid transit for modern cities. Urban Mass Transit 2:13–15 (in Chinese)
3. Zhu L (2003) Analysis on power supply system and current collection mode for urban mass transit. Electr Locomot Mass Transit Veh 26(3) (in Chinese)
4. He H, Zhu J, Li J (2004) The state-of-the-art of hybrid electric vehicle. Veh Power Technol (2):31–33 (in Chinese)
5. Liu Z (2010) Research on traction drive system for 100% low-floor vehicles and relevant equipment. Urban Rapid Rail Transit 23(5) (in Chinese)
6. Su J, Li H (2011) Study on the application of Li4Ti5O12 battery on EMUs. Electr Drive Locomot (4):26–27 (in Chinese)
7. Li J, Jiang J (2013) Battery electric vehicles charger model. Electr Mach Control 17(4):64–71 (in Chinese)

8. Cai F, Zhang Y (2009) DC power supply with DSP-based dual closed-loop control. Electr Power Autom Equip 29(10) (in Chinese)
9. Hu G, Duan S, Cai T et al (2012) Control performance analysis of Lithium-ion battery charger based on model parameter fitting. Trans China Electrotech Soc 27(2):146–152 (in Chinese)

Chapter 14
Research of Fault Location Method for Metro Traction Power Supply System

Xiaoming Song, Jinghan He, Tony Yip, Brian Kirby and Xia Yang

Abstract Accurate fault location is significant to the city metro catenary fault investigation process. A theoretical analysis of the fault location method based on impedance calculation has been performed, using the voltage and current quantities from both end of the line, and the error sources were thoroughly studied. However, in practice the double-terminal electrical quantities used in the calculation may not be synchronous to one another, due to the different protection operating times at either ends of the fault. Another new location algorithm was proposed, which uses genetic algorithm to solve the problem. The simulation results show that the proposed new method has improved location accuracy.

Keywords DC traction system · Fault location method · Impedance calculations · Genetic algorithm

14.1 Introduction

With the rapid development of urban railway transit system, new problems for the security and reliable operation of the power supply system have been identified. In order to enhance the reliability of metro power supply system, improvements can be considered from two aspects. On the one hand, a more reliable protection algorithm should be designed so as to reduce or eliminate protection maloperation due to the increase of the lines and the traffic density. On the other hand, fault

X. Song (✉) · J. He · T. Yip
School of Electrical Engineering, Bejing Jiaotong University, Beijing, China
e-mail: 12121535@bjtu.edu.cn

B. Kirby · X. Yang
Alstom Grid Technology Center, Shanghai, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_14,
© Springer-Verlag Berlin Heidelberg 2014

location techniques are needed, which can estimate the fault location and help to speed up the repair process after fault clearance by the protection devices.

At present, mature fault location methods or devices for DC traction power supply system have not yet been put into practical application. However, the catenary of metro power supply system as a type of transmission line can draw lessons from what methods are commonly used in power system fault location [1, 2]. Especially, the application of HVDC transmission line fault location method research has a very important reference value [3].

14.2 DC Railway System Configuration

In urban railway traffic systems, the DC traction supply system has been widely used. The DC traction system of Beijing Metro in China is operating at a third rail voltage of 750 V DC supplied from 10 kV AC distribution station. A typical DC railway system is shown in Fig. 14.1.

The modeling of DC transit system is implemented based on the following principles:

- The track section is energized by two substations. Each traction substation has two sets of 12-pulse rectifier units and the winding of two rectifier-transformers have $\pm 7.5°$ phase-shifting to form a 24-Pulse equivalent rectification system.
- During short-circuit conditions, the upstream and downstream catenaries both provide short-circuit currents to the fault point, thus it is necessary to include both the upstream and downstream traction networks into the model.
- The phenomenon of skin effect for rails is significant in that the electrical parameters of rails are variable during the short-circuit transient process [4]. This paper has included skin-effect in the distributed parameter model of the traction network during simulation.

14.3 Simple Impedance Method for Fault Location

A variety of fault location algorithms have been developed and proposed over the years. These algorithms can be broadly classified into two categories. One is based on impedance; the other is based on traveling wave [5, 6]. Both types of algorithms have their advantages and inherent defects. By comprehensive analysis and comparison of the two methods with respect to their applicable scope, accuracy, and implementation, the paper decided to use impedance method to achieve fault location and analysed the effective measures to improve the accuracy.

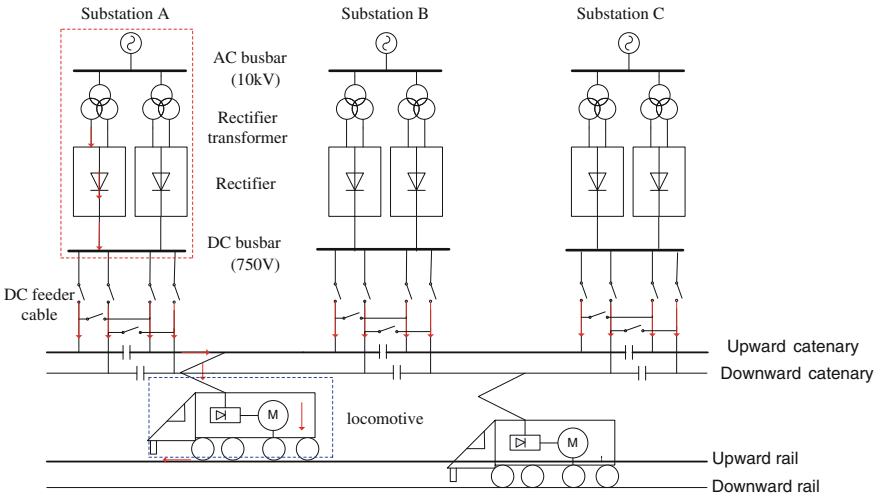


Fig. 14.1 Structure of DC railway power supply system

14.3.1 Theoretical Analysis

For bilateral DC traction power supply system, when short-circuit fault occurs, the fault current contribution at the catenary is from several nearby traction substations. According to Ref. [7] which calculates the steady-state short-circuit current, 96 % of the short-circuit current is from the four nearest traction substations on both sides of short-circuit point, and 80 % comes from the two closest stations. For the sake of the convenience for analysis, this paper researched in accordance with the short-circuit current from the two closest substations and considered the four substations when modeled to ensure the accuracy.

Figure 14.2 shows the short transient circuit diagram. Since the current value used in calculation is close to steady-state current, while the inductance only has impact on the current rising rate, so it can be ignored in the derivation analysis to simplify the calculation model. The simplified equivalent circuit diagram in Fig. 14.3 is obtained from Fig. 14.2 using $\Delta - Y$ transformation:

The following equations are obtained from the circuit using mesh-current analysis:

$$\begin{aligned} I_{s1} &= \frac{(R_2 + R_3) \cdot V_{s1} - R_3 \cdot V_{s2}}{(R_1 + R_3)(R_2 + R_3) - R_3^2} \\ I_{s2} &= \frac{(R_1 + R_3) \cdot V_{s2} - R_3 \cdot V_{s1}}{(R_1 + R_3)(R_2 + R_3) - R_3^2} \end{aligned} \tag{14.1}$$

R_1, R_2 and R_3 are the equivalent parameters that can be computed according to the figures, and the fault location can be numerically expressed as the percentage

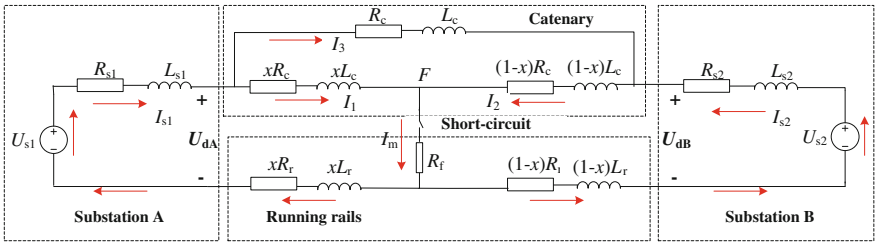


Fig. 14.2 Equivalent circuit of short transient for fault location determination

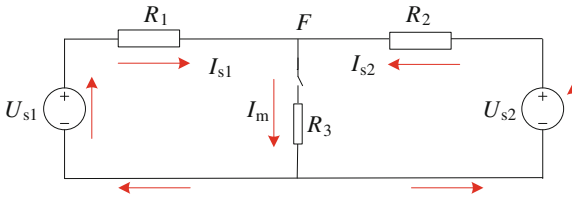


Fig. 14.3 Simplified equivalent circuit diagram

of x, with fault resistance R_f. The currents I_1 and I_2 that flow through the fault catenary can then be deduced, considering that $V_{s1} = V_{s2}$, $R_{s1} = R_{s2}$. From the complex equations above, the fault location formula can be eventually written as below

$$\frac{I_1}{I_2} = \frac{I_{s1} - I_3}{I_{s2} + I_3} = \frac{R_s + \frac{1-x}{2}R_c + (1-x)R_r}{R_s + \frac{x}{2}R_c + xR_r} \approx \frac{1-x}{x} \tag{14.2}$$

Length from substation A can be expressed as

$$L_1 = x \cdot L \tag{14.3}$$

L is the interval length of two adjacent substations. The fault location can be realized using Eqs. (14.2) and (14.3).

14.3.2 Simulation and Results

This paper uses MATLAB/SIMULINK to model the Beijing subway 750 V DC traction system. An internal fault is assumed to occur between the catenary and running rails as shown in Fig. 14.4. To obtain more accurate result, we chose the second-order Butterworth low-pass filter to reduce the harmonics and to calculate the mean values of the data before locating the fault.

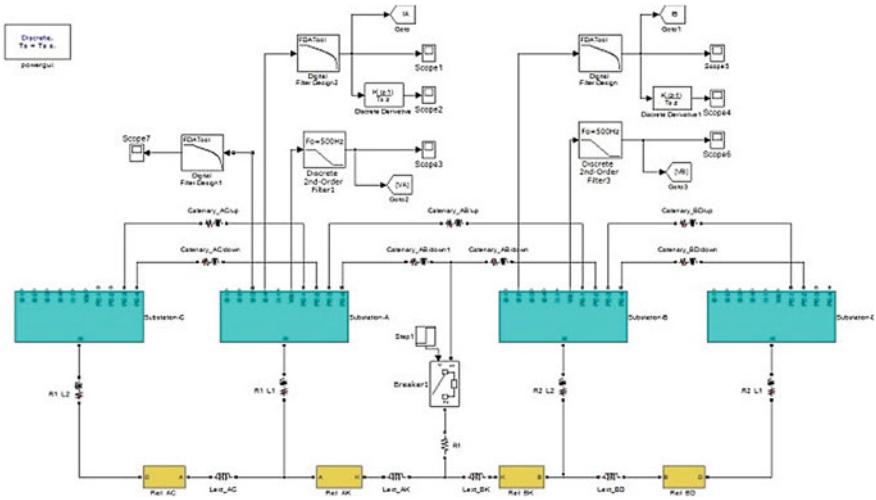


Fig. 14.4 Complete MATLAB/simulink model of short-circuit fault

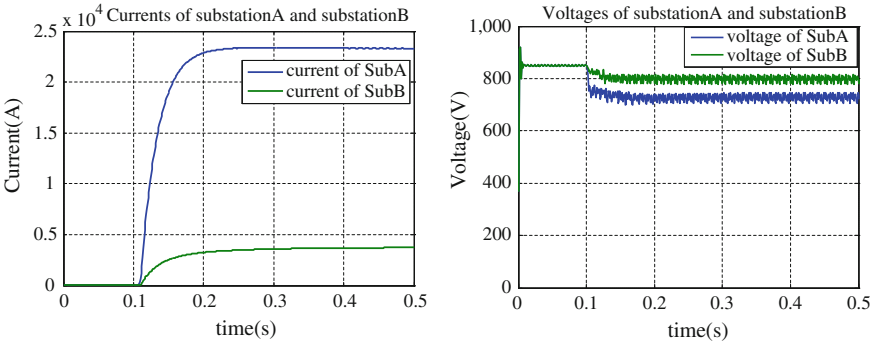


Fig. 14.5 Short-circuit currents and voltages of substation A and substation B

As shown in Fig. 14.5, a short-circuit fault located at 10 % of the power supply range has occurred at time 0.1 s. We use the data of 50 ms after the occurrence of short circuit as the optimum value for fault location. The short-circuit currents of substations A and B are 19,150 A, and 2,547 A, respectively at 0.15 s (0.05 s after fault). Due to space limitations, this article only presents the voltage and current waveforms of 10 % short distance, the other fault location results are shown in Table 14.1.

Table 14.1 Simulation results of fault location based on the simple impedance method

L_1/L (%)	I_A/A	I_B/A	Simulation Result/m	Error/m
10	19,150	2,547	352.1	52.1
20	12,570	3,394	637.7	37.7
30	9,215	4,110	925.3	25.3
40	7,212	4,896	1,213	13.0
50	5,875	5,877	1,500.2	0.2

14.3.3 Improvement Algorithm Based on Simple Impedance Method

Analysis of this simple impedance method has identified certain sources of error. For example, the equation derivation process simplifies a series of conditions, and assumes that the output voltages on both sides of the traction are equal. In fact, the external characteristics of rectifier unit in traction substation lead to the output voltages on both sides vary greatly. Multistage line presented in [8] is commonly used to describe the 24-pulse rectifier unit external characteristics.

As an example of short-circuit fault located at 10 % in Fig. 14.5, the current of substation A at 0.15 s is 19,150 A, the voltage is 718 V; while the current of substation B is 2,547 A and the voltage is 798 V. The DC output voltages are different which also explains why the accuracy of fault near the midpoint is higher than that happens near the ends. In order to avoid the voltage influence, the paper introduces double-end voltages into calculation, and put forward an improved expression as follows (Table 14.2):

$$\frac{I_1}{I_2} \approx \frac{(1-x)\cdot U_{s1}}{x \cdot U_{s2}} \tag{14.4}$$

14.4 A New Fault Location Method Based on Genetic Algorithm

In practical application, the double-ended short-circuit currents and voltages used in the above methods are not steady-state data and are not time synchronous. The DC-circuit breakers tripped on both ends of the traction do not coincide in time. If we acquire the electric quantities based on the time period just before the protection operation, it will create a nonsynchronization problem.

Table 14.2 Fault location results using the improved method

L_1/L (%)	I_A	V_A	I_B	V_B	Result	Error
10	19,150	718	2,547	798	320.3	20.3
20	12,570	729	3,397	791	597.2	−2.8
30	9,215	765.1	4,110	789	905.8	5.8
40	7,212	774.7	4,896	786	1,203	3.0
50	5,875	780.9	5,877	780.9	1,500.2	0.2

14.4.1 Theoretical Analysis

The short-circuit transient model of traction network in Fig. 14.2 is again used for this analysis. Based on the distributed parameter model, the fault location can be deduced according to the principle that the voltages at the fault point calculated by both sides should be equal. The voltage drops due to fault resistance calculated at substations A and B, respectively are shown as below:

$$\begin{aligned} U_{FA} &= U_{dA} - I_1 \cdot xR_c - xL_c \cdot \frac{dI_1}{dt} - (I_1 + I_3) \cdot xR_r - xL_r \cdot \frac{d(I_1 - I_3)}{dt} \\ U_{FB} &= U_{dB} - I_2 \cdot (1-x)R_c - (1-x)L_c \cdot \frac{dI_2}{dt} \\ &\quad - (I_2 - I_3) \cdot (1-x)R_r - (1-x)L_r \cdot \frac{d(I_2 - I_3)}{dt} \end{aligned} \tag{14.5}$$

As the simulation on feeder current and voltage is sampled, the above time-domain differential equations can be rewritten into difference equations to get the voltage difference calculated on each side:

$$\begin{aligned} U_{FA}(k) - U_{FB}(k) = U_{dA}(k) - I_1(k)_c - xL_c \cdot \frac{I_1(k) - I_1(k-1)}{T_s} - (I_1(k) + I_3(k)) \cdot xR_r \\ - xL_r \cdot \frac{I_1(k) - I_1(k-1) + I_3(k) - I_3(k-1)}{T_s} \\ - U_{dB}(k) + I_2(k) \cdot (1-x)R_c \\ + (1-x)L_c \cdot \frac{I_2(k) - I_2(k-1)}{T_s} + (I_2(k) - I_3(k)) \cdot (1-x)R_r \\ + (1-x)L_r \cdot \frac{I_2(k) - I_2(k-1) - I_3(k) + I_3(k-1)}{T_s} \end{aligned} \tag{14.6}$$

Beijing Metro Line 2 uses Secheron SEPCOS as the measurement and protection integrated equipment for DC power supply system. It can record the short transient current and voltage data by being installed in the secondary wiring part of DC feeder switchgear. U_{dA}, U_{dB}, I_1, I_2, and I_3 in the above formula can be obtained

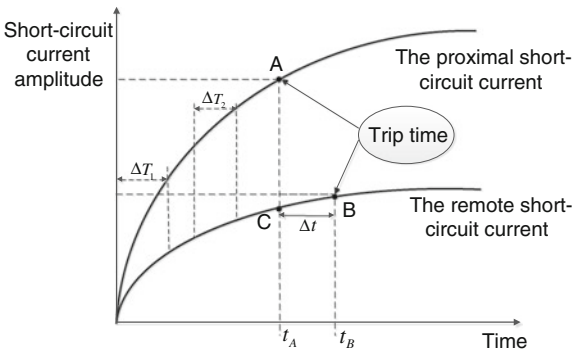


Fig. 14.6 Short-circuit current tripping time of double-side

through the simulation model, and the device of SEPCOS can also collect and record these electrical data.

When a short-circuit fault occurs, we can retrieve the previously stored voltage and current sampled values from a number of time durations ΔT_1, ΔT_2,…,ΔT_n before the trip point as shown in Fig. 14.6. The calculations from each time duration are then averaged and the percentage of x can be obtained.

$$\begin{cases} \min \quad f(x) = \dfrac{\sum_{k=1}^{N} [U_{FA}(k) - U_{FB}(k)]^2}{N}, \quad (N = \dfrac{\Delta T_i}{T_s}, \text{where } T_s \text{ is sampling time}) \\ \bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} \end{cases} \tag{14.7}$$

The algorithm can solve the fault locating equations with genetic algorithm (GA) which is a global optimization approach that has great advantages in detecting distribution of solutions of a complicated problem. The built-in MATLAB Genetic Algorithm Tool provides great convenience for the genetic algorithm optimization problems. The proposed algorithm is proved to be correct and accurate by simulation results of MATLAB.

14.4.2 Simulation and Results

See Table 14.3.

Table 14.3 Simulation results using the new method based on GA

L_1/L (%)	Analysis of genetic algorithm	
	Result/m	Error/m
10	304.8	4.8
20	606.3	6.3
30	906	6.0
40	1,203	3.0
50	1,500.3	0.3

14.5 Conclusion

In this paper, an impedance-based fault location method based on the ratio of two-terminal steady-state currents was analyzed. Through analysis of the origin of ranging error, an improved method considering double-end voltages was proposed. The results show that the fault location precision can be significantly improved. However, in practice the unsteady-state data and nonsynchronized problem was found to introduce errors, due to the protection devices operate before the current reaches steady-state, creating inconsistency in the two terminal circuit breakers' tripping times. A genetic algorithm technique was finally adopted which does not require fault type identification, eliminates the influence of fault resistance and system impedance, and resolves the errors caused by simple impedance method. Simulation results using Matlab/Simulink show the effectiveness of this method.

Acknowledgments This work is supported by Beijing Jiaotong University and ALSTOM collaborative research project: Study on IEC61850 in DC Railway Traction Systems with an Aim to Develop New Protection Devices.

References

1. Saha MM, Iżykowski J, Rosolowski E (2010) Fault location on power networks. Springer, Berlin
2. Cheng W, Xu GQ, Mu LH (2005) A novel fault location algorithm for traction network based on distributed parameter line model. Proc Electr Power Syst Autom 6:14
3. Suonan J, Gao S, Song G, Jiao Z, Kang X (2010) A novel fault-location method for HVDC transmission lines. IEEE Trans Power Delivery 25(2):1203–1209
4. Yu L, He JH, Hu J, Bo ZQ, Li MX, Yip HT, Klimek A (2010) Accurate track modeling for fault current on DC railways based on MATLAB/Simulink. In: Power and energy society general meeting, IEEE, July 2010, pp 1–6
5. Kale V, Bhide S, Bedekar P (2012) Fault location estimation based on wavelet analysis of traveling waves. In: Power and energy engineering conference (APPEEC), Asia-Pacific, IEEE, March 2012, pp 1–5
6. Chang CS, Feng T, Khambadkone AM, Kumar S (2000) Remote short-circuit current determination in DC railway systems using wavelet transform. IEEE Proc Electr Power Appl 147(6):520–526 (IET)

7. Liu W (2009) Urban rail transit train running process optimization and traction power system dynamic simulation. Doctoral dissertation, Southwest Jiaotong University, Chengdu (in Chinese)
8. Zhou W-W, Chen D, Chen L, Wang Y (2012) An improved method and application subway fault location. Electrified Railw 23(2):47–50 (in Chinese)

Chapter 15
Study and Implementation of Closed Loop Control Based on Double Synchronous Rotating Frame for EMU Auxiliary Inverter with Unbalanced Load

Huiqing Du, Zhigang Liu, Yao Fu, Lei Wang and Lijun Diao

Abstract As one of the EMU auxiliary inverter's important performances, the symmetry of three-phase voltage puts forward high requirement in the inverter's topologies and control strategies. Based on the method of symmetrical component, this paper builds positive-, negative-, and zero-sequence mathematical model of the three-phase inverter system with unbalanced load, and then analyses the generation mechanism of asymmetry output voltage. Afterwards it is proposed a double-loop SVPWM control strategy based on double synchronous rotating frame, and further analysis of the working mechanism and realization mode is raised. Finally, it turns out to be reliable and feasible to apply the control strategy after being tested with virtual DSP simulation model with MATLAB, and on a 3 kVA experimental prototype.

Keywords Auxiliary inverter · Unbalanced load · Method of symmetrical component · Double-loop control · Double synchronous rotating frame

15.1 Introduction

In addition to provide power to three-phase loads, EMU auxiliary inverter should also provide power to single-phase loads, such as windscreen heater, drain heater, outlet, and so on. Different kinds of single-phase load will lead to three-phase unbalanced voltage. For the air-conditioning compressor, cooling fan, and other three-phase electromagnetic components, if the imbalance of input voltage exceeds the allowable value, it will cause serious overheating of electromagnetic components. This will not only bring greater loss, lower efficiency, and worse

H. Du (✉) · Z. Liu · Y. Fu · L. Wang · L. Diao
School of Electrical Engineering, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 11117361@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_15,
© Springer-Verlag Berlin Heidelberg 2014

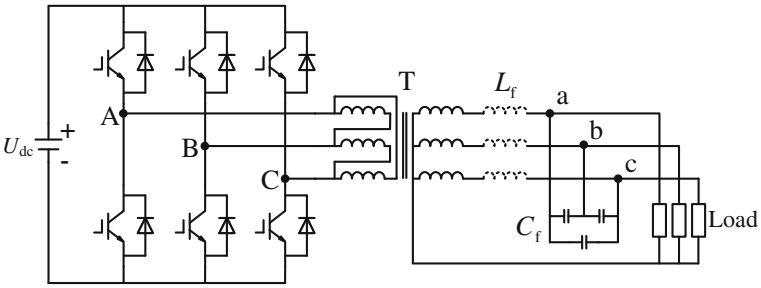


Fig. 15.1 The normal topology of EMU high-power auxiliary inverter

performance of the system, but also with the degree of imbalance continuing to increase, it will damage the equipment [1]. Therefore, the analysis and study of EMU auxiliary inverter's control strategies and topologies with unbalanced load is of great significance [2, 3].

The usual three-phase four-wire inverter topologies can be divided into the following four kinds, three-phase four-wire inverter with split capacitors topology [4, 5], neutral point forming transformer topology [6], three-phase full-bridge inverter with $\triangle/Y_0$ type isolation transformer topology [7], and three-phase four-leg topology [8, 9]. Although there are limited categories of topologies of three-phase four-wire inverter, its control strategies are relatively more [10–13].

EMU high-power auxiliary inverter usually adopts three-phase full-bridge inverter with $\triangle/Y_0$ type isolation transformer topology, as Fig. 15.1 shown. Based on the topology, this paper expounds production mechanism of output voltage asymmetry. And by using a novel controller, it further eliminates the output voltage distortion caused by unbalanced load. Finally, the simulation results by virtual DSP model of MATLAB Simulink show the feasibility of the control method, and it's verified by experiment.

15.2 Unbalanced Mechanism

15.2.1 Mathematical Model of Unbalanced Load

By dividing the asymmetric structure into several superimposed symmetric structures, it will be able to analyze the entire system with basic circuit theory. By using the method of symmetrical component [14], the instantaneous unbalanced output voltage can be divided into positive-, negative-, and zero-sequence voltage components, as formula (15.1) shown. And in order to facilitate the analysis, Table 15.1 shows the names of variables in this paper and their physical meanings.

Table 15.1 The definitions of variable names

Variable name	Physical meaning
Subscript p, n, z	Positive-, negative- and zero-sequence component
Subscript *j*, *J*	a, b, c and A, B, C
u_a, u_b, u_c	Inverter output phase voltage
u_A'', u_B'', u_C''	Fundamental voltage of three-phase inverter converted to the secondary
i_a', i_b', i_c'	Transformer secondary output current
i_a, i_b, i_c	Inverter load current
$R_{1\sigma}''$, $X_{1\sigma}''$	Transformer primary leakage impedance converted to the secondary
R_m'', X_m''	Transformer magnetizing impedance
R_{m0}'', X_{m0}''	Transformer zero-sequence magnetizing impedance
$R_{2\sigma}$, $X_{2\sigma}$	Transformer secondary leakage impedance
L_f, C_f, Z_j	Inverter output inductor, capacitor and load impedance
L_T, R_T, R_{T0}	$X_{1\sigma}'' + X_{2\sigma} + L_f$, $R_{1\sigma}'' + R_{2\sigma}$, $R_{10} + R_{2\sigma}$
L_{T0}	$X_{10} + X_{2\sigma} + L_f$
R_{10}, X_{10}	Parallel equivalent impedance of $R_{1\sigma}''$, $X_{1\sigma}''$ and R_{m0}'', X_{m0}''

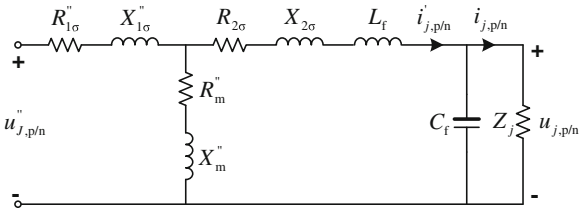


Fig. 15.2 The positive- and negative-sequence equivalent circuit of the transformer

$$\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} u_{a,p} \\ u_{b,p} \\ u_{c,p} \end{bmatrix} + \begin{bmatrix} u_{a,n} \\ u_{b,n} \\ u_{c,n} \end{bmatrix} + \begin{bmatrix} u_{a,z} \\ u_{b,z} \\ u_{c,z} \end{bmatrix} \tag{15.1}$$

Based on the T-form equivalent circuit of the transformer, the system can be divided into positive-, negative-, and zero-sequence equivalent circuit, as shown in Figs. 15.2 and 15.3.

Ignoring the transformer's magnetizing impedance and line impedance, we can get the differential equations of positive- and negative-sequence voltage and output filter inductor current component, respectively shown as formulae (15.2) and (15.3).

$$\frac{d}{dt}\begin{bmatrix} u_{a,p/n} \\ u_{b,p/n} \\ u_{c,p/n} \end{bmatrix} = \begin{bmatrix} \frac{1}{C_f} & 0 & 0 \\ 0 & \frac{1}{C_f} & 0 \\ 0 & 0 & \frac{1}{C_f} \end{bmatrix} \left(\begin{bmatrix} i'_{a,p/n} \\ i'_{b,p/n} \\ i'_{c,p/n} \end{bmatrix} - \begin{bmatrix} i_{a,p/n} \\ i_{b,p/n} \\ i_{c,p/n} \end{bmatrix} \right) \tag{15.2}$$

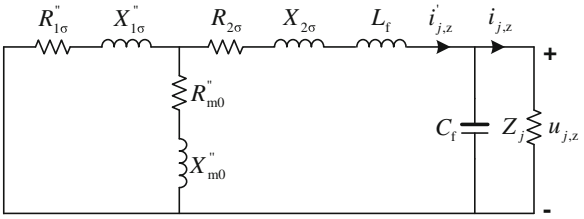


Fig. 15.3 The zero-sequence equivalent circuit of the transformer

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} i'_{\mathrm{a,p/n}} \\ i'_{\mathrm{b,p/n}} \\ i'_{\mathrm{c,p/n}} \end{bmatrix} = \begin{bmatrix} \frac{1}{L_\mathrm{T}} & 0 & 0 \\ 0 & \frac{1}{L_\mathrm{T}} & 0 \\ 0 & 0 & \frac{1}{L_\mathrm{T}} \end{bmatrix} \left(\begin{bmatrix} u''_{\mathrm{A,p/n}} \\ u''_{\mathrm{B,p/n}} \\ u''_{\mathrm{C,p/n}} \end{bmatrix} - \begin{bmatrix} u_{\mathrm{a,p/n}} \\ u_{\mathrm{b,p/n}} \\ u_{\mathrm{c,p/n}} \end{bmatrix} \right) - \begin{bmatrix} \frac{R_\mathrm{T}}{L_\mathrm{T}} & 0 & 0 \\ 0 & \frac{R_\mathrm{T}}{L_\mathrm{T}} & 0 \\ 0 & 0 & \frac{R_\mathrm{T}}{L_\mathrm{T}} \end{bmatrix} \begin{bmatrix} i'_{\mathrm{a,p/n}} \\ i'_{\mathrm{b,p/n}} \\ i'_{\mathrm{c,p/n}} \end{bmatrix} \tag{15.3}$$

For three-phase full-bridge inverter with $\triangle/Y_0$ type isolation transformer topology, the zero-sequence component of load current is not 0. The relationship of zero-sequence components between output voltage and load current is as shown in formulae (15.4) and (15.5), from which it can be deduced that the zero-sequence component is mainly determined by transformer leakage impedance, zero-sequence magnetizing impedance and filter inductor.

$$\frac{\mathrm{d}u_{j,z}}{\mathrm{d}t} = \frac{1}{C_\mathrm{f}}\left(i'_{j,z} - i_{j,z}\right) \tag{15.4}$$

$$\frac{\mathrm{d}i'_{j,z}}{\mathrm{d}t} = -\frac{u_{j,z}}{L_{\mathrm{T0}}} - \frac{R_{\mathrm{T0}}i'_{j,z}}{L_{\mathrm{T0}}} \tag{15.5}$$

15.2.2 Analyze the Reason of Unbalance

The three-phase inverter usually adopts the controller in synchronous rotating d-q coordinate system, whose rotational angular frequency is based on the grid fundamental angular frequency ω_f. Through Park transformation, we can get positive- and negative-sequence d-q axis components in rotating coordinate system of three-phase output voltage, which are shown in formula (15.6).

$$\begin{bmatrix} u_{\mathrm{od,p}} & u_{\mathrm{od,n}} \\ u_{\mathrm{oq,p}} & u_{\mathrm{oq,n}} \end{bmatrix} = \begin{bmatrix} \frac{3}{2}U_{\mathrm{M,p}}\sin\varphi_\mathrm{p} & \frac{3}{2}U_{\mathrm{M,n}}\sin(2\omega_\mathrm{f}t+\varphi_\mathrm{n}) \\ -\frac{3}{2}U_{\mathrm{M,p}}\cos\varphi_\mathrm{p} & \frac{3}{2}U_{\mathrm{M,n}}\cos(2\omega_\mathrm{f}t+\varphi_\mathrm{n}) \end{bmatrix} \tag{15.6}$$

wherein, $U_{\mathrm{M,p}}$ and $U_{\mathrm{M,n}}$ donate the amplitude of positive- and negative-sequence fundamental voltage, φ_p and φ_n donate the initial phase angle of positive- and negative-sequence voltage. From the above equation, for the positive-sequence component, the reference and feedback values of output voltage are DC

components. Therefore, by using the integral action of proportional-integral (PI) controller, we can force the positive-sequence output variables to track the references with no static error. However, the negative-sequence component exists in form of AC component, whose angular frequency is $2\omega_f$. The PI controller can't eliminate the error of negative-sequence output voltage. According to the regulation of positive-sequence component, a novel PI control in reverse synchronous rotating frame is proposed to weaken the error of negative-sequence component. Thus the negative-sequence component comes into DC form, which has the same moderating effect as positive-sequence component.

15.3 Control Strategy with Unbalanced Load

15.3.1 Double-Loop Control Strategy

In this paper, we adopt novel double-loop control strategy, and the feed-forward introduction of voltage command improves the system's open-loop gain. The control diagram of three-phase four-wire auxiliary inverter is shown in Fig. 15.4. With this strategy, the negative-sequence distortion of inverter with unbalanced load is eliminated, and the balanced output voltage waveform is got.

For ideal three-phase sinusoidal voltage signals, they should not contain negative-sequence component, so the reference value of negative-sequence d-q axis component is 0. However, the reference value of positive-sequence d-q axis component depends on the desired voltage amplitude and initial phase of the signal. In three-phase four-wire inverter system, whether it is positive- or negative-sequence component, there exists coupling relationship between d-q axis current and voltage. In the case of heavy load, the designed voltage loop is difficult to completely eliminate the influence of current coupling, which will have serious disturbance of output voltage. So it is necessary to decouple the d-q axis current by multiplying a factor $L_f\omega_f/U_{dc}$ in positive-, negative-sequence component after sampling and conversion, as Fig. 15.4 shown.

15.3.2 Method of Extracting Variables

The method of extracting positive- and negative-sequence d-q axis component of three-phase unbalanced signal mainly includes two kinds, and their flowcharts are shown in Fig. 15.5. Wherein, the signal got by method (a) exists of a quarter delay of the modulating waveform period. Its oscillation is comparatively large, but the dynamic response speed is relatively fast. The signal waveform got by method (b) is smooth, but it responses slowly, and exists of one period delay.

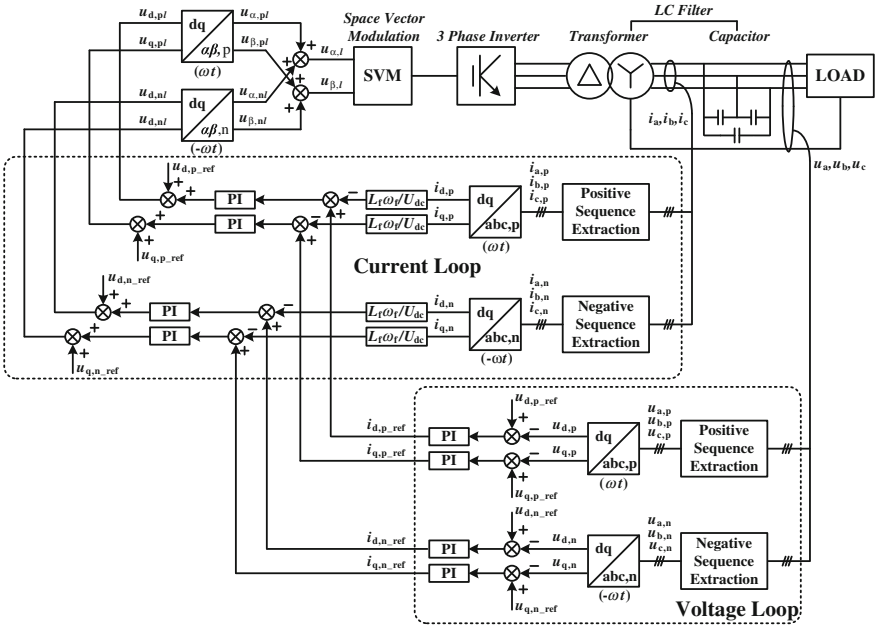


Fig. 15.4 The control block diagram of auxiliary inverter with unbalanced load

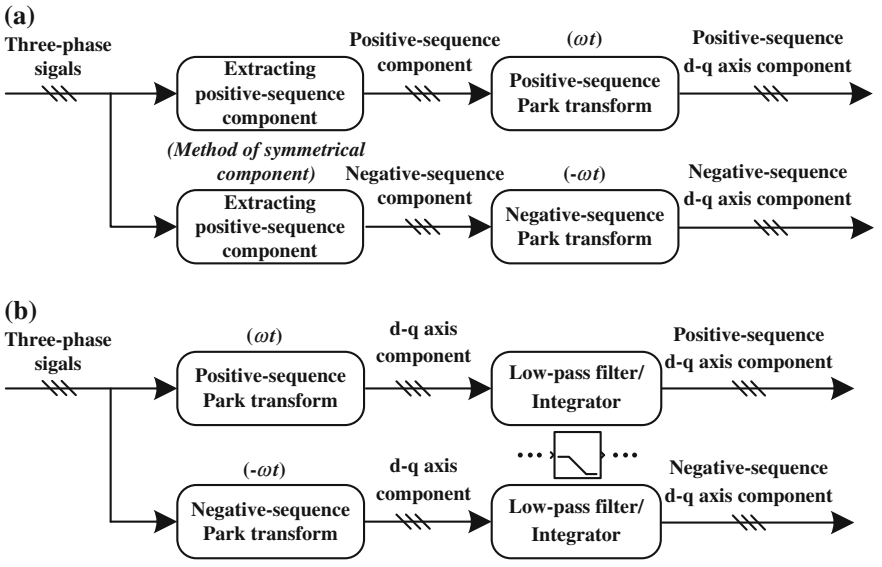


Fig. 15.5 The flowchart of extracting positive- and negative-sequence d-q axis component

For EMU auxiliary inverter, whose load is changing at all times, it proposes high requirement of the system to possess favorable track of the load. This shows that method (a) is more suitable for this system, and through designing proper low-pass filter it can eliminate the interference, thus the integrated control effect improves.

15.3.3 Implementation of Control System

As the control strategy in this paper includes numerous complex elements, such as signal decomposition, coordinate transformation, and variables regulation. In order to facilitate the implementation of digital control system, it is necessary to take a reasonable approach to simplify the complex procedures, meanwhile to ensure the control precision, response speed, and operation reliability of the system.

Through the method of symmetrical component, the sampled instantaneous output voltage can be decomposed into positive- and negative-sequence component, as formulae (15.7) and (15.8) shown. The three-phase current can also be decomposed by the same method.

$$\begin{bmatrix} u_{a,p} \\ u_{b,p} \\ u_{c,p} \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} - j\frac{\sqrt{3}}{6}\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} \tag{15.7}$$

$$\begin{bmatrix} u_{a,n} \\ u_{b,n} \\ u_{c,n} \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} - j\frac{\sqrt{3}}{6}\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} \tag{15.8}$$

To realize the transformation, we adopt 90° angle transient method, which has significant advantage in real-time digital control system. It will not take large storage resource of the microcontroller, and the algorithm is simple, which facilitates the realization of the controller. Even when the system is instable, it can still get good control effect. The flowchart of extracting positive-sequence component by 90° angle method is shown in Fig. 15.6, where j (i.e., 90° phase shifted) in the above formula can be digitized by all-pass filter (APF). Similarly, there is a diagram for extracting negative-sequence component, only by modifying the gain factor.

Through positive- and negative-sequence Park transformation, we can get d-q axis DC components, which facilitate decoupling control. The expressions of positive- and negative-sequence Park transformation are shown as follows, where $\theta = \omega_e t$, and ω_e represents angular frequency of synchronization signal.

$$C_{abc,p/dq} = \frac{2}{3}\begin{bmatrix} \sin(\theta_t) & \sin\left(\theta_t - \frac{2}{3}\pi\right) & \sin\left(\theta_t + \frac{2}{3}\pi\right) \\ \cos(\theta_t) & \cos\left(\theta_t - \frac{2}{3}\pi\right) & \cos\left(\theta_t + \frac{2}{3}\pi\right) \end{bmatrix} \tag{15.9}$$

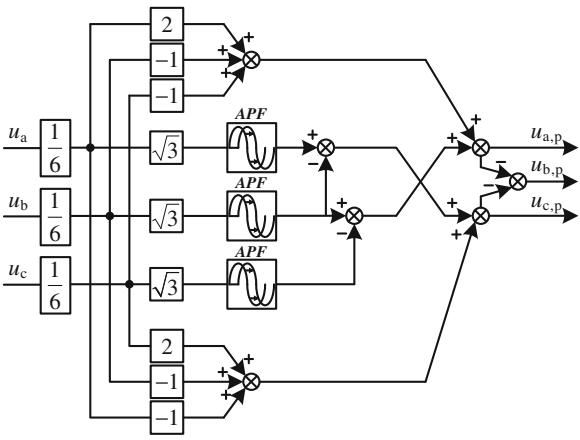


Fig. 15.6 The diagram of extracting positive-sequence component

$$C_{\mathrm{abc,n/dq}} = \frac{2}{3}\begin{bmatrix} \sin(\theta_t) & \sin\left(\theta_t + \frac{2}{3}\pi\right) & \sin\left(\theta_t - \frac{2}{3}\pi\right) \\ \cos(\theta_t) & \cos\left(\theta_t + \frac{2}{3}\pi\right) & \cos\left(\theta_t - \frac{2}{3}\pi\right) \end{bmatrix} \tag{15.10}$$

Because the system is influenced by controller delay, filter delay, frequency perturbation and other factors, the initial phase angle is often not a fixed value. So we propose a three-phase software phase-locked loop (SPLL) based on synchronous rotating coordinate system [15, 16], and by using the PI controller it has strong anti-interference ability and good tracking results, as shown in Fig. 15.7. Wherein, U_r represents the boundary value between linear regulator and PI regulator, which determines the adjustment bandwidth and stability margin. And n_e represents the enumerated data of the controller's inner timer, whose reference benchmark N_{ff} is the maximum count value when $\omega_e = \omega_f$. And n_{con} represents adjusting step length, which determines the tracking speed of SPLL, furthermore, because the frequency fluctuation of output voltage is so small that $n_{con} \ll N_{ff}$.

Through a series of calculations and adjustment, we finally get positive-, negative-sequence d-q axis control quantities, which should be converted into control variables in two-phase stationary coordinate system by positive- and negative- anti-Clark transformation. Thus, they can participate in the control of SVPWM modulation process. The specific transformation method is shown as follows:

$$C_{\mathrm{dp,p}/\alpha\beta} = \begin{bmatrix} \cos(\omega_e t) & -\sin(\omega_e t) \\ \sin(\omega_e t) & \cos(\omega_e t) \end{bmatrix} \tag{15.11}$$

$$C_{\mathrm{dp,n}/\alpha\beta} = \begin{bmatrix} -\cos(\omega_e t) & \sin(\omega_e t) \\ \sin(\omega_e t) & \cos(\omega_e t) \end{bmatrix} \tag{15.12}$$

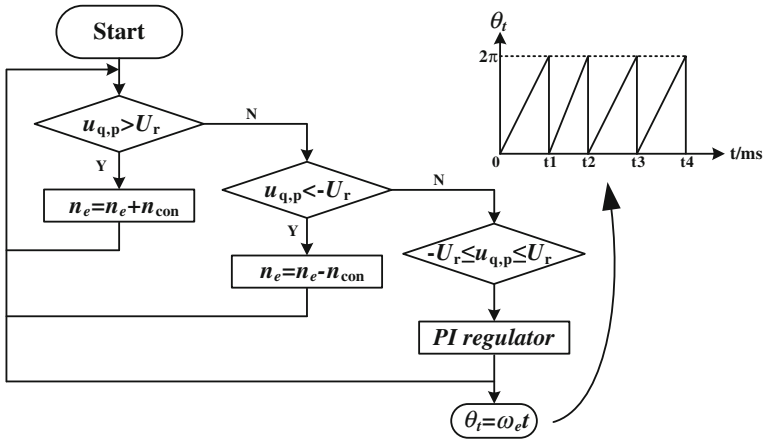


Fig. 15.7 The flowchart of three-phase SPLL

15.4 Simulation and Experimental Verification

According to the above analysis and design, based on MATLAB simulation tools, this paper constructs virtual DSP simulation model in Simulink environment to test and verify the proposed control strategy. The regulatory effect on unbalanced three-phase output voltage is contrasted between conventional voltage, current double-loop control and the novel control in this paper.

Among the EMU AC380 V medium-voltage bus electricity loads, three-phase electromagnetic components such as air conditioning compressors, cooling fans and other equipment, account for more than 75 % of total capacity, and nearly 15 % of rated power is supplied to battery charger. Such loads mainly care about the supply quantity of line voltage, which is the main detection object. And according to the two load situations, namely phase A-rated load, phase B, C-50 % rated load and phase A-rated load, phase B, C-no load, simulations are separately compared (Figs. 15.8, 15.9).

Under different load situations, the comparison results are given in Table 15.2, in which the method of calculating the voltage unbalance mainly refers to the relevant regulations in IEC61000-4-30.

The simulation waveforms and data form the above comparison show that with conventional closed-loop control strategy, the maximum difference of peak value of three-phase output line voltage is 16.4 V, and the degree of imbalance is closed to 3 %. However, when using double-loop control strategy based on double synchronous rotating frame, the capability of adjusting to unbalanced load has been significantly enhanced. The maximum difference of peak value drops to 3.6 V, and the degree of imbalance stabilizes is less than 1 %, which greatly increases the quality of output voltage.

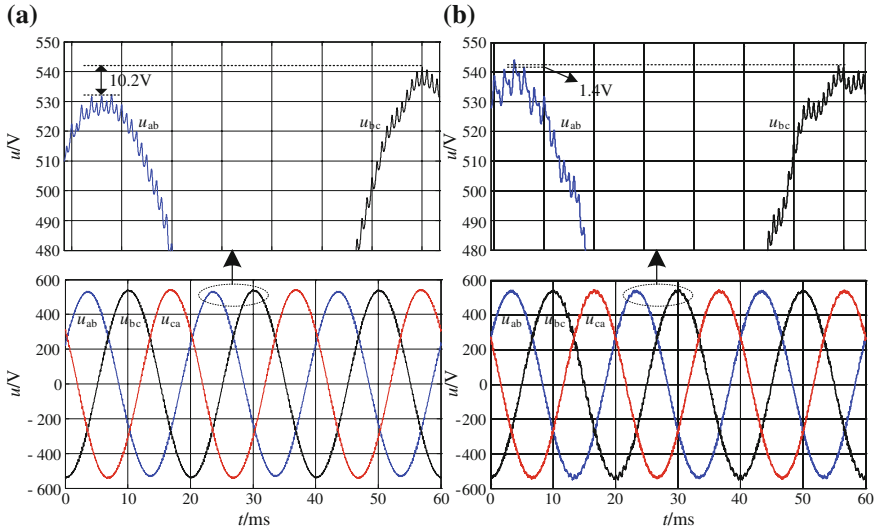


Fig. 15.8 Output voltage comparison with phase A-rated load and phase B, C-50 % rated load

Table 15.2 Output voltage and unbalanced degree comparison results

Control strategy	Load condition	u_{ab}/V	u_{bc}/V	u_{ca}/V	Unbalanced degree
Conventional control strategy	A-100 %, B/C-50 %	371.4	378.6	382.0	1.652 6 %
	A-100 %, B/C-0 %	368.2	379.8	383.7	2.457 3 %
Novel control strategy	A-100 %, B/C-50 %	379.3	380.3	380.3	0.175 4 %
	A-100 %, B/C-0 %	376.0	378.6	380.3	0.660 7 %

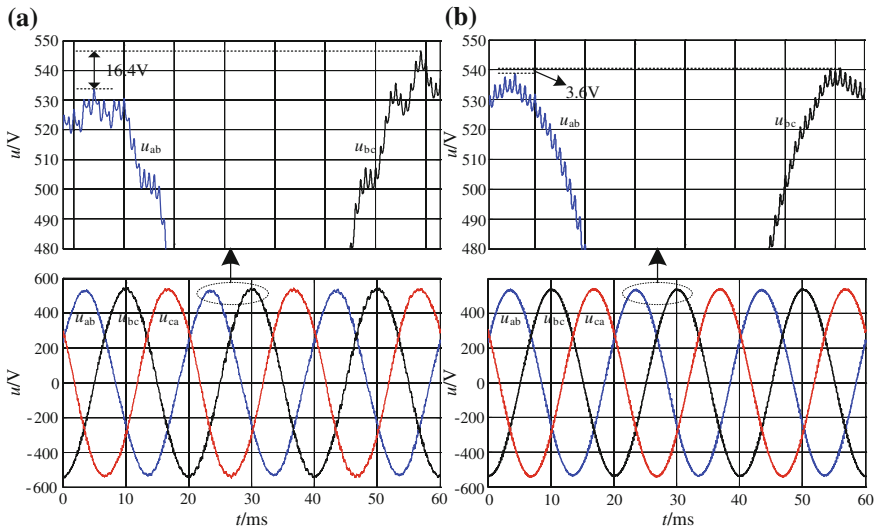


Fig. 15.9 Output voltage comparison with phase A-rated load and phase B, C-no load

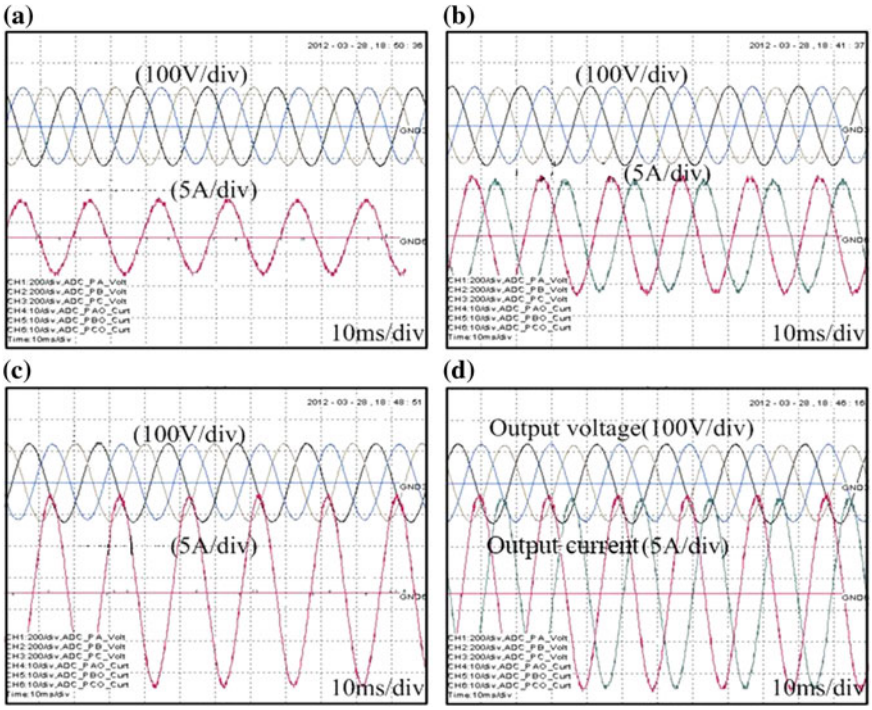


Fig. 15.10 The experimental waveform with unbalanced load

In order to verify the correctness of the above simulation results, this paper conducts multiple experiments with unbalanced load based on a 3 kVA experimental prototype. The experimental results are shown in Fig. 15.10.

Figure 15.10a and c show the inverter's experimental waveforms with different single-phase loads, and Fig. 15.10b and d show the experimental waveforms with different two-phase loads. Seen from the above experimental results, the three-phase output voltage can remain balanced with the proposed control strategy, and its degree of imbalance is less than 1 %. So the system is provided with formidable capacity to adjust to unbalanced load.

15.5 Conclusion

Based on the complex load conditions of EMU high-power auxiliary inverter, this paper focuses on the analysis of output waveform distortion with unbalanced load, and proposes a novel control strategy to improve the quality of output voltage. This paper first uses the method of symmetrical component to construct positive-, negative-, and zero-sequence mathematical model, and then analyses the cause of

generating unbalanced voltage. Furthermore, synthesizing the system topology, a double-loop control strategy on double rotating coordinate frame is proposed, and its all-digital implementation method is given. Finally, the simulation and experimental results verify that the proposed control strategy can ensure higher quality output voltage of the system with unbalanced load, even with single-phase load, which can effectively reduce the degree of three-phase output voltage.

Acknowledgments This work is supported in part by the China National Science and Technology Support Program under Grant 2013BAG21QB00, the National Natural Science Foundation of China under Grant U1134204 and the Fundamental Research Funds for the Central Universities under Grant 2012JBM099.

References

1. Zhang W, Wang H, Ren Y et al (2010) Investigation on control of three-phase grid-connected inverters under unbalanced grid voltage conditions. Trans China Electrotech Soc 25(12):103–110 (in Chinese)
2. Duan X, Luo Y, You J et al (2012) Control strategy based on impedance analysis for three-phase inverter with unbalanced load. Trans China Electrotechn Soc 27(12):56–62 (in Chinese)
3. Xu Y, Zheng J, Mei J et al (2012) Negative sequence voltage feed forward compensation unbalanced control strategy for threephase PWM rectifier based on one cycle control. Trans China Electrotechn Soc 27(11):183–190 (in Chinese)
4. Chen J, Diao L, Zhu K et al (2012) Research on control strategy of three-phase four-wire metro auxiliary converter. J China Railw Soc 34(4):34–38 (in Chinese)
5. Cross AM, Evans PD, Forsyth AJ (1999) DC link current in PWM inverters with unbalanced and non-linear loads. IEE Proc-Electr Power Appl 146(6):620–626
6. Tong Y, Yan Y (1993) The analysis of neutral point forming transformer. J Nanjing Univ Aeronaut Astronaut 25(3):341–350 (in Chinese)
7. Borup U, Enjeti PN, Blaabjerg F (2001) A new space-vector-based control method for UPS systems powering nonlinear and unbalanced loads. IEEE Trans Ind Appl 37(6):1864–1870
8. Gu H, Wang D, Shen H et al (2011) Research on control scheme of three-phase four-leg inverter[J]. Power Syst Prot Control 39(24):41–46 (in Chinese)
9. Zhang R, Prasad VH, Boroyevich D et al (2002) Three-dimensional space vector modulation for four-leg voltage-source converters. IEEE Trans Power Electron 17(3):314–326
10. Kong X, Wang J, Peng L et al (2003) The control scheme of three-phase voltage-source inverter output waveform based on internal model theory. Proc CSEE 23(7):67–70 (in Chinese)
11. Nian H, Zeng R (2011) Improved control strategy for stand-alone distributed generation system under unbalanced and non-linear loads. Renew Power Gener IET 5(5):323–331
12. Luo Y, Duan X, Meng F et al (2012) High performance repetitive controller for eliminating periodic disturbance of inverter with unbalance load. Power Energy Eng Conf (APPEEC) 2012:1–4
13. Nian H, Zeng R (2009) Research on a novel control strategy for unbalanced loads. Power Energy Eng Conf (APPEEC) 2009:1–4
14. Fortescue CL (1918) Method of symmetrical coordinates applied on solution of poly-phase networks. Trans Am Inst Electr Eng (AIEE) 37:1027–1140

15. Zhang Z, Li H, Zhang X et al (2011) Simulation and modeling of software phase-locked loop based on single/double synchronous coordinate system. Power Syst Prot Control 39(11):138–144 (in Chinese)
16. Fan S, Wang Z (2012) Grid synchronization technology for grid converters of wind power generation. Power Syst Prot Control 40(24):100 (in Chinese)

Chapter 16
Fluid and Thermal Analysis of Power Li-Ion Battery Pack and Experimental Verification

Dafen Chen, Jiuchun Jiang, Yaojuan Duan, Zhanguo Wang and Feng Wen

Abstract Fluid flow and heat transfer of a power li-ion battery pack which adopts the air cooling method with an electric fan are predicted using a computational fluid dynamics (CFD) software. The airflow rate distributions at coolant passages are analyzed quantificationally, which will affect temperature distribution of the battery pack. Temperatures at the center of battery surface that are located in the middle of each battery modules are monitored using thermal couples under the same driving cycle as simulation. The simulation results are close to the experimental ones, which indicate that the calculated results can be used as the basis of developing thermal management strategy of battery pack and optimization scheme of fluid flow.

Keywords Power li-ion battery pack · Thermal analysis · Fluid flow analysis

16.1 Introduction

Hybrid power traction locomotive (HPTL) can effectively reduce the fuel consumption and the associated emissions, which meets the national environment protection and energy policies. Therefore it has wide market prospects [1]. The component for energy storage in an HPTL mainly includes super capacitors, battery packs, and their combinations [2]. The energy storage component should have features of high specific energy and specific power, excellent high current

D. Chen (✉) · J. Jiang · Y. Duan · Z. Wang
National Active Distribution Network Technology Research Center (NANTEC), Beijing Jiaotong University, Beijing, China
e-mail: 11117360@bjtu.edu.cn

F. Wen
Beijing Epower Electronic Co., LTD, Beijing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_16,
© Springer-Verlag Berlin Heidelberg 2014

charge–discharge characteristics, long cycle life, environment friendly, and so on [3]. Because of its excellent performance, li-ion batteries become one of the most promising candidates of energy storage component in HPTLs. Actually, HPTLs such as HD300-910 already adopted li-ion battery packs as its energy storage device [4]. Generally, to ensure safety and cycle life of li-ion batteries, the environment temperature should be within the range of −25 to 55 °C during discharging and within 0–55 °C during charging [5]. Thus the cooling, heating, and thermal isolation schemes should be taken into account during designing of battery packs. Some researchers simulated the fluid and/or thermal fields of li-ion battery packs [6–9], but they barely verified the results in a real battery pack. In this paper, fluid and thermal fields of a power li-ion battery pack were simulated and we also monitored temperatures of a real battery pack to validate the simulation results.

16.2 Li-Ion Battery Pack Structure and CFD Model

The structure of the analyzed li-ion battery pack is shown in Fig. 16.1.

There are 88 batteries in the battery pack, which is connected in the way of 2 parallel and 44 series. They are divided into four modules, each module having 22 batteries. The modules are named as module A, module B, module C, and module D from inlet toward outlet. The gap between two batteries and the gap between battery and wall are all 1 mm. Each gap is an air passage, so there are 23 air passages in one module.

Because the battery pack structure is almost symmetry, in order to reduce calculation cost and running time, only half of the battery pack is constructed in the CFD model as in Fig. 16.2. Some geometrical features that are negligible for the fluid and thermal distribution are also simplified. When the fan is open, air from inlet flows to the bottom of battery modules, then flows through the air passages to the top of battery modules, and finally pumps out by the fan at the outlet. Between the bottom of battery modules, there are baffles with holes on them to change the airflow direction that makes the fluid distribution more uniform.

16.3 Flow Field Simulation Results of Battery Pack

Figures 16.3a,b show the airflow in the battery pack after the fan is started.

Because the area of inlet is relative smaller, the air velocity near inlet is higher than in other locations. The air from inlet is divided into four parts after it meets the baffles between module A and module B. The main part is squeezed to the ends of the baffles and flows into the next module. Another part flows through the gap in the batteries which is located in middle of module A and comes to the top of the battery pack directly; the third part forms a vortex, returning to the two corners

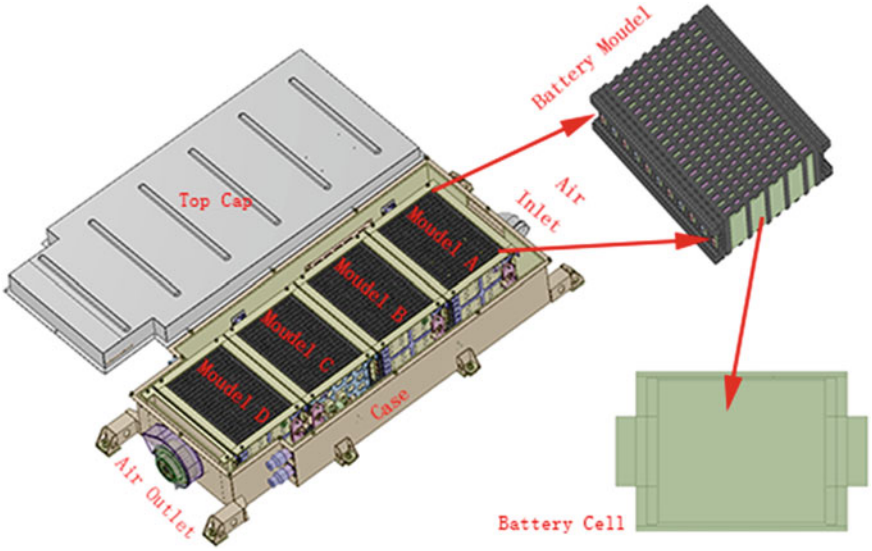


Fig. 16.1 Battery pack structure

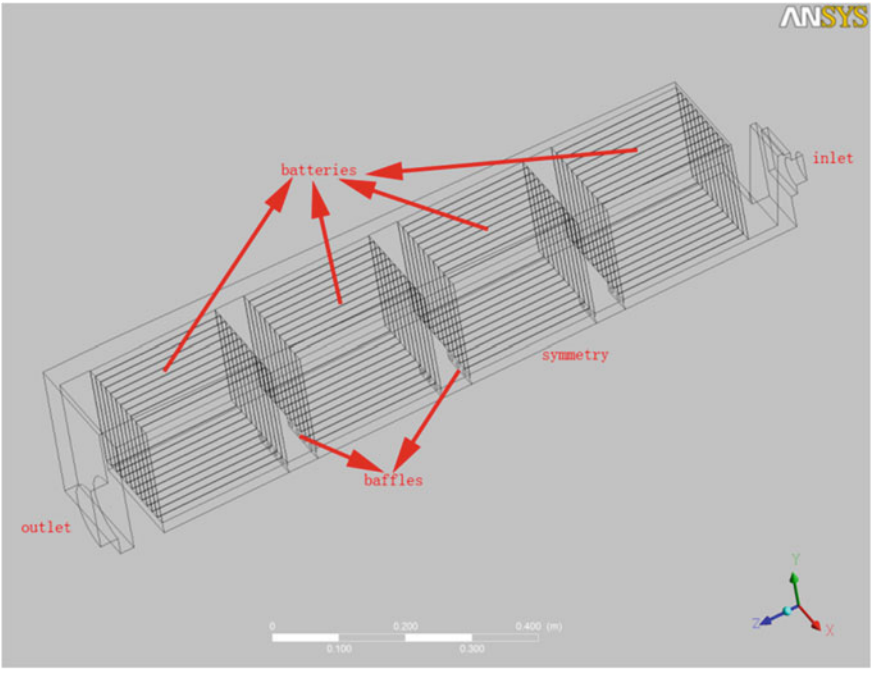


Fig. 16.2 Simplified CFD model

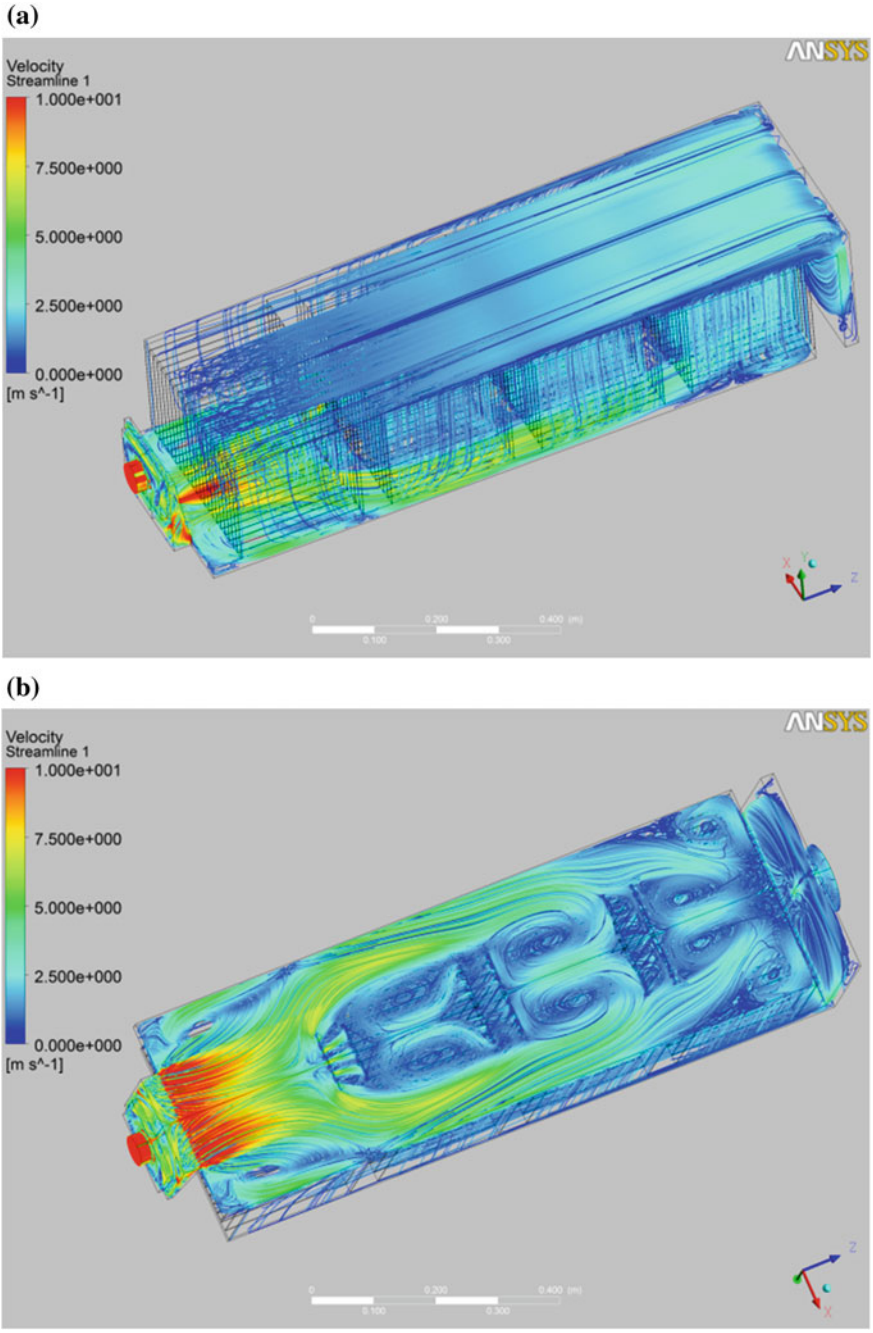


Fig. 16.3 **a** Streamline distribution in battery pack (*top view*). **b** Streamline distribution in battery pack (*bottom view*)

that are close to the inlet side and then flows to the top of module A; the rest flows through the holes on baffles into the module B directly.

Due to the block of baffles between modules A and B, the air that flows to other modules from the ends of the baffle has a relatively large velocity component that is perpendicular to the battery pack side walls. This velocity component leads airflow toward side walls. After rebounding by the wall, this part fluid has a velocity component in the direction pointing to symmetry plane. However, since part of the energy has been consumed during rebound, the velocity component is weakened. So the baffles between modules B and C capture a portion of the fluid flow back to the middle bottom of module B, and the baffles between modules C and D capture a portion of the fluid flow back to the middle bottom of module C, the part not captured by baffles flow to the bottom of module D.

The air at the top of modules is pumped out by the fan through a gradually increased air duct.

In order to regulate the battery temperature within a desired range and to reduce the battery cell to cell temperature variation in a battery pack, we often analyze the total flow rates and the uniformity of the fluid distribution. The fan parameters already fixed in this case can limit the temperature in the desired range sufficiently, so we mainly analyze the uniformity of the fluid field.

The heat dissipation power taken away by air that flows through the batteries gaps can be calculated as follows:

$$P_h = C_{\text{air}} q_{\text{air}} \left(\int_{A_{\text{out}}} T_{\text{out}} - \int_{A_{\text{in}}} T_{\text{in}} \right)$$

where P_h represents the heat dissipation power (J), C_{air} represents specific heat capacity of air (J/kg/ °C), q_{air} represents airflow rate(kg/s), T_{out} and T_{in} represent the air temperature at outlet and inlet surfaces respectively (°C).

Because all batteries are parallel from the view of airflow path in this battery pack, the capacity of heat dissipation is decided by the airflow rate. The airflow rates of all coolant passages are shown in Fig. 16.4.

The airflow rates are almost the same in modules C and D, while in modules A and B the airflow rates have a remarkable difference. The airflow rate is higher at the location near the side walls than the gaps located in the middle of module B. In module A, the gaps in the middle and two sides have higher airflow rates, while No.6 and No.19 air gaps have the lowest airflow rates. Module D has the highest total airflow rates, and module B has the lowest total airflow rates; the total airflow rates of module A are higher than module B and lower than module C. The unconformity of airflow rates at each coolant passages may lead to increase in the maximum temperature difference in battery pack, which will be analyzed in Sect. 16.4.

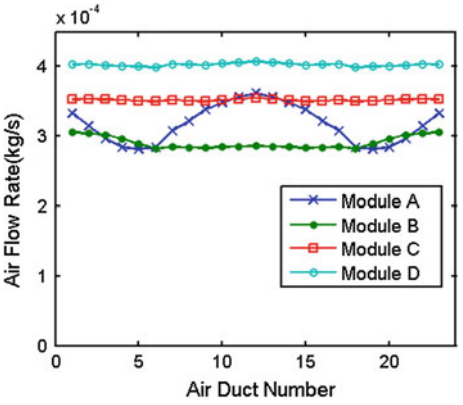


Fig. 16.4 Airflow rate in battery gaps

16.4 Thermal Field Simulation Results of Battery Pack

In order to verify whether the battery system can operate safely for a long time, the charge–discharge driving cycle at about SOC = 60 % (1C charge for 45 s and 1.5C discharge for 30 s) is simulated. The heat generation rate of single battery is calculated as follows:

$$P_{\text{avg}} = \frac{I_{\text{chg}}^2 R_{\text{chg}} t_{\text{chg}} + I_{\text{dch}}^2 R_{\text{dch}} t_{\text{dch}}}{t_{\text{chg}} + t_{\text{dch}}}$$

where P_{avg} is the average heat generate rate, I is the current, R is the total resistance, t is time, the subscript chg represents charge, and dch is discharge.

The initial temperature of battery pack is 27 °C and the fan is shut down. When the maximum temperature of the battery exceeds 40 °C the fan is started up. (approximately after the time is 6,000 s in the Simulation). The battery temperature distribution at the moment of the fan startup is shown in Fig. 16.5.

It can be seen from the figures that before the fan is on, the temperature difference is small; the highest temperature region is located at the top of modules A, B, and C; the cooler areas are the top of module D and the bottom of all modules. Because the top of module D is near the outlet side, its top space is the largest and the outlet air distance is shorter and the cooling capacity is stronger than other modules. Though the fan is off at this point, the bottom of the module temperature is lower because there is a certain airflow in the battery box and the hot airflow toward the tip. Because module A is near the air inlet and has heat exchange with the inlet air, thus the bottom of the battery module A has lower temperature than the bottom of other modules.

After the fan is started (Fig. 16.6a), the temperature of most regions in the battery system decreases due to forced air cooling; Because the airflow rate between the batteries is not equal, the temperature difference in the battery system increases; the hotter region corresponds to the small airflow rate area as shown in

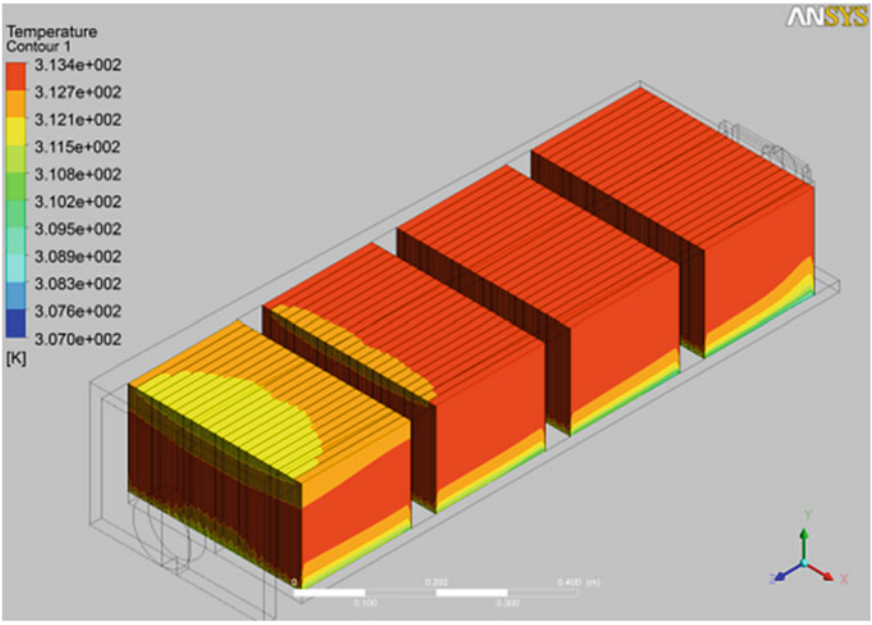


Fig. 16.5 The temperature distribution of battery surfaces at the moment of fan startup ($t = 6000$ s)

Fig. 16.4. The temperature at the bottom of the battery is lower, at the top temperature is higher, and the maximum temperature difference is about 8 °C (Fig. 16.6b).

Figure 16.7 shows the temperature curves of the center location at the surface of the battery which is located at the middle of each module. Before opening the fan, the temperature in different regions inside the battery pack is basically the same, while after turning on the fan, the significant temperature difference among the areas appears and the airflow rate is higher where the temperature is lower. The final temperature gap is about 3 °C.

16.5 Experimental Results and Discussion

Several temperature sensors are installed at the center of the each module in the actual battery pack, and the battery pack is charged and discharged using the same driving cycle condition with the set of the simulation conditions. Finally the temperature curve is obtained and shown as in Fig. 16.8a.

When the fan is shut down, the temperature difference of the battery at different modules is already obvious which may be due to the heat generated rates difference. After turning on the fan, the decline rate of temperatures shows that the

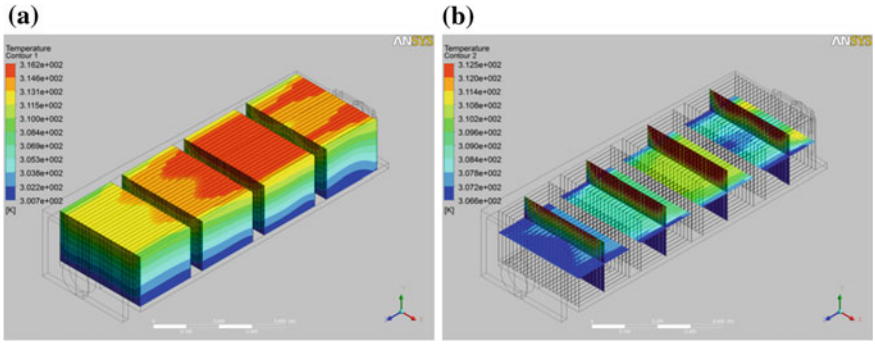


Fig. 16.6 **a** (*left*) The temperature distribution at battery surfaces after the fan startup 3,000 s (t = 9,000 s). **b** (*right*) The temperature difference inside batteries after the fan startup 3,000 s (t = 9000 s)

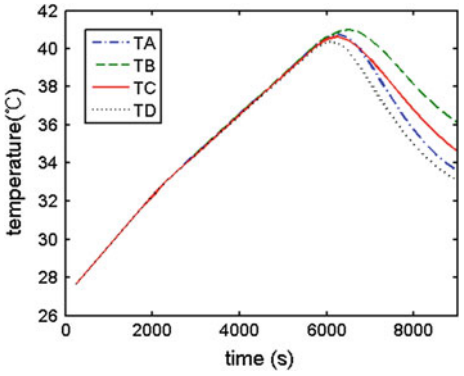


Fig. 16.7 Temperatures of each module center (simulate results)

temperature decline speed of module D is the fastest, temperature decline speed of module A and module C are almost the same, and module B is the worst.

As shown in Fig. 16.8b, the absolute error between simulation results and the experimental results is less than 4 °C, presenting that although the simulation ignores some factors including the part geometry of actual system, inconsistency between the batteries and the battery pack sealing, the simulated results are basically consistent with the experimental results.

16.6 Conclusion

A model of CFD is built to analyze the fluid field and thermal field of a power li-ion battery pack. The simulated results agreed well with the experimental results. From the simulation results we can conclude that: in battery pack that

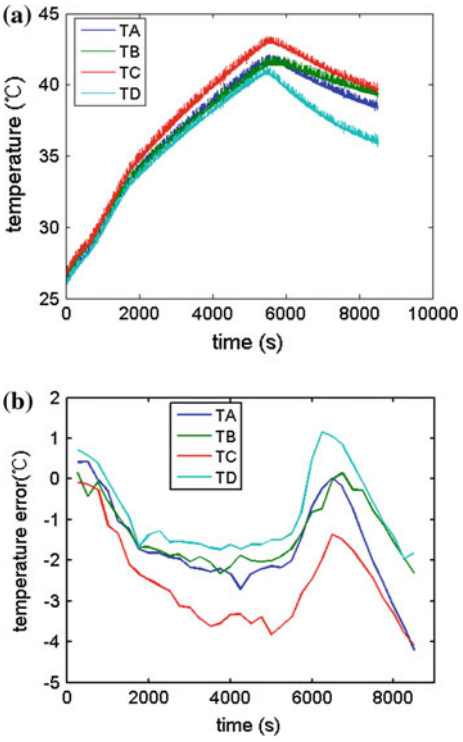


Fig. 16.8 **a** Temperatures of each module center (experimental results). **b** Temperature errors of simulated and experimental results

adopts parallel cooling construction, the fluid flow rate flow through each coolant passage affects the thermal distribution directly, thus improved uniformity of the fluid flow rate is very important. In addition, the inconsistency of batteries and the seal condition will affect the distribution of fluid field and thermal field and should be taken into account during battery pack design and manufacture.

Acknowledgments The work is supported by the National Key Technology R&D Program (2013BAGL1B00) and the Fundamental Research Funds for the Central Universities (2013YJS081).

References

1. Tian R, Kang Y (2012) Overview of the development of the hybrid locomotive at home and abroad. Diesel Locomotives 10:1–3 (in Chinese)
2. Liu Y (2012) New energy strategy thinking about rail power traction. Electr Locomotives Mass Transit Veh 35(5):1–4 (in Chinese)
3. Ye D, Peng C, Liu S (2012) The development of hybrid traction shunting locomotive. Electr Drive Locomotives 5:17–20 (in Chinese)
4. Tiant Z (2011) Japan HD300-901 type hybrid locomotive prototype. Foreign Diesel Locomotive 5:10–14 (in Chinese)

5. Ma X, Gao D (2012) Application feasibility study on power lithium-ion battery for double energy locomotive. Electr Locomotives Mass Transit Veh 35(6):53–56 (in Chinese)
6. Heesung P (2013) A design of air flow configuration for cooling lithium ion battery in hybrid electric vehicles. J Power Sources 239:30–36
7. Sungjin P, Dohoy J (2013) Battery cell arrangement and heat transfer fluid effects on the parasitic power consumption and the cell temperature distribution in a hybrid electric vehicle. J Power Sources 227:191–198
8. Li X, He F, Ma L (2013) Thermal management of cylindrical batteries investigated using wind tunnel testing and computational fluid dynamics simulation. J Power Sources 238:395–402
9. Xun J, Liu R, Jiao K (2013) Numerical and analytical modeling of lithium ion battery thermal behaviors with different cooling designs. J Power Sources 233:47–61

Chapter 17
A Novel Real-Time Tension Monitoring Method for Overhead Contact System

Xin Meng, Jinghan He, Xiaojun Wang and Guomin Luo

Abstract This paper presents a novel real-time tension monitoring method (RTMM), based on angular displacement of gear wheel and vibration of concrete weights, used to detect the service state of catenary without physical contact with the contact line. The classical detection of contact wire which relies on the inspection vehicle is not suitable for real-time monitoring. The method chooses the characteristic parameters which could reflect the tension of catenary with non-contact to determine whether the wire could still work normally or not. The paper also gets a test result which uses the new method to detect the tension of catenary. The result shows the method could monitor the catenary force safely and effectively.

Keywords Overhead contact system · Tension monitoring · Iterative calculation

17.1 Introduction

With fast development of high-speed railway in China, the overhead contact method (OCS) is an important part of the traction power supply method [1]. The structure of OCS is complex relatively, so if there is something wrong with the OCS, such as line breaks and certain part loose, the power supply section which the fault located in will lose energy as the consequence of impacting on the whole railway method. Therefore, it is extremely significant to design and realize a novel real-time monitoring system (RMS) for monitor the service state of OCS.

Tension, as a basic requirement for OCS, has been specified. Table 17.1 shows the tension values in different countries with different speed levels. Fatigue wear

X. Meng (✉) · J. He · X. Wang · G. Luo
Beijing Jiaotong University, Beijing, China
e-mail: 10117335@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_17,
© Springer-Verlag Berlin Heidelberg 2014

Table 17.1 Tension values for different countries

Country	Suspension way	Running speed (km/h)	Tension value (kN)
France	Simple catenary	300	20
Germany	Stitched catenary	350	27
Spain	Stitched catenary	350	31.5
China	Stitched catenary	350	27

leads to the change of the tension. Change in the tension value may influence the catenary-pantograph interaction, thereby affecting high-speed train's safety. But it's hard to monitor the tension value on-line, because contact line always vibrate when pantograph pass or there is a windy weather.

17.2 Tension Monitoring Technique

The technology of OCS geometry parameters, including tension force, detection which measures the relevant parameters by optics and laser has been used for more than 40 years in the detection field [2]. And a series of OCS geometry parameters detection methods used in different situations have been invented after a great deal of researches [3]. For normal traction method with low speed, measuring the value of tension in some months, could meet the practical needs enough. Nevertheless, as to the high-speed lines designed for maximum speeds of more than 300 km/h, more attention should be paid to the tension which could reflect the operational status of the OCS [4].

At present, the most representative monitoring methods of OCS are designed by Italian and German. Italian methods, which are called noncontact detection mode, emphasize on the test of concrete weights. German methods, which are called contact detection mode, emphasize on test the angular of gear wheel. Detection methods of OCS made in Italy mainly adopt laser irradiation, servo tracking and image processing technology, which can measure the weight of concrete weights and the height of the contact lines accurately [5]. However, the tension cannot be detected by this method precisely which cannot adapt to the high-speed railway contact method because of the huge amount of image processing. Detection methods of the OCS developed by Germany focus on studying the contact pressure, which could reflect the change of tension, and hard point [6], by this method, accurate dynamics parameters and intuitive judgment of the OCS can be gained. But this method does not suitable for China not only because it is too expensive, but also because it owns many disadvantages.

The paper provides a novel tension monitoring method based on the vibration of concrete weights and the angular of gear wheel to realize real-time monitoring for high-speed railway. The paper also provides a test result which use this new method on real high-speed railway to prove the validity of the method.

17.3 Description of New Method

17.3.1 The Gear Wheel Angular Introduction

Tension compensation device, as an accessory device, provides enough tension for contact line through pulley wheel which ratio is 1:3 [7]. Tension compensation device directly connected with the contact line through the insulator, so it has the same dynamic characteristics [8]. The method determines the contact line's tension through monitor the force on the tension compensation device. Figure 17.1 shows the structure of tension compensation device.

From Fig. 17.1, in the normal condition, the gear wheel has a drop angular which to make both sides of lines keep balance. So, from the Newton's law, the moment balance equation is

$$\vec{F_T} + \vec{F_C} = \vec{F_p} \tag{17.1}$$

where:

F_T = the tensile force of the tension line

F_C = the tensile force of the contact line

F_p = the tensile force of the accessory pole

So the angular θ mark in Fig. 17.2, named gear wheel angular, can be considered

$$\tan\theta = \frac{F_T}{F_C} \tag{17.2}$$

17.3.2 Concrete Weights Introduction

Concrete weights, as the end of the accessory device, make the tensile force of whole system stability. The change of the concrete weights could make the tension of catenary change immediately.

The force balances equation of concrete weights with no vibration is:

$$F_{T0} = m_C g \tag{17.3}$$

where:

F_{T0} = Initial value of tension (N)

m_C = Weight of concrete weights, constant (kg)

g = Constant $g = 10\ \mathrm{m/s^2}$

When contact line vibrates with acceleration a

$$a = a(t) \tag{17.4}$$

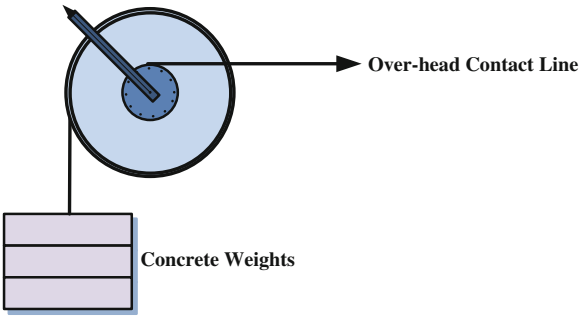


Fig. 17.1 Structure of tension compensation device

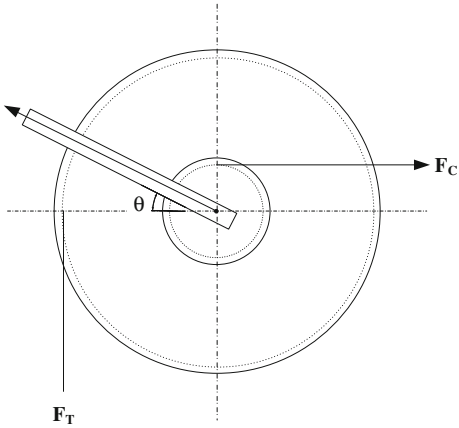


Fig. 17.2 The stress analysis diagram

So, at any time, there is force equation:

$$F_T = m_C[g + a(t)] = m_C a(t) + F_{T0} \tag{17.5}$$

In fact, the acceleration of concrete weights $a(t)$ is easy to be detected by the acceleration sensor, because there is no electric and magnetic interference. From the detected $a(t)$, the value of tension (F_T) is also obtained by (17.5).

But considering about the environment influence, Eq. (17.5) could not equal all the time. So, Eq. (17.5) can be change into:

$$F_T = m_C[g + a(t)] = m_C a(t) + F_{T0} + \Delta \tag{17.6}$$

where: Δ is the errors existing in environment and measurement.

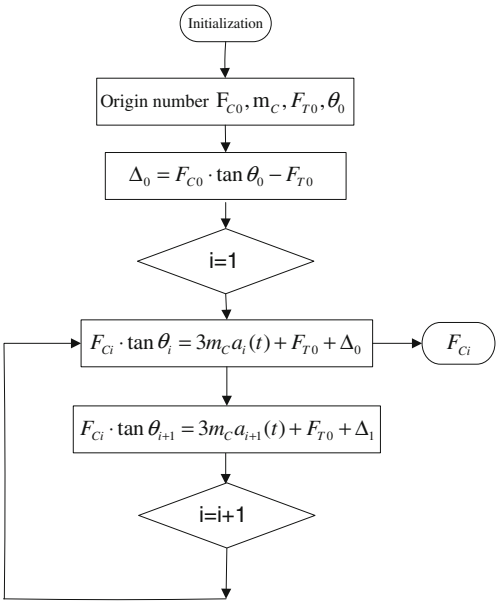


Fig. 17.3 The flow chart of the method

17.3.3 Calculation Description

From Eqs. (17.2) and (17.6), the following relationship can be considered:

$$F_C \cdot \tan\theta = m_C a(t) + F_{T0} + \Delta \tag{17.7}$$

In the Eq. (17.7), m_C and F_{T0} are the basic parameters of contact system, and θ and a can be measured by simple sensors, so F_C can be calculated by Δ. But the errors Δ changes all the time and there is no way to measure it. So in order to calculate F_C, this paper gives an iterative calculation method. There is the flow chart of the method mentioned in this paper in Fig. 17.3.

17.4 Experiment and Result for the Method

In order to research the performance of novel method, this paper conducts an experiment on high-speed railway to prove. The basic parameters are given by Table 17.2.

In this experiment, the monitors are installed both on the accessary pole and the concrete weights. One man-made fault is given at the end of the contact line, and make gear wheel angular and acceleration of concrete weights changed. The monitors can record the value of the angular and acceleration. Sampling frequency is 1 kHz. Figures 17.4 and 17.5 show the gear wheel angular and the acceleration of the contact line.

Table 17.2 Basic parameters for experiment high-speed railway

Part name	Unit	Value of number
Concrete weights (m)	kg	900
Initial tensile force (F_{T0})	kN	9
Initial tensile force for contact line (F_{C0})	kN	27
Gear wheel angular (θ)	°	18.4

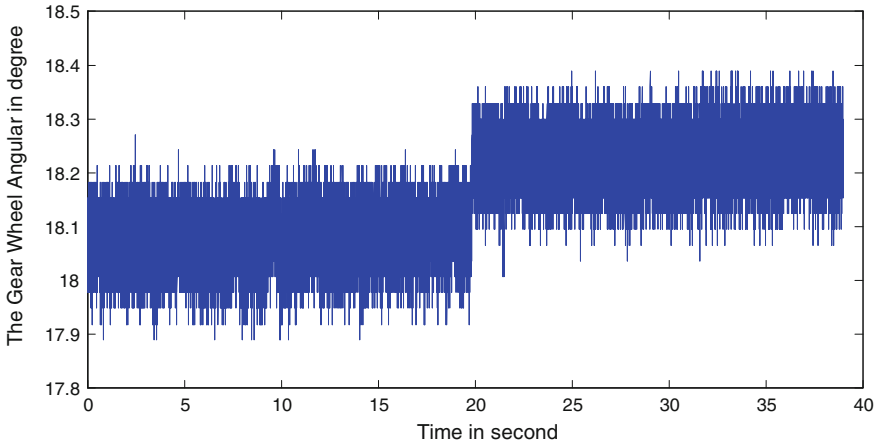


Fig. 17.4 The gear wheel angular

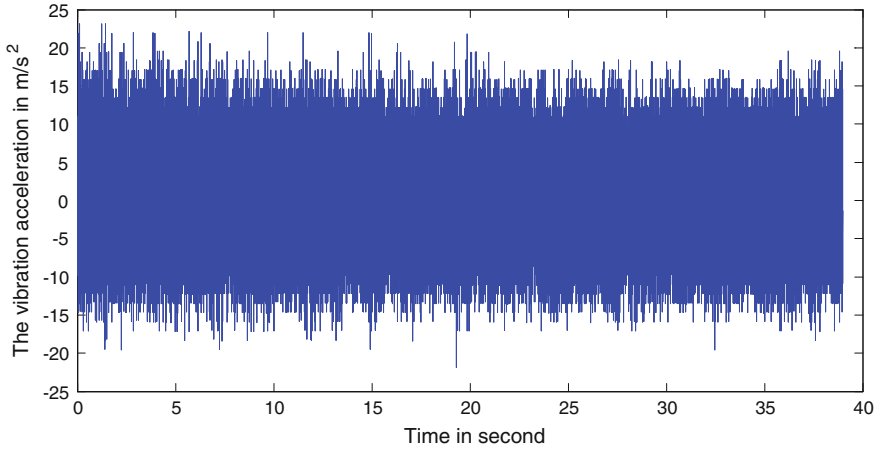


Fig. 17.5 The vibration of the concrete weights

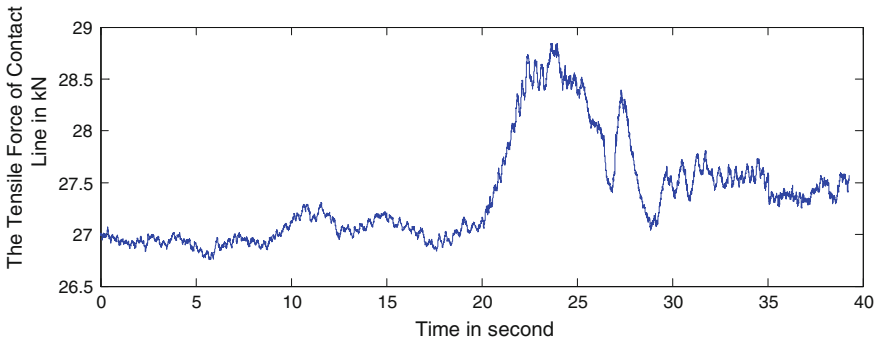


Fig. 17.6 The tensile force of the contact line

A man-made fault is given at nearly 20 s, the change of gear wheel angular is significant, and because of the damping effect, the vibration of the concrete weights is changed not obviously. According to the flow chart mentioned above in Fig. 17.3, the value of tensile force of the contact line $F_C$could be calculated as shown in Fig. 17.6.

The calculation result shows that: there is a fault at 20 s; the fault makes the tension of contact line gradually changed from one step to another step. As the figure shows, the tension of contact line up to 28.5 kN from 27 kN, then after a series of fluctuations, the tension get back to normal value again. All these process is in line with the actual situation, so the novel tension monitoring method mentioned in this paper is effective.

17.5 Conclusions

This paper has presented a novel monitoring method used to monitor the tension of contact line system. One important objective of the designed method is that the measurements are performed without physical contact with the contact line. In the paper, concrete weights vibration and gear wheel angular have been chosen as characteristic parameters to calculate the tension of contact line. This method can detect contact line's working state through comprehensive judgment. To validate the method's feasibility, an experiment has been installed and the satisfactory results of the test prove the designed method gets the proposed functions.

Acknowledgments This work is supported by the National High-tech R&D Program of China ('863'Program) (No. 2011AA11A102) and the Fundamental Research Funds for the Central Universities (NO. 2013YJS086).

References

1. Huang X, Sun Q, Han X (2008) An on-line monitoring method of temperature of conductors and fittings based on GSM SMS and Zigbee. In: 3rd IEEE conference on industrial electronics and applications. pp 1522–1527
2. Hofler H, Dambacher M, Dimopoulos N, Jetter V (2004) Monitoring and inspecting overhead wires and supporting structures. *Intell Veh Symp* pp 512–517
3. Fararooy S, Mair C (2004) Review of railway overhead wire geometry measurement methods. *Curr Collect High Speed Trains Semin* (Ref No 1998/509), pp 4/1–4
4. Poetseh G, Evans J, Meisinger R (2007) Pantograph/eatenary dynamics and control. *Veh Method Dyn* 28(2–3):159–95
5. Huang C, Zhang J, Fan X (2007) Real-time safety status monitoring and failure early warning method for railway locomotive. *Electr Drive for Locomotives* 4:62–66 (in Chinese)
6. Wu C, Lu J, Zhou H (2007) The on-line real-time detection and broken line accident early warning method on railway catenary cable stress load. *Control Autom* 19:102–103+157 (in Chinese)
7. Liang Y (2012) Steady-state thermal analysis of power cable methods in ducts using streamline-upwind petrov-galerkin finite element method. *IEEE Trans Dielect Electr Insul* 19:1
8. Kawase Y (2000) Heat analysis of a fuse for semiconductor devices protection using 3D finite element method. IEEE Trans Magn 36(4):1377–1380

Chapter 18
Locomotive DC600 V Power Supply Design and Control Redundancy Improvement

Yan Xia, Chunmei Xu and Lei Wang

Abstract DC600 V rectifier power supply provides DC power for the auxiliary power equipment (such as air conditioner, water heater, etc.) of the electric locomotive. Because of its reliability and stability, DC600 V rectifier power supply is widely used. This article focuses on the design of the main circuit of DC600 V and the improvement of the control redundancy.

Keywords DC600 V · Half-controlled rectifier · Redundancy control

18.1 Introduction

With the continuous development of the railway enterprise, people have a higher requirement for comfort when they are travelling. It is associated with the increasing civilian equipment on passenger trains, such as air conditioning, water heater, etc. The operation of these equipment needs the train to supply power. The DC600 V power supply system could not only satisfy the requirement of the driving demand, but also steadily supply the power [1].

18.2 Design of the Circuit

18.2.1 Selection of the Rectifier Bridge

The railway passenger locomotives use a single-phase power supply, which is suitable for the DC600 V power supply system to use a bridge-phase controlled rectifier circuit. The bridge-phase controlled rectifier circuit is divided into single-

Y. Xia (✉) · C. Xu · L. Wang
School of Electrical Engineering, Beijing Jiaotong University, No.3 Shangyuancun, Beijing, China
e-mail: xia333yand@163.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_18,
© Springer-Verlag Berlin Heidelberg 2014

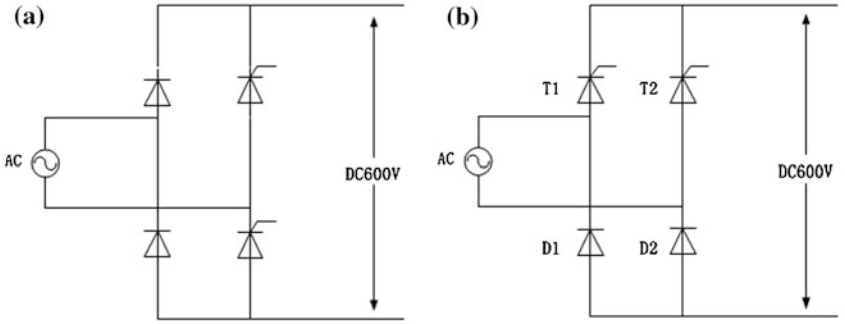


Fig. 18.1 Half-control bridge rectifier circuit. **a** Series connection, **b** Parallel connection. Comparing these two patterns, (**a**) has two advantages over (**b**)

phase full control circuit and single-phase half-controlled bridge rectifier circuit. Compared to the former, the half-controlled circuit has the advantage of simple structure and simple control. In addition, the half-controlled circuit can totally meet the requirement, compared to the full controlled circuit. Besides, the full controlled circuit has the disadvantage of low power factor when the supplied power is low. In conclusion, DC600 V rectifier circuit should use the single-phase half-controlled rectifier circuit.

The single-phase half-controlled rectifier circuit can be divided into two kinds; the series connection, and the parallel connection. It is shown in Fig. 18.1.

① Under certain conditions, the parallel connection pattern will "lose control," but the series connection pattern will not make it happen. Working with the parallel connection pattern, if we cut off the trigger pulse suddenly or increase the phase control angle α to 180°, "lose control" will happen. The essence of this phenomenon is out of control of the SCR. For example: in Fig 18.1b, when T1 is conducted, we cut off the trigger pulse or increase the phase control angle α to 180°, then T2 cannot be conducted. It leads to the situation that when the input AC u is in the positive half-cycle, T1 and D2 are conducted, which make the DC output; when the input AC u is in the negative half-cycle, D1 substitutes for D2 to be the rectifier tube and T1 continues to conduct current if only the load inductance is large enough. At this time, the output voltage is 0 V. Hence, whatever the condition, T1 is always conducted, which will make it so hot that it may be damaged. However for the series connection pattern, the after-flow is finished by D1 and D2. In this situation, SCR will be shut off when the circuit naturally commutates. The diode which can withstand high current can easily be made. With the series connection pattern, the SCR will not be damaged for the long conduction. Because of these reasons, "lost control'' can be prevented [2].

② For the series connection pattern, the average value and the root mean square value of the current which passes through the diode is larger than that which passes through the SCR when the phase control angle α is larger than 0. The rectifier bridge has high requirement for the diode, because when the train starts, the starting current

is 2 times the rated current. At the same time, when the system starts, the phase control angle α will be large and the rectifier voltage will be low. So considering the average value and the root mean square value of the component, the rating of the diode is larger than that of the SCR. The large rating diode is easy to be made [3].

In conclusion, DC600 V phase control rectifier circuit should use the half-control series connection rectifier circuit which is more efficient and has more advantages.

18.2.2 Selection of the Rectifier Diodes and SCRs

The main circuit uses the half-phase control method. The bridge uses two diodes and two SCRs. The analysis of the withstand voltage and current is as below:

1. withstand voltage–
   When the system is working and the diodes or SCRs do not conduct, the voltage of the input AC860 V is on the tubes. At this time, the largest reverse withstand voltage can be calculated as follows (18.1):

$$\sqrt{2}U_2 = \sqrt{2} \times 860 = 1216\,\text{V} \tag{18.1}$$

Considering the redundancy (18.2):

$$U \geq (1.5 \sim 2) \times \sqrt{2}U_2 = (1824 \sim 2432)\,\text{V} \tag{18.2}$$

2. withstand current—
   When the system is working with full load, the largest current passing through the tubes is 667 A. So the effective value of the current can be calculated as follows (18.3) [4]:

$$I = \sqrt{\frac{1}{2}}I_d = 471.6\,\text{A} \tag{18.3}$$

So the nominal current can be calculated as follows (18.4) [4]:

$$I_T = (1.5 \sim 2) \times 471.6 = (707.4 \sim 943.2)\,\text{A} \tag{18.4}$$

3. At the same time, considering the heat caused by the loss, DC600 V system should use the forced air cooling.

In conclusion, the SCRs are finally selected as KPA1200-36, whose nominal voltage is 3600 V and nominal current is 1200 A; the diodes are finally selected as ZPA2000-36, whose nominal voltage is 3600 V and nominal current is 2000 A.

18.2.3 Design of the Whole Circuit

The power capability of DC600 V phase control rectifier power system is designed as 400 kW*2. The whole circuit includes two main rectifier circuits which are respectively called Circuit I and Circuit II. They are all the same. Each power capability is 400 kW, totally 800 kW. The system gets power from the auxiliary winding of the traction transformer. The input voltage is AC860 V and the output DC600 V is paralleled at the train DC600 V bus [5, 6]. The two main circuits are controlled by the control system to supply the power for the train. From the input to the output, each circuit includes these components as: the main contactor; the fuse; the RC absorbing circuit; the overpressure varistor; the current sensor; the voltage sensor; the rectifier bridge; the LC filter circuit; the ground fault protection circuit, etc. According to the design requirements of the DC600 V phase control rectifier power and the selection of the rectifier bridge, the whole circuit DC600 V phase control rectifier power is designed as in Fig. 18.2 [7].

18.3 Simulation of the Main Circuit

Considering the design of the main circuit, the simulation model is built as in Fig. 18.3.

The AC side is set an AC power, which supplies the power to the system. The bridge concludes 2 SCRs and 2 diodes. The pulse of the SCRs is supplied by the control system. The DC side consists of a smoothing reactor L and a filter capacitor C, which remove the harmonics.

The parameter of the main circuit is set as: L = 9 mH, C = 2*5700 uF. When it is working with full load, the simulation load resistor is 0.9 Ω.

DC600 V phase controlled rectifier power uses the method of closed-loop voltage control to make the output voltage stable. The output voltage of the main circuit is detected and then sent to the control system to be dealt with. The simulation approach is as follows: When the detected voltage is larger than 600 V, the module "Relational Operator" will not produce drive pulses and the SCRs will be shut off. So the output voltage will gradually decrease. When the detected voltage is smaller than 600 V, the module "Relational Operator" will produce drive pulses and the SCRs will be conducted. So the output voltage will gradually increase to 600 V again. So repeatedly, the output voltage will maintain 600 V with slight fluctuation. By adjusting the PI parameter Kp or Ki, the response speed and oscillation can be controlled.

In this simulation, the final PI parameter is set as in Fig. 18.4 and at this time the wave of the output voltage is the best.

At this time, Kp = 0.7, Ki = 5, the output voltage and current is as in Fig. 18.5 (the upper is the voltage wave and the lower is current wave).

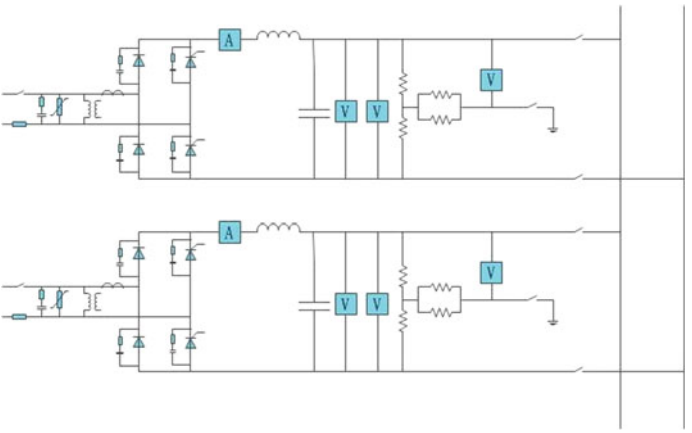


Fig. 18.2 DC600 V power supply main circuit

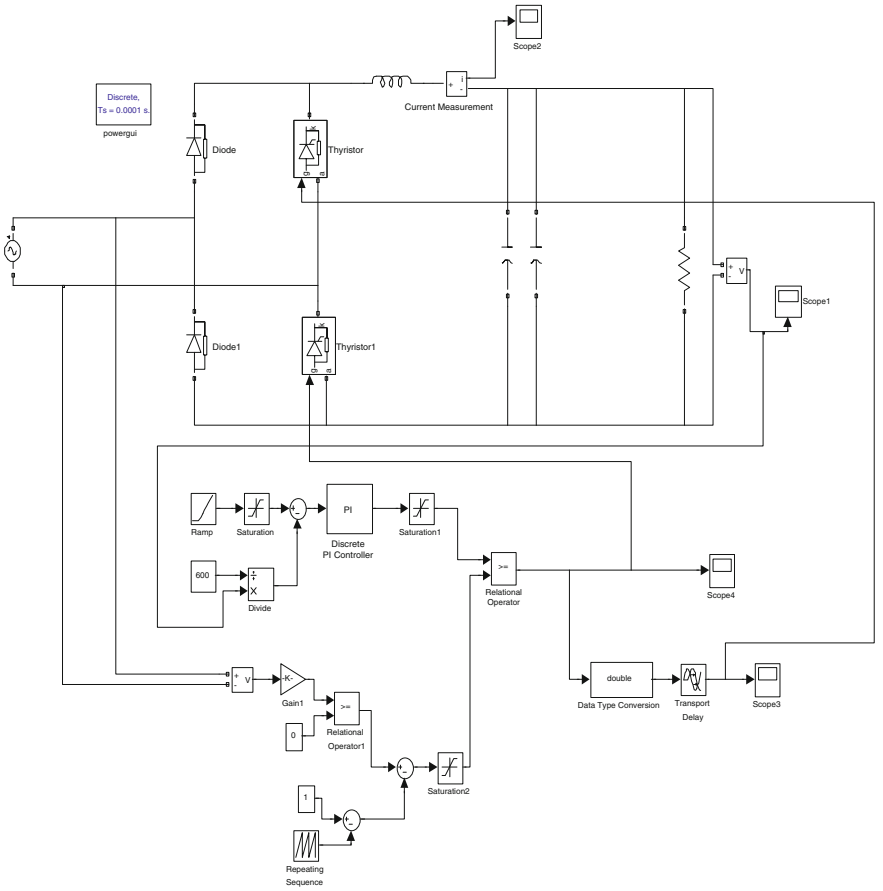


Fig. 18.3 DC600 V simulation model

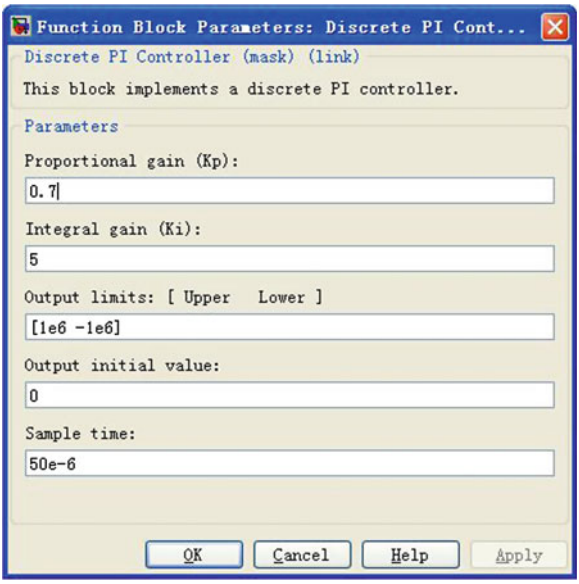


Fig. 18.4 PI parameter

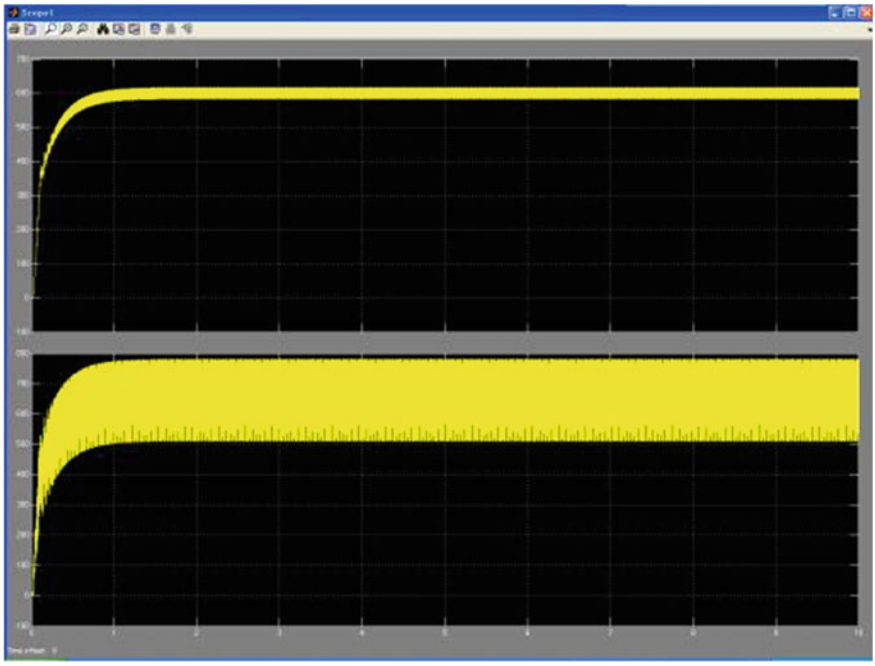


Fig. 18.5 Output voltage and current wave

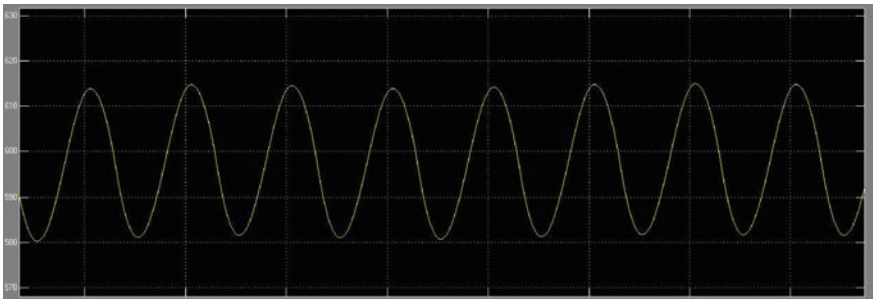


Fig. 18.6 Output voltage wave

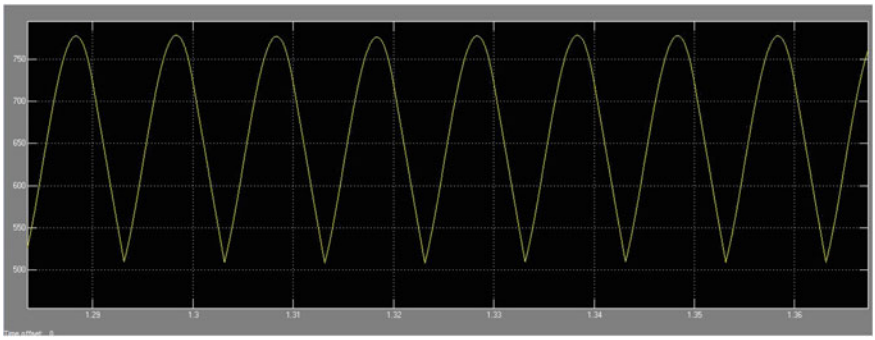


Fig. 18.7 Output current wave

From this figure, we can see that it takes about 1 s to reach 600 V. After 1 s, the output voltage wave is stable.

The stable output DC600 V wave is as in Fig. 18.6.

From this figure, we can see the output voltage fluctuates between 580 and 620 V and the average value is 600 V, which meets the requirement of ripple wave.

The stable output current wave is as in Fig. 18.7.

The current is 667 A and the ripple is ±130 A.

From the simulation result, we can see that when L = 9 mH, C = 2*5700 uF, the output power meets the requirement of the design and the dynamic response of the output wave is very good. So the selection of L, C is suitable, which can meet the requirement.

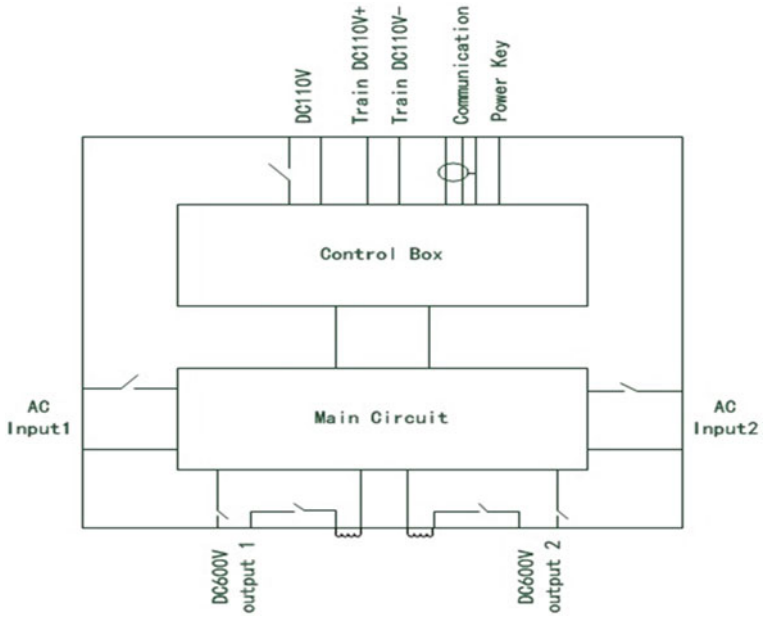


Fig. 18.8 Cabinet structure schematic diagram

18.4 Redundancy Improvement of the Control System

DC600 V phase control rectifier power is finally designed as a power cabinet. Except for the main circuit, there is a control box to control the main circuit. The structure is as in Fig. 18.8.

A control box needs to control the two main circuits. The traditional way only needs one control system. For us, in order to improve the reliability of the control box, we design a double control system plan, which makes the redundancy improvement. This means that there are two main control boards in the control box, which are called *main control board 1* and *main control board 2*. They have the same function. Each can control the two main circuits. *Main control board 2* acts as the redundancy of the *main control board 1.* When the *main control board 1* breaks down, the system will change to the *main control board 2.* Thus it makes sure that the system can operate normally.

The DC600 V system works as follows: *main control board 1* and *main control board 2* work as redundancy for each other. Each can control the main circuit to operate. When it operates normally, the *main control board 1* works as the main control system, which makes the DC600 V power system supply power. The *Main control board 2* works as the standby control system. Only when the *main control board 1* does not work, the system will make a change to use the *main control*

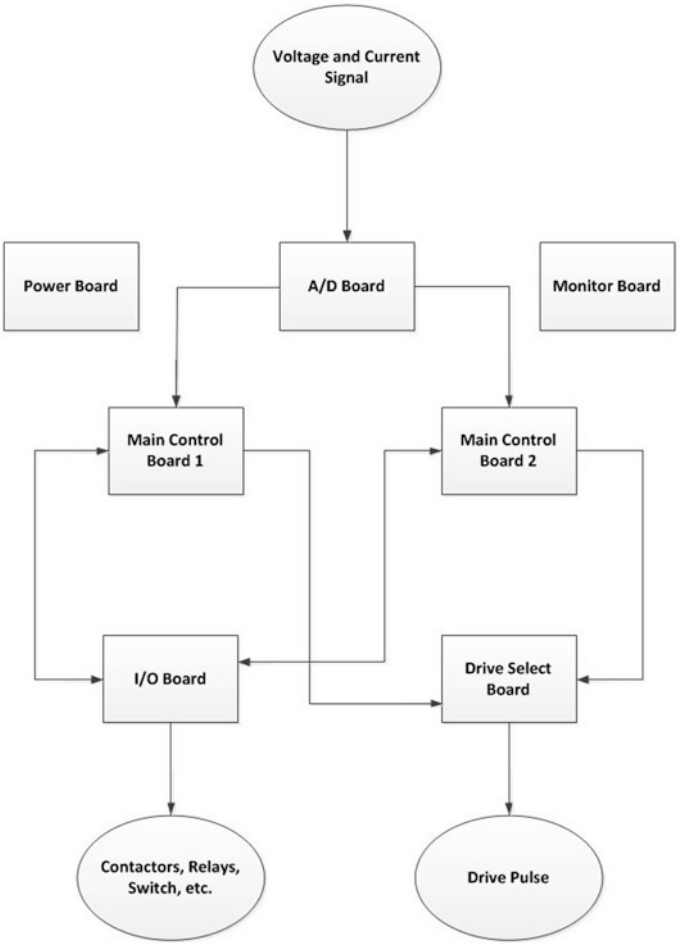


Fig. 18.9 Control system diagram

board 2 to replace it. The relationship of the boards in the control box is as in Fig. 18.9. Thus there are two control systems for the DC600 V system to make sure that it operates steadily. So it makes the redundancy improvement.

Acknowledgments This work was supported by the Beijing City Traffic Industry Science and Technology Project under grant E13M00020. We would like to thank the anonymous reviewers for their comments and suggestions.

References

1. Xu L (2001) DC600 V train power supply system.Railw Veh 39(3):13–16 (in Chinese)
2. Liu Z (2004) Power electronics. Tsinghua University Press, Beijing Jiaotong University Press, Beijing, pp 53–90 (in Chinese)
3. Luo B (2003) Passenger train DC600 V power system optimization suggestions for improvement. Railw Veh 41(3):31–32 (in Chinese)
4. Zhang X, Li H (1998) Actively develop the locomotive to the passenger train power supply technology. Electr Drive Locomot 3(5):53–56 (in Chinese)
5. Wu Q (2003) Power supply system of passenger train. Electr Drive Locomot 2(5):54–61 (in Chinese)
6. Wu Q (2002) Locomotive auxiliary inverter technology development. Electr Drive Locomot 1(3):4–7 (in Chinese)
7. Li G (2003) Auxiliary power supply system of high-speed train at home and abroad. Electr Drive Locomot 1(5):57–61 (in Chinese)

Chapter 19
The Research on BEV-Based Urban Passenger Transport Environment

ZhiQiang Rao, ZiJian Zhuang and Bin Ning

Abstract In view of the mass severe pollution and frequent haze encountered by cities, to find a solution to the bottleneck in the urban passenger transport environment, environment-friendly battery electric vehicles (BEVs) are proposed for urban public transit as a substitute for high-energy consumption high-pollution traditional buses. The immense emission reduction potential and superior overall performance of battery electric buses were yielded after years of follow-up analysis and online measurement, offering scientific guarantee for the coordination between the urban transport environment and resources.

Keywords BEV · Urban passenger transport · Effect of emission reduction · Transport environment

19.1 Introduction

As the national economy sprawls and urbanization progresses, riding on vehicles has become a dominant way of travel in recent years. Different from developed economies where people normally travel in their private cars, public transit tools are still the first choice for the Chinese population. However, mass severe air pollution and haze has been reported recently in Central East China, attributable to the excessive emission of production or domestic waste gasses. Pollutants from motorized vehicles account for approximately 25 % of atmospheric pollution.

Z. Rao (✉) · Z. Zhuang
China Center for Industrial Security Research, Beijing Jiaotong University, Beijing 100044, People Republic of China
e-mail: rao_hua1@163.com; 657176805@qq.com

B. Ning
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, People Republic of China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_19,
© Springer-Verlag Berlin Heidelberg 2014

Besides CO_2, hazardous gasses contained in vehicle exhaust emissions include carbon oxide, hydrocarbon (HC), and nitrogen oxide (NO) which are the main contributors to air pollution. Besides, the enormous quantity of automobiles in service will mean huge pressure on energy and environment [1]. Traditional petroleum-based automobile industry will be faced with great challenge. China will be facing even fiercer challenge in its petroleum safety and transportation energy. Energy consumption is becoming a bottleneck for the nation's transportation business. To improve the urban air environment and quality, one of the best solutions could be battery electric buses.

19.1.1 Performance of Electric Motor Coaches

Battery electric cars made in China outscore large international automobile manufacturers like Citroen [2]. We have gained key expertise for battery electric vehicles (BEV) and capacity for industrialized package production, having manufactured in small lots and sold them outside China. Electric motor coaches and cars can be fast charged, needing 15 min–2 h for electric cars, and 3 h for electric motor coaches to come to 75–80 % of the battery capacity. Table 19.1 presents the performance of BEV made in China.

Batteries currently used in electric vehicles include lead-acid batteries, NI-MH batteries, lithium batteries, and super capacitors. Table 19.2 compares the performance among these batteries.

19.2 Measurement of Actual Power

The power of battery electric buses when operating on real roads was measured during 09:44–10:34 am period, over a range of 17 bus stops. The up and down trip directions of the buses are roughly comparable [3]. 2989 s time data totaling 10107 datasets were obtained during the obtained, and the transient output voltages and currents of the buses were taken. The transient power was calculated according to Eq. (19.1).

$$P(W) = U(V) \times I(A) \tag{19.1}$$

where, P is the power, W; U is the voltage, V; I is the current, A.

Figure 19.1 shows the voltage variation of battery electric buses during the up trip. 620 V is the stack voltage at departure. During the travel, the bus voltage is in the range of 500–650 V and changes with the speed.

Figure 19.2 shows the current variation of the bus during its up trip. When the current is positive, the bus gave off current. When the current was negative, the bus can recover current.

Table 19.1 Statistics of BEV performance in China

Type BEV	Max speed (km/h)	Driving range (km)		Charge time	Operation energy consumption (kWh/km)
		Uniform speed (40 km/h)	Load		
Battery electric car	>125	150–400	90–00	6–10 h	0.07–0.12
Battery electric motor coach	80–90	150–300	50–80	7–10 h	1.0–1.2
Super capacitor motor coach	>50	–	3–5	1.5–3 min	1.2–1.4

Table 19.2 Energy saved per 100 km when BEVs are used in urban operation management sector

	Oil cost per 100 km by original bus (yuan)	Electricity cost per 100 km by BEV (yuan)	Energy saved (yuan)
Bus	203.28	49.5	153.78
	44L × 4.62	150 kWh × 0.33	

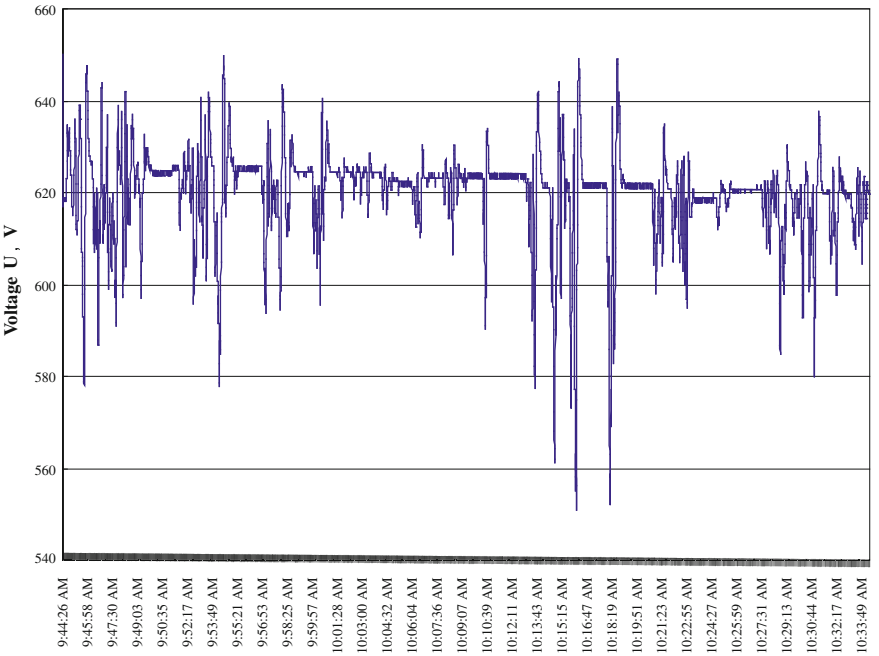


Fig. 19.1 Voltage variation chart

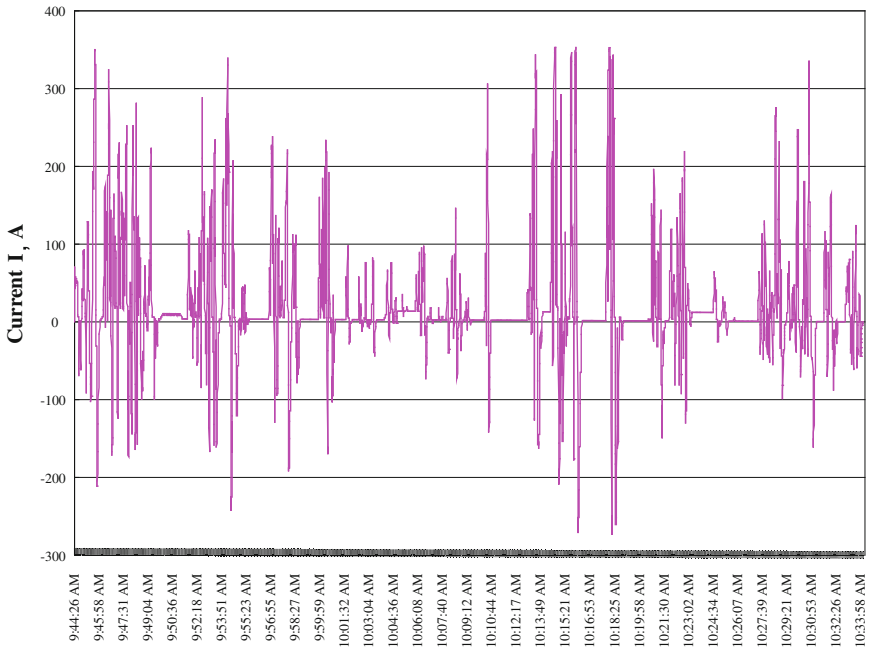


Fig. 19.2 Current variation chart

19.2.1 Transient Power Change of Vehicles

From Eq. (19.1), the transient power variation of the bus during the travel can be derived as shown in Fig. 19.3. When it travels, it has to stop frequently at bus stops and its load changes frequently too. Normally, when the speed of a vehicle increases, so does the power it has to output which is therefore positive [4]. When the vehicle travels at a slower speed, the power needed reduces and is therefore negative. In the figure, the power of the vehicle changes with the load and was in the range of −200–200 kW. For an energy vehicle, when the bus slows down, it can recover part of the energy, thus saving part of the energy consumption.

According to the transient power variation of vehicles, the cumulative power during the travel can be derived. Figure 19.4 shows the cumulative power variation of vehicles. As a whole, the bus power consumption is on the rise. In some of the periods, the cumulative power lowers as the vehicles recover the energy. According to measurement, the electricity consumed by the entire up trip is 9.72 kWh.

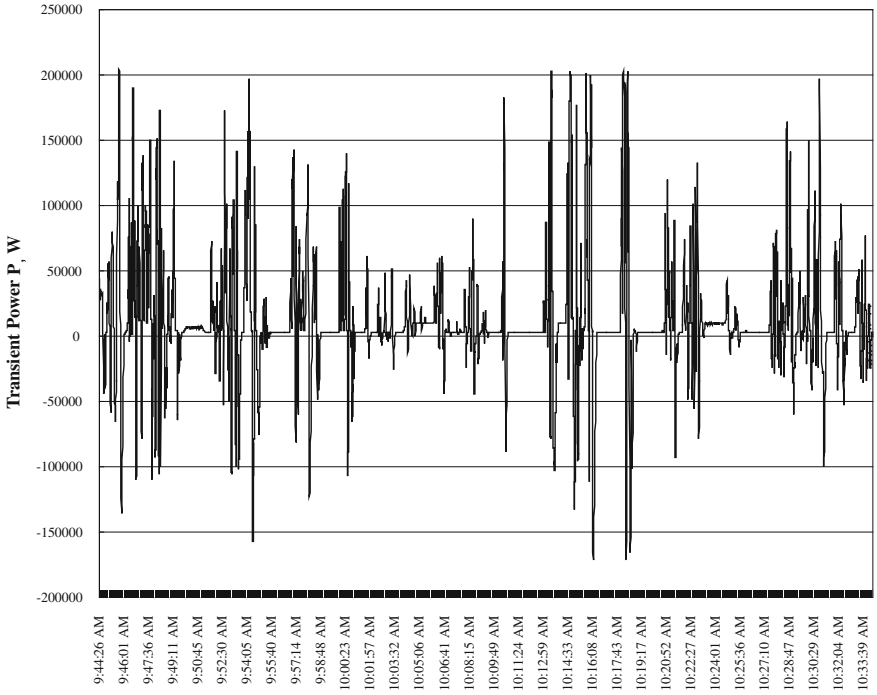


Fig. 19.3 Transient power variation chart

19.3 Comparison of Energy Efficiency

In terms of the energy consumption per 100 km, an air-conditioned (A/C) capacitor bus consumes 47 kWh electricity per 100 km than a conventional A/C bus. In addition, a capacitor bus also recovers an average of 17 kWh electricity by braking [5]. That is, it consumes 64 kWh less electricity than a conventional bus, meaning it is quite economical [6]. Table 19.2 Compares Analysis indicates that energy saved per 100 km when BEVs are used in urban operation management sector.

19.4 Effect of Emission Reduction

Table 19.3 compares the daily and annual average pollutant emission per unit line between conventional buses and energy vehicles as well as the emission reduction by energy vehicles as a substitute for buses used for the line [7]. The composite figures herein assume that the given line has all types of conventional buses which are distributed proportionally according to the percentage of the existing vehicles

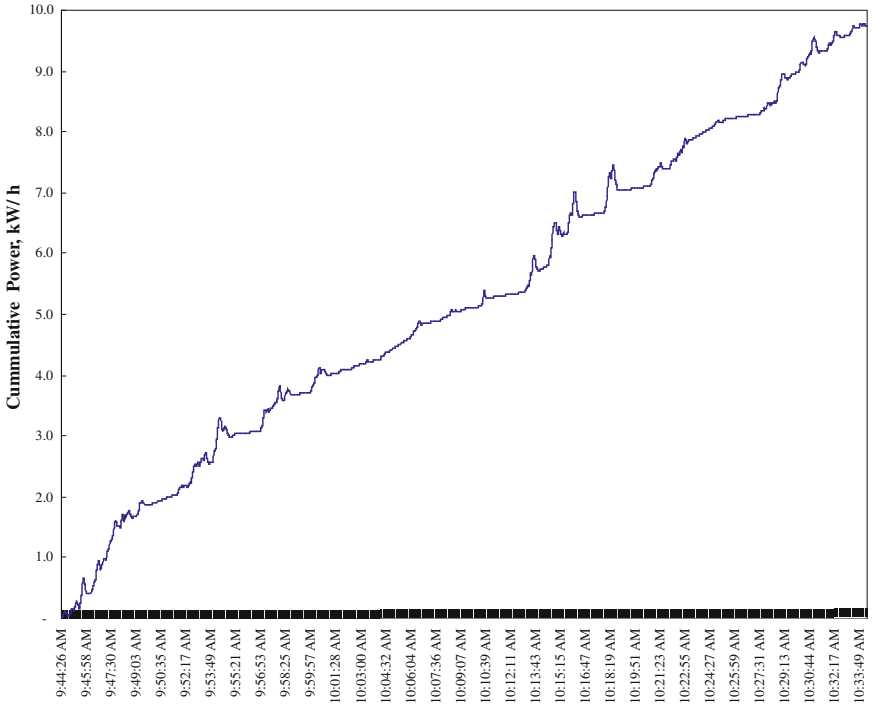


Fig. 19.4 Cumulative power variation chart

serving the city, i.e., 19.8 % for GI previous, 29.7 % for GI standard, 50 % for GII standard, and 0.5 % for GIII standard.

The simulation result well reflects the emission conditions of conventional buses when operating on real roads. Comparative simulation analysis as shown in Fig. 19.5.

Figure 19.6 compares the transient emission between battery electric buses and conventional buses. To make it simple, GII standard buses, which have the largest inventory among its counterparts, are used as our research target. The gray zones are the transient emission differences between GII standard buses and battery electric buses. The transient emission of battery electric buses is significantly lower than that of GII standard buses, implying significant environmental performance. When their power is negative, that is, when they recover energy, battery electric buses are reducing emission in another way. According to estimation, these buses can recover about 36 % of the energy at one trip, equaling to 36 % less remote emission.

Using battery electric vehicles as a substitute for public transition lines based on conventional vehicles has great potential for emission reduction. Figure 19.6 Analysis indicates that assuming a bus line consists of conventional buses of different technical levels in a certain city on a proportional basis, when super

Table 19.3 Emission per line energy vehicles versus conventional buses

Description	Type of vehicle	Daily avg emission (kg)				Annual avg emission (t)			
		CO	HC	NOx	PM	CO	HC	NOx	PM
Emission level	GI previous	33.2	8.3	61.0	2.7	12.1	3.0	22.3	1.0
	GI	22.1	5.6	40.7	1.8	8.1	2.0	14.8	0.7
	GII	18.1	5.0	31.6	0.7	6.6	1.8	11.5	0.2
	GIII	9.5	3.0	22.6	0.5	3.5	1.1	8.2	0.2
	Composite[a]	22.2	5.8	40.1	1.4	8.1	2.1	14.6	0.5
	Super capacitor	1.0	0.1	4.4	0.2	0.4	0.02	1.6	0.07
Emission reduction potential	GI previous	32.2	8.3	56.6	2.5	11.8	3.0	20.7	0.9
	GI	21.1	5.5	36.3	1.6	7.7	2.0	13.2	0.6
	GII	17.1	4.9	27.2	0.5	6.2	1.8	9.9	0.2
	GIII	8.5	2.9	18.2	0.3	3.1	1.1	6.6	0.1
	Composite[a]	21.2	5.7	35.7	1.2	7.8	2.1	13.0	0.4

[a] *Note* "Composite" assumes that a given line has all types of conventional buses which are distributed proportionally to the percentage of vehicles currently used in the city [8]

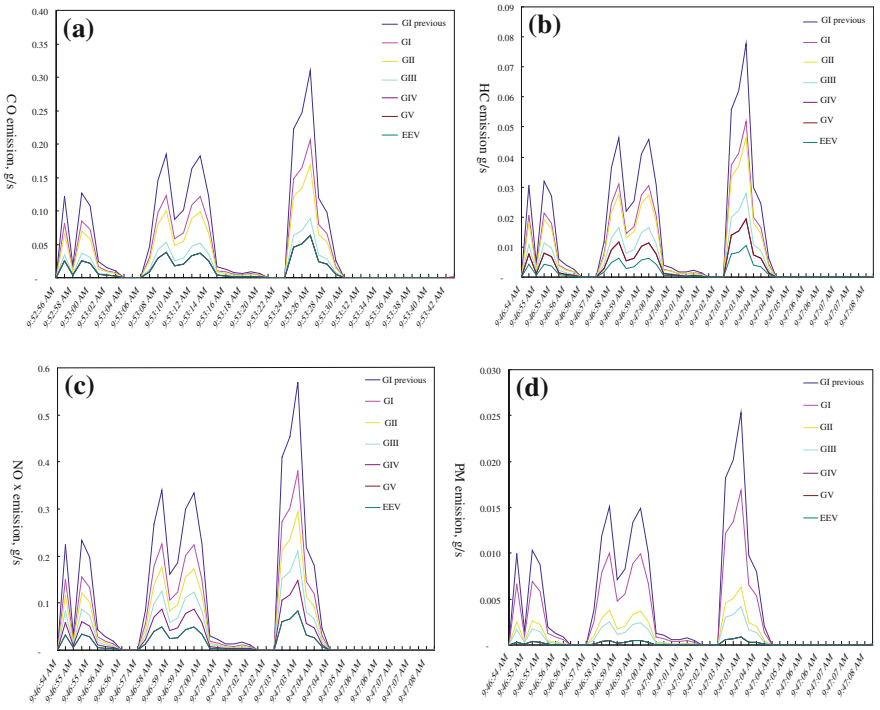


Fig. 19.5 Transient pollutant emission BEVs versus conventional buses **a** CO emission **b** HC emission **c** NOx emission **d** PM emission

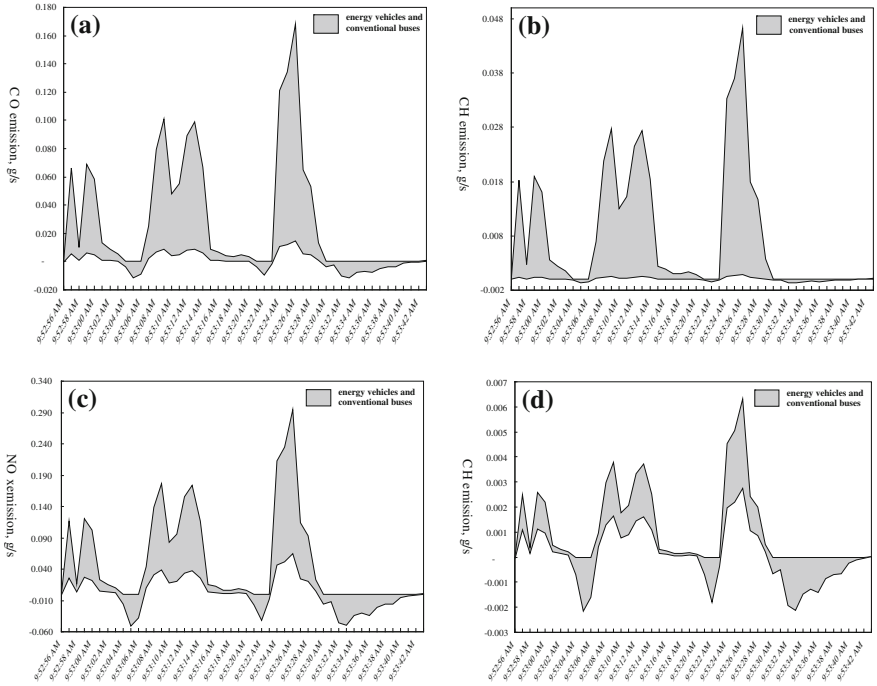


Fig. 19.6 Transient emission difference between energy vehicles and conventional buses **a** CO emission **b** HC emission **c** NOx emission **d** PM emission

capacitor vehicles are used as a substitute for this line, the CO, HC, NOx, and PM emissions per day will reduce by 21.1, 5.7, 35.7, and 1.2 kg, respectively, and the total CO, HC, NOx, and PM emission per year will reduce by 7.8, 2.1, 13.0, and 9.4t, respectively. These are equivalent to a reduction of 95, 99, 89, and 87 %, respectively [9]. The older the vehicles used in the substituted line, the higher the emission reduction effect.

19.5 Conclusions

After follow-up analysis of the immense emission reduction potential displayed by battery electric buses of different technical levels over many years, the superior overall performance of battery electric buses is identified, thereby offering scientific basis for decision making on the commercial operation of these vehicles under different physical conditions (vehicle conditions, load conditions, and road conditions).

References

1. Ozawa S, Tsukamoto K, Maeda T (1976a) Model experiments on devices to reduce Pressure wave, RTRI Research Report No. 990, Railway Technical Research Institute, Japan National Railways, Japan (in Japanese)
2. Ozawa S, Moritoh Y, Maeda T, Kinoshita, M (1976b) Investigation of pressure wave radiated from a tunnel exit, TRRI Research Report No. 1023, Railway Technical Research Institute, Japan National Railways, Japan (in Japanese)
3. Gawthorpe RG (1978) Aerodynamics in railway engineering, Part 3: Aerodynamics problems in the tunnels. Railway Eng Int 3(3):41–47
4. Hara T (1960) Aerodynamics force acting on a high speed train at tunnel entrance. Bull JSME 26:1581–1586
5. Paolo DS, Francesco F (2001) Bus Service optimization with fuel saving objective and various financial constraints. Part A: Policy and Practice. Source: Transp Res 35(2):157–176
6. Melke J (1988) Noise and vibration from underground railway lines: proposals for a prediction procedure. J Sound Vib 120(2):391–406
7. Zhi-qiang Rao HX (2010) Application research on road strain sensor about up-anchors based on circuit bridge compensation method. The international conference on electrical and control engineering June 26–28, 2010 in Wuhan. China. EI. P 5047–5050
8. Zhi-qiang Rao HX (2009) Study based on roof anchor shank strain of three gorges underground power plant under safe lifting load, International Symposium on Risk Control and Management of Design, Construction and Operation in Underground Engineering (IRCM2009) October City. P.R. Chnia. EI. P 476–P480
9. Paolo DS, Francesco F (1995) Bus service optimization and car pricing policies to save fuel in urban areas. Part A: Policy and Practice. Source: Transp Res 29(5):345

Chapter 20
Design and Verification of Hybrid Power Box of 100 % Low-Floor LRV

Li Ming, Li Minggao, Shi Junjie, Yang Guang and Shao Nan

Abstract The hybrid technology is one of the key technologies of the 100 % low-floor light rail vehicles. Based on technology needs of hybrid power box, a set of power battery and super capacitor adapted to different working conditions was designed, and energy control strategy in noncatenary areas was put forward, taking the dynamic performance requirements including the target vehicle speed, mileage, acceleration, gradient, and dynamic performance requirements of the target vehicle into consideration. Ventilation cooling, dynamic performance, and energy control strategy test results show that the hybrid system could satisfy the requirements of security and dynamic performance of the target vehicle, and also ensure higher energy use efficiency.

Keywords Hybrid power technology · Power battery · Super capacitor · Parameter matching

20.1 Preface

In recent years, hybrid power vehicles and relative technologies have made great breakthrough in many countries [1–3]. Hybrid power storage components, such as super capacitor and power battery, have been more and more widely used. Hybrid power technologies are applied earlier in the hybrid electric vehicle industry. In literature [4], a control system of hybrid power electric vehicle was designed which contains super capacitor and storage battery. Experimental results showed that the hybrid power vehicles were able to combine the advantages of batteries and super capacitors. It could meet the performance requirements for starting and

L. Ming (✉) · L. Minggao · S. Junjie · Y. Guang · S. Nan
P&T Research Center, Tangshan Railway Vehicle Co, 3# Changqian Rd, Fengrun District, Tangshan, China
e-mail: sjc-liming@tangche.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_20,
© Springer-Verlag Berlin Heidelberg 2014

acceleration, and improve the energy recovery efficiency when braking. Literature [5] proposed an energy management strategy of four-step method for hybrid power vehicle, and established an energy management strategy which was based on fuzzy rules and genetic optimization algorithm. Nowadays, in the rail transportation industry, hybrid power railway vehicles are also developing rapidly, the power of which is provided by the traditional power supply system, super capacitors, and battery-powered. Countries including Japan, Germany, France, and Spain have many research experiences in hybrid power railway vehicle technologies, which contain the traditional power systems, fuel cells, power batteries, super capacitors, and others, such as the MITRAC item of Bombardier which adopts super capacitor energy-saving devices, Sitras-HES item of Siemens and ACR item of CAF Company which adopt the power supply mode including both super capacitor and power battery.

20.2 Hybrid Power Schemes for 100 % LRV Vehicles

For different lines, this paper proposed different hybrid power schemes as shown in Table 20.1. As is limited by the current technology level of power battery, only the second and third scheme could be implemented.

The hybrid power system and light rail vehicle that this paper developed could meet operational requirements of the second and third scheme. The hybrid power box was fully integrated with advantages of both super capacitor and power battery, which could achieve a larger starting acceleration and a longer mileage. This paper also sets the functions for the hybrid power system: power batteries could charge super capacitor and super capacitor could absorb energy when the vehicle braking, to ensure that the vehicle could start and stop frequently and energy used efficiently.

In this paper, the matching calculation and scheme design were processed for hybrid power vehicle system of 100 % low-floor light rail vehicle. The energy control strategy was developed, and the feasibility of the system was verified through simulation and experimental tests.

20.3 Design of Hybrid Power Box

In this paper, the hybrid power system could realize the function of cooperated power supply by catenary and power storage devices (including super capacitors and power batteries). When it is in the catenary areas, the catenary will supply power to traction inverter, and charge up for the super capacitors and power batteries. When it is in the noncatenary areas or the catenary has no power, the super capacitors and power batteries will respectively provide power to the traction

Table 20.1 Comparison of several hybrid power schemes

The noncatenary areas	Schemes of storage parts	Feasibility
First scheme: the whole line is noncatenary	Super capacitor + power battery	Low
Second scheme: the lines between two stops is noncatenary	Super capacitor + power battery	High
Third scheme: grade crossings is noncatenary	Super capacitor	High

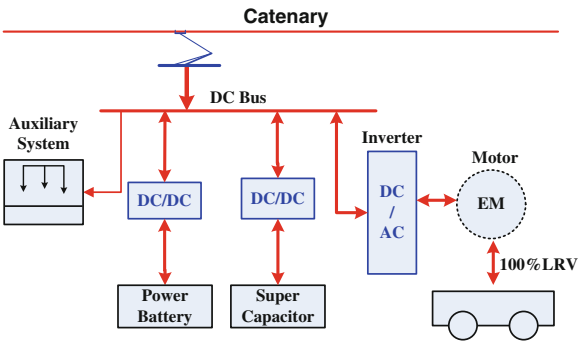


Fig. 20.1 Block diagram of the main circuit

inverter to drive the traction motor through the DC/DC converter. Block diagram of the main circuit was shown in Fig. 20.1.

In order to achieve vehicle dynamic performances and economic optimization of the vehicle, this papers take mileage, maximum speed, maximum gradient, and acceleration performance as operating performance indicators in the noncatenary areas, which could help to determine the power and energy needs of the hybrid power system in different conditions, and then rationally determine the parameters of super capacitors and power batteries and also power supply control strategies.

20.3.1 Parameter Matching Calculation of Hybrid Power System

The vehicle has three motor trains and one trailer train, each motor train configures four motors, and the rated power is 50 kW.

(1) Characteristic calculation of traction system
Traction force F_t (kN) [6, 7] is:

$$F_t = \frac{mgw_0}{1000} + F_g + F_f + \frac{G_d}{3.6} \cdot \frac{dv}{dt} \tag{20.1}$$

In the formula, w_0—the unit basic resistance, N/kN;

F_g—Running resistance of the vehicle on slope, $F_g = mg \sin\alpha$, kN;
F_f—Additional resistance caused by air, kN;
G_d—Total vehicle mass including rotation mass, t;
v—Vehicle speed, km/h.
Total output power of traction motor [8, 9] is

$$P = \frac{F_t v}{3.6\eta} \tag{20.2}$$

In the formula, η—efficiency of the transmission system.

(2) Operation performance indicators of the vehicle
This paper take gradient $\sin\alpha$ similarly to i, so acceleration a of the vehicle could be expressed as:

$$a = \frac{dv}{3.6dt} = \frac{1}{G_d}\left(F_t - \frac{mgw_0}{1000} - mg \cdot i - F_f\right) \tag{20.3}$$

According to the external characteristics of the drive motor, the motor output power corresponds to each speed in constant torque region and motor output power in constant power region could be determined. And then parameters including the motor rated power, peak power, rated torque, and rated rotate speed could be determined.

20.3.2 Matching Calculation of Energy Storage Devices

Power batteries have high specific energy and low specific power, while super capacitors have low specific energy and high specific power. So in this paper, power batteries were considered as the power for continuos operation on straight track, and the super capacitors were considered as the power for acceleration, running uphill, and energy recovering when running downhill and braking.

The principles to confirm the number of batteries and super capacitors are shown in Table 20.2.

The IFP33 lithium battery of Suzhou Xingheng Power Co. and the MAXWELL BMOD0165 super capacitor module were adopted in this paper. The nominal voltage and capacity of the battery cell is 3.2 V and 40 Ah, respectively. The cell's largest energy and maximum power is 0.0683 kWh and 0.56 kW, respectively. The nominal voltage and capacity of the super capacitor cell is 48 V and 165 F, respectively. The cell's largest energy and maximum power is respectively is 0.0304 kWh and 45.87 kW.

(1) Requirement on the number of power battery for running mileage
Figure 20.2 shows the total energy required for different mileages at different speeds, where the figures on the curve represent the minimum numbers of

Table 20.2 The principles to confirm the number of batteries and super capacitors

Constraints	Power batteries	Super capacitor
Power constraints	The maximum output power and minimum number of batteries determined by the maximum speed	Peak power and minimum number of super capacitors determined by the acceleration of vehicle
Energy confinement	The total output energy of power battery and minimum number of batteries determined by the maximum mileage	The required energy during acceleration and minimum number of super capacitors determined by the acceleration of vehicle
Series voltage constraints	The number of series power batteries determined by the cell voltage of the power battery and the total voltage	The number of series super capacitors determined by the cell voltage of the super capacitor and the total voltage

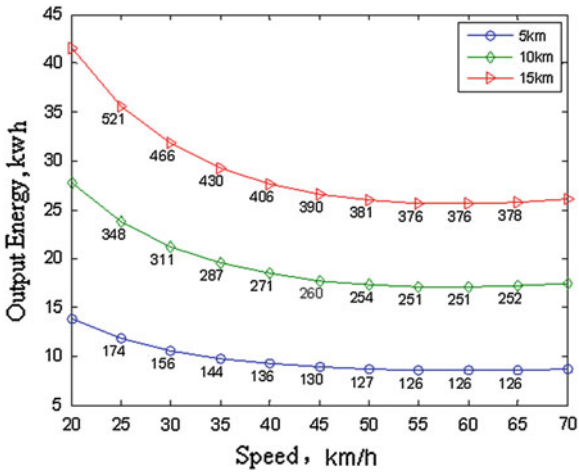


Fig. 20.2 The relationship between speed and total energy consumption while continuous operation

power batteries relative to the energy needed in different speed and mileage. Depending on the requirements of output voltage, the configuration of the battery pack was selected as 150 batteries in series, the operating voltage is in the range of 340–480 V.

(2) Requirement for super capacitors from the starting acceleration
Figure 20.3 shows the total energy needed when the vehicle accelerates to a specified speed in different accelerations, where the figures on the curve represent the minimum number of super capacitors when the vehicle accelerates to the desired speeds. It shows that when the power demand increases, the number of super capacitors increases.

(3) Requirement for super capacitors from maximum slope
Assumed that the vehicle operates on the slope, entirely driven by the super capacitors. It is shown in Fig. 20.4 that the traction power and the quantity of

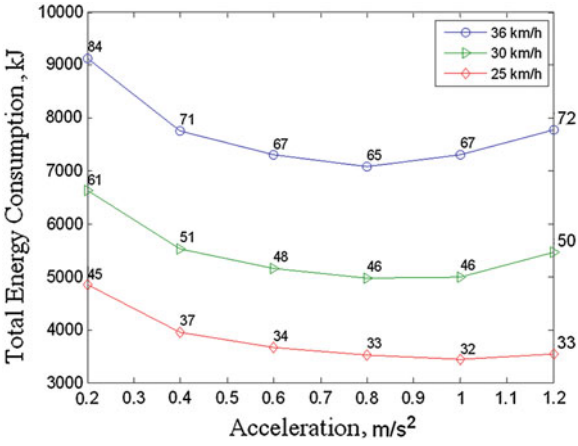


Fig. 20.3 The corresponding relationship between acceleration and total energy consumption

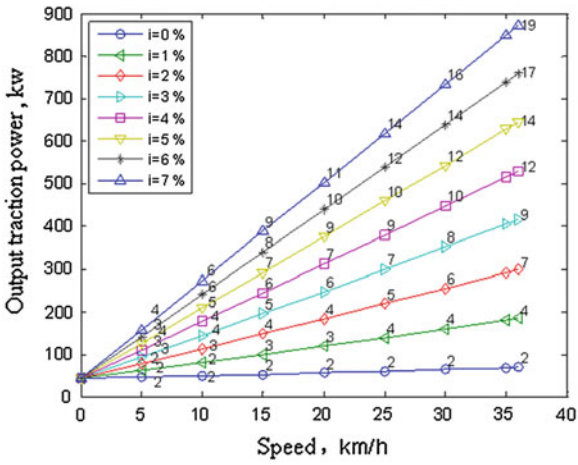


Fig. 20.4 Different slopes and speeds corresponding to power and quantity of super capacitor

the super capacitor are corresponding to the different gradients and different speeds. Yet, the gradients have great influence on the traction power, if the vehicle was equipped with 40 super capacitors that makes up of 10 series and 4 parallel super capacitors, it could meet the power requirement for 50 ‰ gradient.

After parameter matching optimization calculation, it is determined that two power boxes are needed on the vehicle, the configuration adopts 10 series and 4 parallel super capacitors, 150 series and 2 parallel power batteries totally, to meet the requirement that the vehicle could accelerate from static to 30 km/h in accelerations during 0.2 and 1.2 m/s^2, to meet the power demand for slope operation and the mileage, and also achieve high energy efficiency. On this basis, the control strategy of hybrid power system was determined: when it is in the catenary areas, the catenary will supply power to power

boxes; When it is in the noncatenary areas or the catenary has no power, the power boxes will provide power to the vehicle. The super capacitors work on high-power conditions, and the power batteries work on lower-power conditions both when charging up and discharging. Different operating conditions, different types of energy control strategy would be adopted when energy and power demand are different, for a more rational allocation and efficient recovery of energy.

20.3.3 Power Performance Evaluation on Hybrid Power Box

(1) Energy calculation when charging the super capacitor with lithium batteries
When effective energy of the super capacitor (1.045 kWh) is exhausted, it needs energy for roughly 1.088 kWh, taking into account the efficiency of charging up the super capacitor with battery box.
Since the power battery charges the super capacitor with the constant current, and it could not continue too long in a large current, so the charging current is selected as 120 A (3C). According to the constant current discharge characteristic curve of the Xingheng power battery, the battery voltage drops very slowly in the initial stage. And because of the super capacitor has a fast charging speed, it is assumed that the battery voltage drop is 0 in this paper. For example, if the battery voltage is 470 V, the charging time is 86 s.

(2) Estimation for the start and stop times
In this paper, the most simple power supply control strategy was selected that lithium batteries and super capacitors supply power singly. Then the start and stop performances were analyzed when the train accelerating from static condition on straight track. The total capacity of super capacitor box and the battery box is respectively 1.045 and 13.008 kWh (3C discharging).
The first acceleration: firstly, the super capacitor supplies power. When the super capacitor's effective energy is exhausted, the train could reach the speed of 30 km/h, consuming 18 s. Then the power battery was switched to supplying power. It totally cost energy for 2.1 kWh, where the super capacitor supplies 1.045 kWh, power battery supplies 1.055 kWh.
The secondly acceleration: as the super capacitor's energy is basically exhausted, the energy for acceleration process is supplied all by the power battery, the speed can reach 17 km/h. This process takes 60 s, consuming power 1.483 kWh of power battery.
Without considering the other processes (nonacceleration) of energy consumption, the start and stop times by power battery is about 8, the energy calculative formula is $(13.008-1.055)/1.483 = 8.06$.
In summary, the configuration of the vehicle hybrid power box is available to complete one time to speed up from static to 30 km/h, and eight times to speed up to 17 km/h on the straight track.

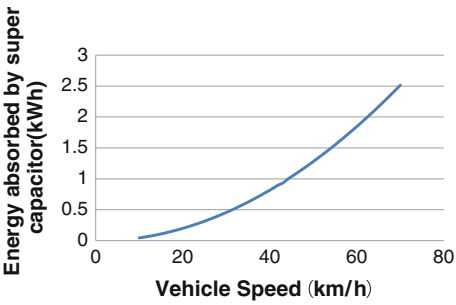


Fig. 20.5 Energy estimation that super capacitor absorbs in different speeds when braking

(3) Brake energy recovery
Considering the efficiency of the power to wheel circumference, traction motor, DC/DC converters and super capacitor, the paper takes the braking energy recovery efficiency of super capacitor is 0.6 (conservative value).
In Fig. 20.5, the figures are estimation values for recovery energy of the super capacitor in different speeds when braking. Because the capacitor could absorb energy for 1.045 kWh, when the vehicle speed exceeds 46 km/h, the braking energy could not be completely absorbed by the super capacitor. At this time, the energy absorbing solution by both super capacitor and battery should be considered.

Based on the above calculations, the hybrid power system could realize the following performance:

(1) In temporary parking cases, the lithium batteries could charge up super capacitor, providing higher power for starting and accelerating;
(2) In the cases that the vehicle starts and stops frequently between the two stations, the vehicle could start nine times frequently, and run 2 km distance in the noncatenary areas (the most extreme conditions) without recharging;
(3) Taking into account that most of the station intervals are less than 1 km, the largest station interval does not exceed 2 km, the driver could stop at any time to add energy to ensure the smooth operation of the vehicle;
(4) The vehicle could achieve braking energy recovery in less than 45 km/h speed.

20.3.4 Structure Optimization of the Power Box

Since hybrid power box is placed on the top of vehicle, it has an relatively high requirement for both the ventilation cooling and the degree of protection. DC/DC converters, battery box, and super capacitor box were designed with different approaches. They are all separate cabinets, avoiding the electrical and mechanical

Table 20.3 The improved design that super capacitor box and power battery box

	Power battery box	Super capacitor box	DC/DC converter box
Cooling methods	Forced air cooling, ventilation cooling begins when the temperature exceeds 40 °C	Natural cooling	Forced air cooling, ventilation cooling begins when the temperature exceeds 50 °C
Box structures	Extend spacing between battery cells, enhancing performance of ventilation cooling; add air filter at inlet and arcuate shroud at outlet, ensuring the protection grade	Expand spacing between the super capacitor modules, and isolated from outside, enhance the protection grade greatly	Optimize the flow channel

influence on each other. For ease of maintenance and marketing, the boxes were designed based on modular design principles, as is shown in Table 20.3.

In addition, thermal management systems and power management system of the DC/DC converter box and the battery box were optimized. To prevent rain, dust, and etc. into the boxes, a filter was installed at the inlet, and an arcuate shroud was installed at the outlet.

After CFD calculations on the re-arrangement of the battery pack (Fig. 20.6), the flow resistance is reduced to 280 Pa from the original 410 Pa, air flow of the cooling fan increased by 60 % or more, simulation and test results of the dissipate heat quantity Q reached above 1 kW. The cooling effect advanced obviously.

20.4 Performance Test Verification of Hybrid Power Box

This paper built a performance test platform of hybrid power system, to test and validate dynamic performance and energy control strategy of the hybrid power system, as is shown in Fig. 20.7.

20.4.1 Mileage Test

This test simulates the vehicle's cruising ability when vehicle drives into areas without electricity, at different speeds, until the SOC of the power box is reduced to 50 %. Mileage test data are shown in Fig. 20.8.

Figure 20.8 shows that when the vehicle is running at higher speed in the straight track, the mileage would be longer, and better economy.

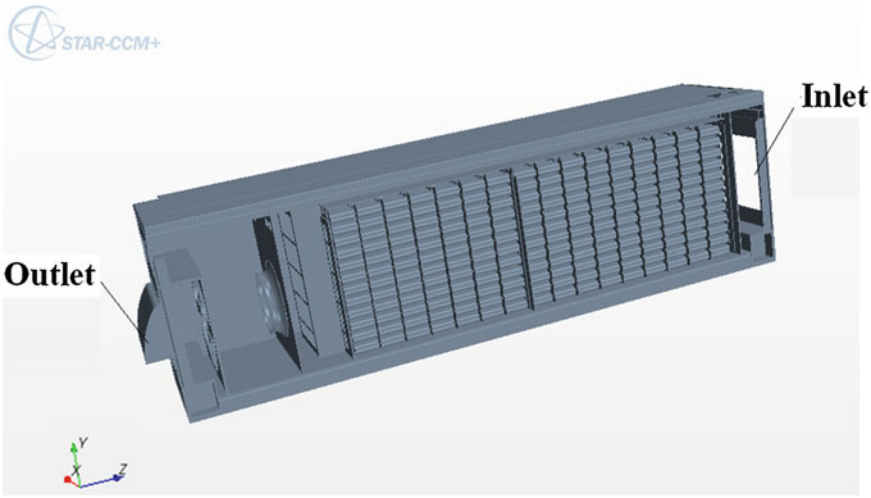


Fig. 20.6 Simulation model

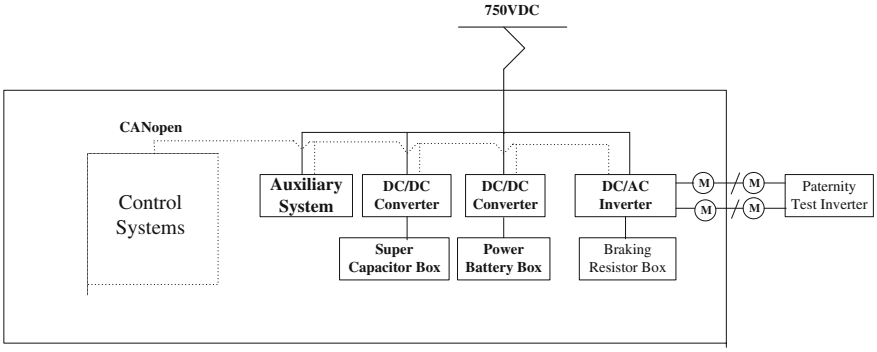


Fig. 20.7 The test platform of hybrid power system

20.4.2 The Ability Test of Running Uphill Slope

The test data of the ability to run uphill slope when the vehicle drives at different speeds and gradients in noncatenary area was shown in Table 20.4 (Because super capacitor's specific energy is low and could not last too long, and the power battery could provide cruising ability, so it is the basic requirements to meet the long slope operating conditions of the battery box).

Table 20.4 shows that at speed of 5 km/h, the box could meet the demand that the vehicle runs continuously on 40 ‰ gradient slope; at speed of 10 km/h, the vehicle could run continuously on 20 ‰ gradient slope; at speed of 20 km/h, the vehicle could run continuously on 10 ‰ gradient slope; at speed of 30 km/h or 36 km/h, the vehicle could run continuously on 5 ‰ gradient slope.

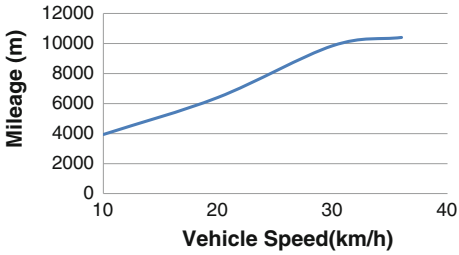


Fig. 20.8 Mileage test data. *Note* Auxiliary power of this test is 16.8 kW, the same below

Table 20.4 The test data of running uphill slope

Vehicle speed (km/h)	Gradient (‰)	Super capacitor discharging time (s)	Battery SOC consumption in 1 min (%)
5	40	43	4.3
10	20	51	5.2
20	10	49	4.4
30	5	49	4.2
36	5	45	4.8

Note It will test the power supply capacity of the hybrid power box when super capacitor's available capacity is exhausted, and then the battery discharge for 1 min

20.4.3 The Ability of Driving Uphill Slope in Areas Without Electricity

This test simulates the acceleration performance of the vehicle on 5 to 30 ‰ slope in noncatenary area, the test data are shown in Tables 20.5 and 20.6.

Table 20.6 shows that: Because instantaneous discharge current of super capacitors is high, it could meet the acceleration performance on relatively high gradient slopes (30 ‰). While the batteries' instantaneous discharge current is low, it could only meet energy requirements of the relatively low slope (5 ‰). So this paper has verified the control strategy is correct that the super capacitors support the slope running, and the batteries supply cruising ability on straight road.

20.4.4 Strategy Validation Test of Braking Recovery Energy in the Noncatenary Areas

When the vehicle is braking, little energy could be absorbed by catenary or batteries in downscale bit, while in the upscale bit, the energy recovery could be rapidly absorbed by super capacitors, which could be used as power supply for acceleration. Based on the vehicle speed and the current limit setting, the energy absorption time could be effectively regulated, reducing shock on the system.

Table 20.5 Test data of super capacitors start on slope in noncatenary area

Acceleration (m/s^2)	Gradient (‰)	Super Capacitor Discharging Time (s)
0.2	5	48
0.2	10	40
0.2	20	31
0.2	30	24
0.4	5	26

Table 20.6 Test data of power batteries start in slope in noncatenary area

Acceleration (m/s^2)	Gradient (‰)	Battery discharging time
0.2	1	11 min 59 s
0.2	2	11 min 1 s
0.2	5	9 min 38 s

Combined braking energy absorption and stops time, this paper takes 200 $\sim$ 300 A as the charging current setting range of super capacitor, 40 $\sim$ 60 A as the charging current setting range of battery.

In the noncatenary areas, the test data and results when power batteries and super capacitors absorbed the energy together show that super capacitors could be charged to 75 % in a brake condition (the vehicle speed is 30 km/h, the charging current is 200 A), which needs 25 s. Energy absorption efficiency is 70 %. This verified the rationality of the aforementioned estimation.

20.5 Conclusions

This paper conducted design of parameter matching and control strategies for the hybrid power box, through the test validation, the following conclusions can be got.

(1) When running on the straight railway, the power and speed of the vehicle, the energy of power battery required were approximately proportional relationship to the mileage. The maximum mileage could be up to 10 km.
(2) The power requirement increases along with the acceleration and the gradient, meanwhile the required number of super capacitors increases.
(3) When running on the slope, both the power and energy requirements are larger. So the vehicle speed or power control strategy is needed, by using power batteries and super capacitors discharge simultaneously, to make up the lack of mutual characteristics for each other.
(4) Combined with theoretical analysis, simulation and experimental verification, the hybrid power box was optimized based on dynamic performance and

temperature test results. The new box has better ventilation and cooling capacity, higher degree of protection, with more security and reliability.

(5) Experimental data show that the hybrid power system equiped on 100 % low-floor light rail vehicles could better achieve dynamic performance and energy efficiency when operating in the noncatenary areas. The box reduces energy consumption, and provides a better guarantee of green energy for the vehicles.

References

1. Ehsani M, Gao YM, Emadi A (2004) Modern electric, hybrid electric, and fuel cell vehicles fundamentals, theory, and design, 2nd edn. CRC Press, New York
2. Chen W, Qian Q, Li Q (2009) Investigation status and development trend of hybrid power train based on fuel cell. J Southwest Jiaotong Univ 44(1):1–6 (in Chinese)
3. Osawa M (2005) The NE Train (New Energy Train) the world's first hybrid railcar. Converter Technol Electric Traction 35(4):16–23
4. Cao B, Cao J, Li J et al (2008) Ultracapacitor with applications to electric vehicle. J Xi'an Jiaotong Univ 42(11):1317–1322 (in Chinese)
5. Zhu Y, Tian G, Chen Q et al (2004) Four-step method to design the energy management strategy for hybrid vehicles. Chin J Mech Eng 40(8):127–133 (in Chinese)
6. Xie W (2010) Electric traction and control. China Railway Press, Beijing (in Chinese)
7. Chen Y, Zhang J, Chen X et al (2012) Dynamics analysis and power allocation strategy research for a new hybrid electric urban rail vehicle. In: The 2nd international conference on mechatronics and applied mechanics, 2012, Hongkong
8. Zheng W (2010) Research on parameter matching method and control strategy of hybrid power vehicle. Harbin Institute of Technology Thesis, p 4 (in Chinese)
9. Zhu Y (2005) Motor and electric drive. Beijing University of Aeronautics and Astronautics Press, Beijing (in Chinese)

Chapter 21
Measurement and Simulation of the Electromagnetic Transients of Lifting Pantograph for an Electric Multiple Units Train

Dan Shi, Mingli Wu, Honghe Zhang, Teng Li, Hui Wang and Kejian Song

Abstract The electric sparks and arcs occurred at the contact of a pantograph and an overhead contact wire is the emission source of electromagnetic interference in electric railways. It has been noticed by the train operators that CRH_2 EMU train has denser electric sparks more frequently than other EMUs when lifting its pantograph. The electromagnetic transients in the high voltage circuit may produce overvoltage and disturb the onboard electronic equipment, as well as emit high frequency electromagnetic interference. Some measurements of the charging transients when CRH_2 EMU lift the pantograph were conducted in the depot. Using PSCAD/EMTDC, a simulation circuit that includes external power source, the pantograph and the high voltage cable on the car roof is established. The simulations to analyze that how the circuit element parameters can affect the electromagnetic transients are given. The calculated results are compared with the measured data to validate the simulation model. One solution is put forward to improve CRH_2 charging transient performance.

Keywords Electromagnetic transients · Electric railway · Overvoltage · High frequency oscillation

21.1 Introduction

In recent years, high-speed railways have been developed rapidly in China. Some new technologies are used to meet the needs of high-speed services, which also bring some new technical problems. One case in point is that the electric sparks and arcs occurred at the contact instant of a pantograph to the overhead contact

D. Shi · M. Wu (✉) · H. Zhang · T. Li · H. Wang · K. Song
School of Electrical Engineering, Beijing Jiaotong Universiry, No.3 Shang Yuan Cun, Beijing 100044, People's Republic of China
e-mail: mlwu@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_21,
© Springer-Verlag Berlin Heidelberg 2014

wire, which are emission sources of electromagnetic interference in electric railways [1–4]. The CRH_2 EMU train has more severe and more frequent electric sparks than other EMUs when lifting its pantograph to achieve the contact wire due to its distinct high voltage circuitry where there is a relatively long section of high voltage coaxial cable in front of the circuit breaker. It is the charging process of the cable that leads to the electric sparks of lifting pantograph.

In this paper, a simulation circuit based on PSCAD/EMTDC is established. How the circuit element parameters can affect the electromagnetic transients including the peak voltage, the oscillation frequency, and the transient duration are analyzed in virtue of simulation. The calculated results are compared with the measured data to validate the simulation model.

21.2 Simple Circuit Analyses of the Electromagnetic Transients During Lifting Pantograph

The circuit breaker is open when the CRH_2 EMU train lifts its pantograph. Then the high voltage circuitry of the EMU can be equivalent to a simple *RLC* series circuit as Fig. 21.1. *R* and *L* represent the source resistance and inductance of the contact wire. The capacitor *C* denotes the high voltage coaxial cable. The switch *S* describes the pantograph, which changes to close when the sliding strip of the pantograph reaches the contact wire. Assuming that the initial energy stored in the capacitor and the inductor is zero, when the switch closes the differential equation of the capacitor voltage u_c is

$$LC\frac{d^2u_c}{dt^2} + RC\frac{du_c}{dt} + u_c = e(t) \tag{21.1}$$

The CRH_2 EMU has a relatively long section of high voltage coaxial cable in front of the circuit breaker. Thus, the distributive capacitance of the coaxial cable has significant effect on the charging oscillation process. In order to facilitate the analysis, the resistance in the circuit can be ignored. Therefore, a simple *LC* circuit can replace the original one. Consider the most adverse condition that the switch closes at the peak voltage. Then the capacitor voltage and the oscillation frequency can be calculated by

$$u_c = U_m(1 - \cos\omega_0 t) \tag{21.2}$$

$$\omega_0 = \frac{1}{\sqrt{LC}} \tag{21.3}$$

respectively.

u_c reaches the maximum at $t = \pi/\omega_0$. Actually, the oscillation process will be damped because the circuit resistance can cause energy loss. The attenuation coefficient is given by

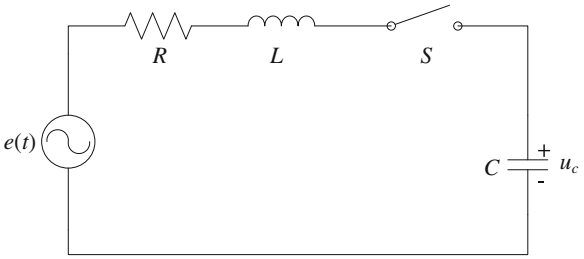


Fig. 21.1 Simple equivalent circuit

$$\alpha = R/2L \tag{21.4}$$

Obviously, greater resistance or smaller inductance can lead to a larger attenuation coefficient so that the transient component of the capacitor voltage decay faster and the transient duration become shorter.

21.3 Field Measured

In April 2013, some measurements of the charging transients when lifting the pantograph were conducted in Qingdaosifang Rolling Stock Factory of China. The CRH_2 EMU has four motor cars and four trailer cars. There is a pantograph on both the fourth and sixth car roof. Three kinds of high voltage cable are used as shown in Fig. 21.2. How do the contact wire voltage and cable charging current change during lifting the pantograph are measured [5].

The voltage signal was collected from the contact wire directly by a voltage divider placed on the top of the tested EMU. The current signals were measured by current clamps. The field measuring pictures are shown in Fig. 21.3 [6].

During the test, the pantograph mounted on the sixth car was lifted and the following electrical quantities were measured:

V_n, Contact wire voltage; I_L, Total cable current.

One of the test results has been shown in Fig. 21.4.

Figure 21.4 indicates that the pantograph contacts the overhead contact wire at around the peak value of 38.9 kV, 50 Hz supply voltage. The transient voltage is the sum of a high frequency component and a power frequency component. The decayed high frequency component represents the capacitive charging process, due to which, the voltage peak value reaches as high as near 65 kV. The oscillatory frequency is about 66.67 kHz and the oscillation duration is approximate 0.2 ms.

It is clear that the overvoltage caused by high frequency oscillation of the CRH_2 EMU is serious. This electromagnetic transient can cause interference to the electronic equipment onboard and will lead to electromagnetic emission to surroundings [7, 8].

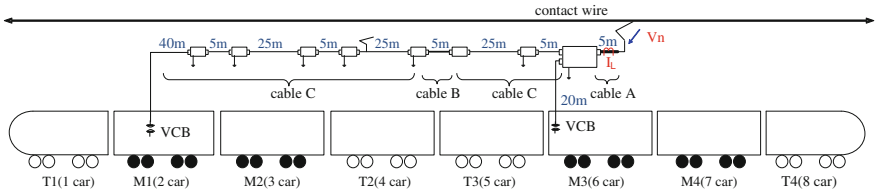


Fig. 21.2 Wire-connection diagram of high voltage cables

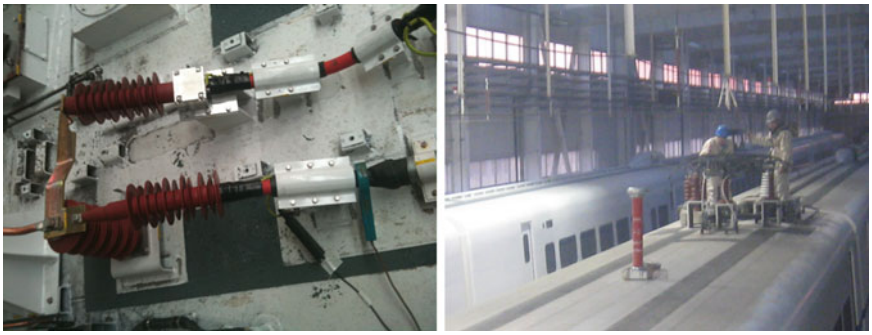


Fig. 21.3 Field measuring pictures

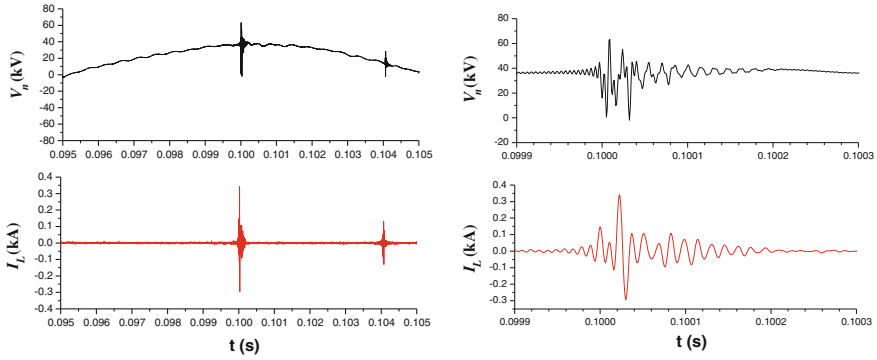


Fig. 21.4 Measured electrical quantities and partially enlarged view

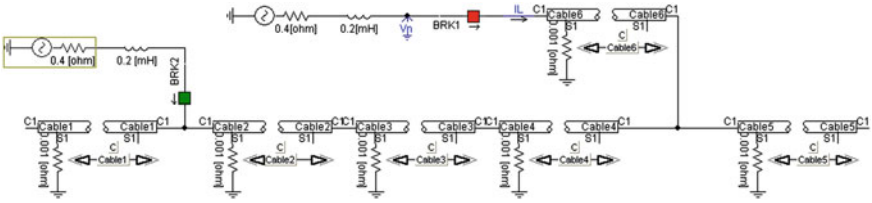


Fig. 21.5 Simulation model of the high voltage circuitry

21.4 Lifting Pantograph Simulation

21.4.1 PSCAD/EMTDC Simulation

The PSCAD/EMTDC model is shown in Fig. 21.5.

In this model, the equivalent inductance and resistance of the power supply are set as 0.2 mH and 0.4 Ω, respectively. Breaker 1 denotes the pantograph on the sixth car and breaker two denotes the pantograph on the fourth car. The high voltage cable configuration is shown in Fig. 21.6. The cable has a four-layer structure including a conductor, an insulating layer, a metal shield layer, and an insulating sheath. The sizes of each layer for the three types of cable are marked in Fig. 21.6.

The simulation results under the condition that breaker 1 close at the peak voltage are shown in Figs. 21.7 and 21.8.

It can be seen from Figs. 21.7 and 21.8 that how do the contact wire voltage and cable currents oscillate during lifting the pantograph. Some characteristic values from the simulation are shown in Table 21.1.

The oscillation frequency, the overvoltage and the transient duration obtained by simulation are substantially identical with the measured values. However, the current value is smaller than the measured data. That may be because that the lifting pantograph process is in fact not an ideal closing switch process. The air will breakdown when the pantograph is approaching to the wire but before a contact occur. The electric sparks or arcs can cause severe interference to the measuring system.

21.4.2 Simulation Analysis

To investigate the effect of circuit parameters on the transients, we change the equivalent inductance and resistance values of the power supply while keeping the cable configuration unchanged. The resistance value is taken as 0.2, 0.4, 0.6, 0.8, 1 Ω and the inductance as 0.1, 0.2, 0.3, 0.4, 0.5 mH, respectively. The simulation results are shown in Figs. 21.9, 21.10 and 21.11. These results indicate that the

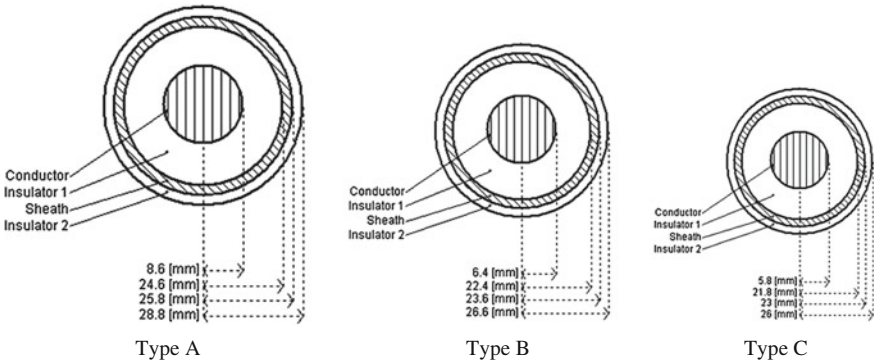


Fig. 21.6 Cable parameters

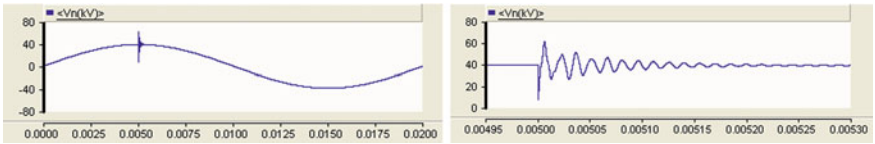


Fig. 21.7 Contact wire voltage V_n

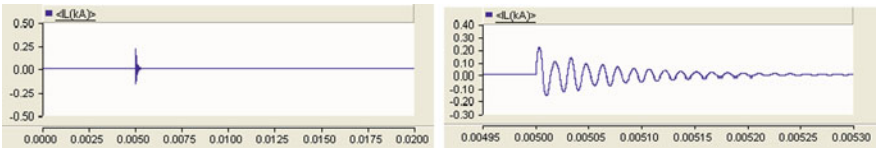


Fig. 21.8 Cable current I_L

Table 21.1 Simulation values and measured data

Characteristic values	Simulation values	Measured data
Oscillation frequency (kHz)	67.79	66.67
Peak voltage (kV)	61.35	63.27
Transient time (ms)	0.25	0.21
Peak I_L (kA)	0.21	0.35

oscillation frequency and the peak value voltage are mainly affected by the inductance, but the transient duration is influenced by both the inductance and resistance. This is consistent with the theoretic analysis.

Based on the simulation, the relationship of the voltage value with the angle of the contact voltage when lifting the pantograph can be obtained. The measured results are consistent with the simulated curve as shown in Fig. 21.12.

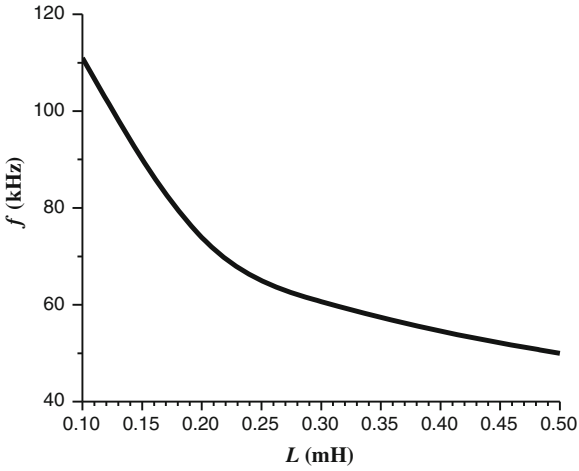


Fig. 21.9 Oscillation frequency

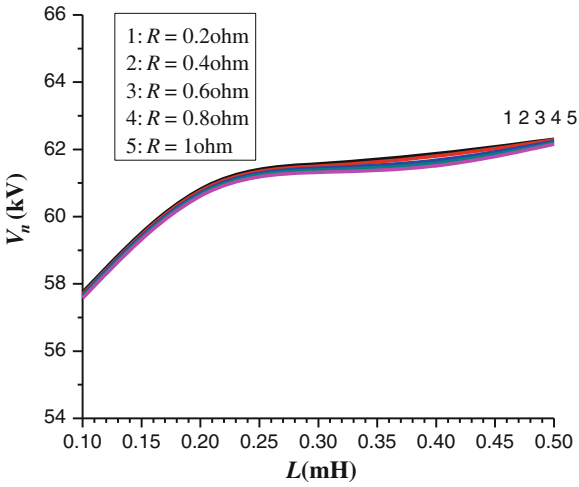


Fig. 21.10 Peak value of voltage

21.5 Possible Solutions to Suppress the Transients

As is shown in Fig. 21.13, two RC absorption circuits can be used to reduce the overvoltage. In this case, the resistance is 50 Ω and the capacitance is 0.2 uF. The simulation results are shown in Fig. 21.14 and concrete value is given in Table 21.2.

Table 21.2 indicates that this solution can restrain the overvoltage. In addition, the transient duration is reduced. But the current peak value increased. This approach is applicable since the *RC* filter can also improve the high frequency performance of the EMU.

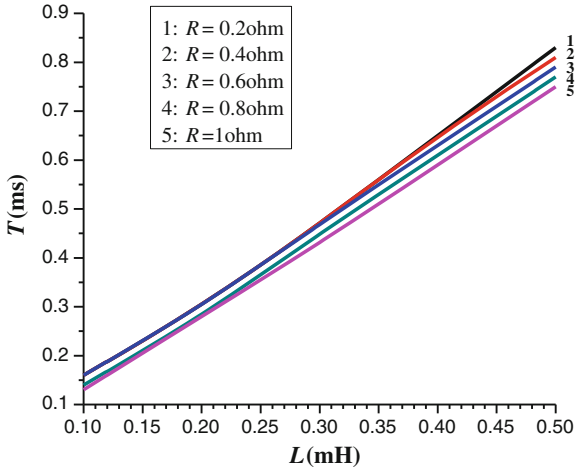


Fig. 21.11 Oscillation duration

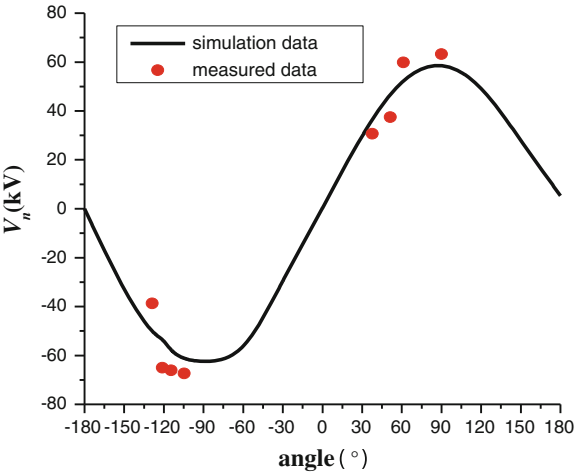


Fig. 21.12 The angle of contact voltage

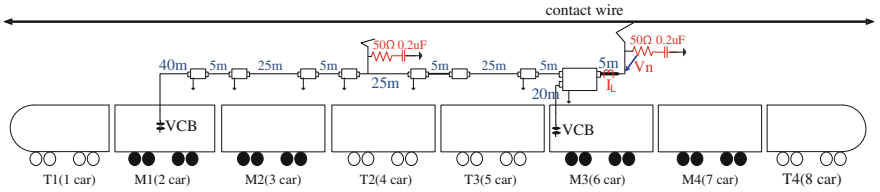


Fig. 21.13 Wire-connection diagram of high voltage cable

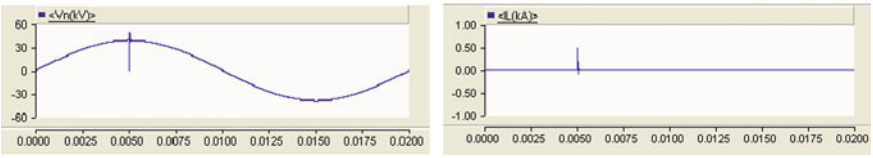


Fig. 21.14 Contact wire voltage V_n and Cable current I_L

Table 21.2 Simulation values and measured data

Peak values	Without RC filter	With RC filter
Peak voltage (kV)	61.35	48.62
Peak I_L (kA)	0.21	0.48
Transient time (ms)	0.25	0.05

21.6 Conclusions

The lifting pantograph sparks and arcs of CRH_2 EMU are due to the charging process of the high voltage cable. The source inductance and resistance have effects on the transient process by affecting the damping component.

The PSCAD/EMTDC simulation model has been verified by field tests, which can help to study the mechanism of the electromagnetic transients and how do the circuit elements affect this process.

One solution to reduce the overvoltage has been proposed. The feasibilities have been validated by simulation.

References

1. Xingchang J (2011) Study on pantograph—catenary arc in the process of rise and drop pantograph. Ph.D. dissertation, Southwest Jiaotong University, Chengdu, China. (in Chinese)
2. Midya S, Bormann D, Larsson A, Schutte T and Thottappillil R (2008) Understanding pantograph arcing in electrified railways—influence of various parameter. In: Proceedings of the. IEEE international symposium on electromagnetic compatibility
3. Midya S, Bormann D, Schutte Z and Thottappillil R (2009) Conducted and radiated emission from pantograph arcing in AC traction system. In: Proceedings of the IEEE power and energy society general meeting
4. Bartlett E.J, Vaughan M and Moore P.J (1999) Investigation into electromagnetic emissions from power system arcs. In: Proceedings of the international conference and exhibition on electromagnetic compability, EMC York, York, UK, pp. 47–52
5. Tellini B, Macucci.M, Giannetti.R, and Antonacci G.A (2001) Conducted and radiated interference measurements in the line-pantograph system. IEEE Trans Instrum Meas, vol. 50, no. 6, pp. 1661–1664

6. Shuguang Z (2010) CRH_2 Electrical Multiple Units. China Railway Press. (in Chinese)
7. Tianzhi L, Guangning W, Lijun Z, Guoqiang G (2011) Pantograph arcing's impact on locomotive equipments. In: Proceedings of the IEEE 57th Holm Conference
8. Junru W and Xueshan L (2010) The influence of the load on electrical locomotive to arc. China Water Transp, vol. 10, no.10, pp. 107–112 (in Chinese)

Chapter 22
Calculations of Leakage Impedance of Rail to Earth in Ballastless Track by Finite Element Method

Teng Li, Mingli Wu, Fan He and Kejian Song

Abstract The leakage resistance of rail to earth is an important parameter of electrified railways, which directly affects the characteristics of traction return current and railway signal track circuit. The ballastless track usually has a much larger value for the leakage resistance of rail to earth compared to ballasted track. In addition, its distributed capacitance effect of rail to earth should not be ignored. The rail has an anharmonic cross-section, and is installed on the spacing support of the concrete integrated ballastless bed in ballastless track structure. The leakage impedance of rail to earth cannot be calculated using a practically simple theory or method, such as the model of thin circular conductor of overhead transmission lines. This paper presents the test and calculation results of the rail leakage impedance in Jin–Qin (Tian jin city–Qin huangdao city) high-speed railway of China. The finite element model is established by using the Comsol Multiphysics simulation software. Comparison of the calculated and measured results shows that the finite element model for calculating the leakage impedance of rail to earth is an effective approach. The field measured data can be used in the future for the related analysis and calculation of traction networks and track circuits.

Keywords Electrified high-speed railway · Ballastless track · Leakage impedance of rail to earth · Finite element model

22.1 Introduction

Jin–Qin (Tianjin-city–Qinhuangdao-city) Passenger Dedicated Line (PDL) with a design speed of about 300 km/h, is an important high-speed electrified railway in northern China; its ballastless track length accounts for 63.1 % of the total length [1].

T. Li (✉) · M. Wu · F. He · K. Song
School of Electrical Engineering, Beijing Jiaotong University, No.3 Shangyuan Cun, Haidian, Beijing, China
e-mail: liteng@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_22,
© Springer-Verlag Berlin Heidelberg 2014

Jin–Qin PDL adopts CRTS II track structure of slab track [2], which is a type of whole track bed. Compared to traditional ballasted track, ballastless slab track has low structural height, good stability, long life, as well as advantages that ballast does not splash under the high-speed train, and the lateral movement of the track bed is small [3]. However, compared to ballasted track rail, leakage resistance of the ballastless track increases, and the distributed capacitance effect of the rail to earth cannot be ignored. Due to the big impact of the high-speed train rolling stock wheels on the rails, the pad between the rail and the sleeper pad is generally thickened in ballastless track, which is electrically insulating; train control signal also requires good insulation between the two rails and rail-to-ground. These factors have led to the larger rail-ground leak resistance. For high-speed railway, appropriate measures must be taken to reduce the track current and rail potential.

With the development of the railway to high-speed heavy load direction, the traction return current correspondingly increases; track circuit signal equipments and the structure of the track bed have also been affected, including the transmission length, the sensitivity, and the security of the track circuit transmission. Suppose the rail self-impedance and traction load current have unchanged values, if the rail leakage resistance to ground is too large, it may cause excessive rail potential. Excessive rail potential will affect the operation of the power supply system performance, personal safety to the station passengers and maintenance personnel, and may also damage signal equipment insulation, and do harm to the safety of running trains [3].

22.2 Theory

The rail line can be regarded as transmission line with distributed parameters [4–6]. The so-called distributed parameters is that the transmission line has a per unit length resistance R, inductance L, capacitance C, and leakage conductance G. The following Fig. 22.1 shows the equivalent circuit for the track in unit length.

When traction current runs through the rails, loss generates in both the internal and external rails. Rail external resistance reflects the leakage active power losses. Distributed capacitance between rails and the ground reflects the external reactive power. Note that this article does not consider the effect generated by the mutual inductance and capacitance effects between the rail line and overhead contact line.

There are a lot of literature on the calculation of the track resistance, inductance. The calculation methods include the equivalent cylindrical method, conductor subdivision scheme [7], the finite element method [4, 8], and so on. But there is little literature on the track-to-ground leakage impedance for finite element analysis. Basic theoretical calculation of the distribution parameters of the rail line can be divided into the analytical method and the finite element method. In the analytical method based on simplified conditions for solving the approximate solution, many parameters haven't been taken into account, such as rail shape. The rail has an anharmonic cross-section, and is installed on the spacing support of the

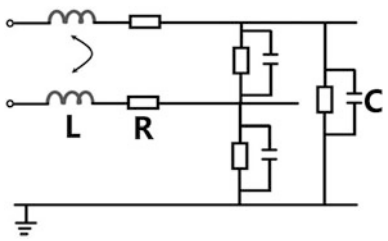


Fig. 22.1 Equivalent circuit for the track in unit length

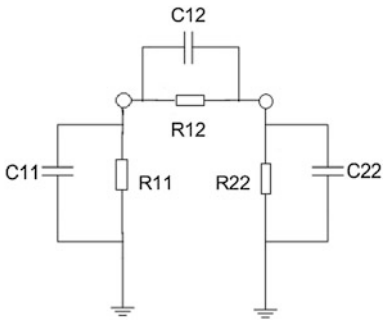


Fig. 22.2 Distributed rail-earth leakage impedance

concrete integrated ballastless bed. The leakage impedance of rail to earth cannot be calculated using a practically simple theory or method, such as the model for thin circular conductor of overhead transmission lines. Considered the actual rail shape, finite element (FE) method subdivides conductor subdivision into many grids and then solves the corresponding physical fields using partial differential equations to calculate the parameters. However, for the consideration of limited information about the materials and the calculation consumption such as time and memory, FE method also needs a certain degree of simplification and estimation. For example, in this paper, considered physical fields are static. Further, the rail inductance is not considered. Calculation of the rail to ground leakage impedance, and then the model verification are the main purposes of this paper, magnetic permeability is thus ignored. Therefore, the unit length rail-earth leakage impedance model is shown in the Fig. 22.2, wherein, R11, R22, C11, and C22 means a single rail leakage resistance or capacitance, R12 and C12 refer to the resistance or capacitance between two rails. Theoretically R11 equals R22, and C11 equals C22. If the electric potential of two rails equals, the leakage impedance of C12 paralleling R12 between two rails can be further ignored; only the rail-ground leakage impedance is considered instead, which is shown in Fig. 22.3.

COMSOL Multiphysics is a finite element analysis and solves software package for various physics and engineering applications, especially coupled phenomena, or multiphysics. The AC/DC Module span electrostatics, magnetostatics, and electromagnetic quasi-statics phenomena [9]. The sub-modules of "Electric Currents," "Electrostatics," and "Magnetic and Electric Fields" can all be used to

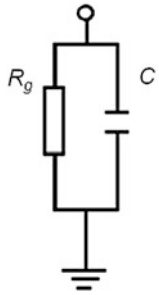


Fig. 22.3 Further simplified leakage impedance

calculate rail line parameters. In this paper, the sub-module of "Electric Currents" is used to calculate leakage impedance without considering the magnetic field, i.e., the rail inductance is ignored. Only the electric field change caused by the electric potential is considered.

The partial differential equations of the "Electric Currents" include:

$$\nabla \cdot J = Q_j \tag{22.1}$$

$$J = \sigma E + j\omega D + J_e \tag{22.2}$$

$$E = -\nabla V \tag{22.3}$$

wherein, J is the current density (A/m^2), E is the electric field intensity (V/m), D is electric displacement field (C/m^2), V is the electric potential (V). In the model, the electric potential of the rail is set to 200 V and the angular frequency ω equals to $2\pi f$, i.e., 314.159 rad/s. There is no space electric charge Q_j considered, i.e., Q_j is 0.

The calculation of leakage impedance of the rail to earth is:

$$Y_0 = \frac{I_{ie}}{V_r} = \frac{\int J_{ie} \mathrm{dS}}{V_r} = G_0 + jB_0 = \frac{1}{R_g} + j\omega C \tag{22.4}$$

$$R_g = \mathrm{Re}(Y_0) \tag{22.5}$$

$$C = \mathrm{Im}(Y_0)/\omega \tag{22.6}$$

where, V_r is the total rail voltage, I_{ie} is the total current injected into the ground per unit rail length, J_{ie} is the current density in the ground over surface dS. r is the leakage resistance of the rail to earth.

The leakage impedance R_g and the distributed capacitance C can be calculated using (22.5) and (22.6) after the admittance being analyzed by (22.4).

22.3 Modeling Results

CRTS II type of slab track is composed of a concrete support base/layer, cement mortar adjustment layer (CA layer), precast concrete track panel, slab connection components, rail fastening clip, etc. Track structure height (internal rail top surface

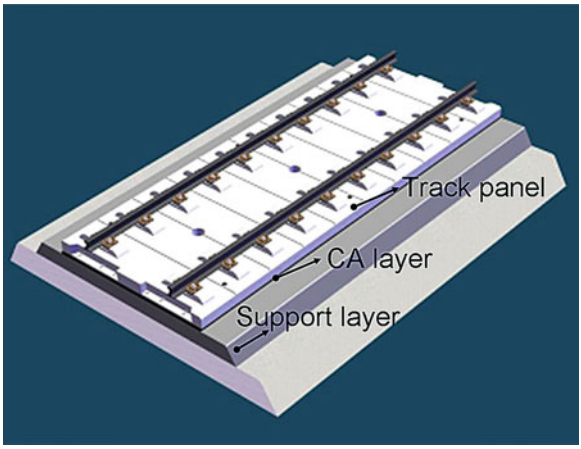


Fig. 22.4 CRTS II slab track structure diagram

to the bottom of the concrete support layer) is 779 mm; track panel width is 2,550 mm, with the thickness of 200 mm and the standard length of 6,450 mm; CA layer theoretical thickness is 30 mm; the support layer made of hydraulic material has the top surface width of 2,950 mm, the bottom surface width of 3,250 mm, and the thickness of 300 mm; each track panel sets 60 transverse tendons with a diameter of 10 mm, and six longitudinal tendons with a diameter of 20 mm [2]. Figure 22.4 shows the general appearance of a standard of the II type slab track. Cross-sectional details of the II type slab track in straight line roadbed section is shown in the following Fig. 22.5.

22.3.1 Model Geometry and Parameters

Result of building a three-dimensional geometry of the II type track slab shown in the Fig. 22.6, which is close to the actual II type track slab geometry. In order to calculate the leakage impedance of the rail to the ground, a ground layer is also built under the concrete support layer with depth of 500 mm, width of 3,650 mm. Cross-sectional details of the model is shown in the Fig. 22.7.

For the program running time and storage memory consideration, the modeled slab belongs to a part of the standard slab since it can be divided into 10 parts with almost the same properties in longitudinal direction, the actual applied model is shown in the Fig. 22.8. The model is meshed into free-tetrahedral grids. Most boundary conditions in the model are set as the default values. Electric potential of ground layer surfaces is set to 0 and electric potential of rail surfaces is set to 200 V. Table 22.1 lists applied material electrical parameters in different layers [10, 11].

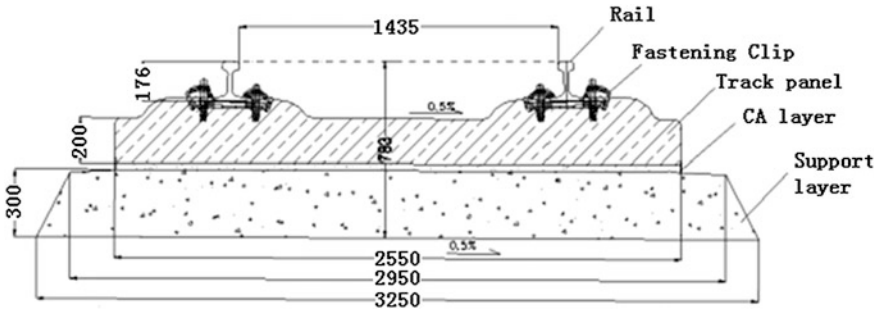


Fig. 22.5 Cross-sectional details of the slab track

Fig. 22.6 Geometry of the II type track slab

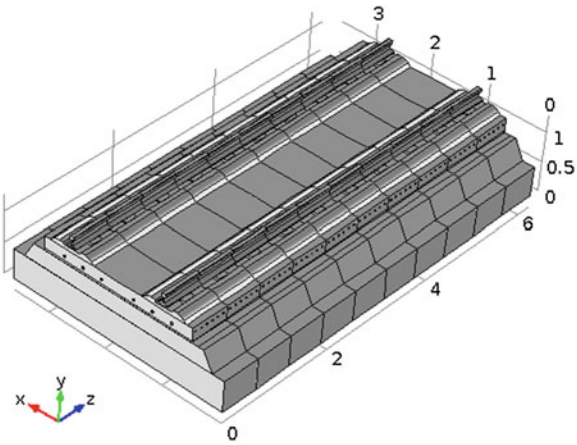
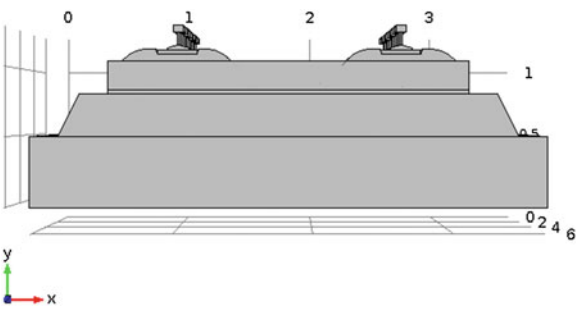


Fig. 22.7 Cross-sectional details of the built model



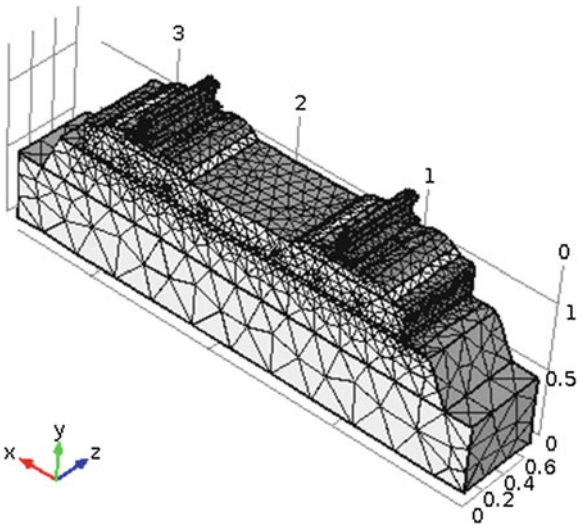


Fig. 22.8 The actual applied model and its mesh

Table 22.1 Applied material parameters in different layers

Materials	Electrical conductivity (S/m)	Relative permittivity
Rail	9×10^5	1
Rubber cushion pad	10^{-5}	10
Track panel	0.001	4
Tendons	8×10^5	1
CA layer	0.002	5
Support layer	0.001	4
Ground layer	0.01	1

22.3.2 Modeling and Calculation Results

Figures 22.9 and 22.10 show the modeling results for the current density norm distribution in cross-sectional and lateral view.

Table 22.2 shows the related modeling results and calculation results for the rail-earth leakage impedance.

22.4 Measurement and Model Verification

The traditional track circuit parameter measurement method is based on transmission line theory and then subsequent on-site measurements. The main rail leakage impedance measurement methods are the "open circuit method" and "short circuit method." Depending on the test equipments, measurement methods

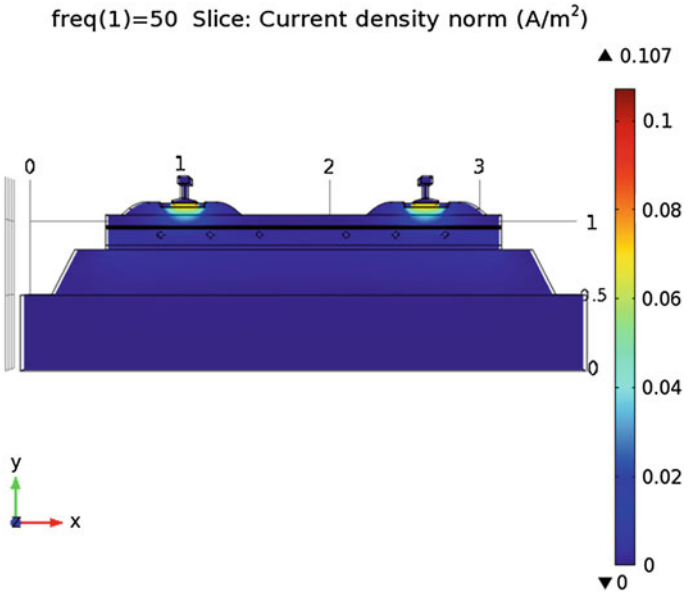


Fig. 22.9 Current density norm *cross view*

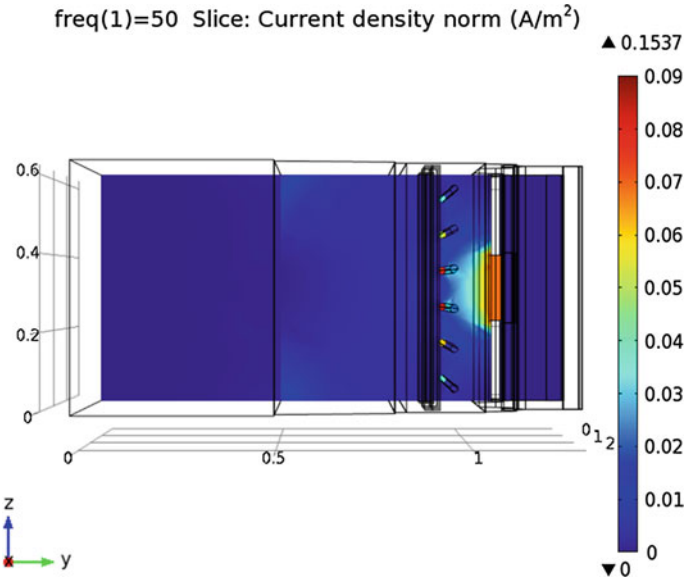


Fig. 22.10 Current density norm *lateral view*

Table 22.2 Modeling and calculation results for the leakage impedance

Electrical parameters	Values
Current density norm of the ground layer top surface (A)	0.00497
Current density, x component of the ground layer top surface (A)	−1.99208e-8 −5.36274e-11i
Current density, y component of the ground layer top surface (A)	−0.00394 −1.06963e-5i
Current density, z component of the ground layer top surface (A)	9.42324e-20 +2.3037e-22i
Resistive losses (W)	0.44334
Leakage resistance (Ω·km)	**2510**
Leakage capacitance (μF/km)	**1.71**

Table 22.3 Measurement results for the leakage impedance

Leakage resistance at 50 Hz (Ω·km)	2483
Leakage capacitance at 50 Hz (μF/km)	1.94

can also be divided into "phase table method," "two voltmeter and an ammeter," "two ammeter and a voltmeter," and other methods [3]. In the actual measurement, combined with Wenner four-probe resistivity test [12] method, we take engineering simplified method, "open circuit method" to measure the rail-to-ground leakage impedance. The actual measured Jin–Qin passenger leakage resistance and distributed capacitance values are shown in the Table 22.3.

Comparing the measured data and modeling data, we found that errors of the leakage impedance between them may caused by the simplification of the model, setting values of the model parameters (e.g., the leakage resistance is affected greatly by the rubber cushion pad), model geometry deviation, and measurement errors, etc. However, the comparison result shows that the model is still valid.

22.5 Conclusion

The calculation results and the measured results are compared with each other. Error factors between them have been analyzed. Results show that, the finite element model for calculating the leakage resistance and the distributed capacitance of rail to earth is an effective approach. The field measured data can be used in the future for the related analysis and calculation of traction networks and track circuits.

Acknowledgments The work was supported in part by the Fundamental Research Funds for the Central Universities of China (**E11JB0033, C13JB00250**) and the BJTU Talent Funds (**E11RC00060**). We would like to thank the anonymous reviewers for their comments and suggestions.

References

1. Li J (2010) Processing Techniques for reinforcing nets of II-type CRTS track slabs for the Tianjin-Qinhuangdao passengers-oriented railway. Traffic Eng Technol Nat Defense 8(5):66–68 (in Chinese)
2. Xue Q, Han JH, Li ZQ (2010) Technology for production of CRTS II track slab for Shijiazhuang-Wuhan passenger dedicated line. Railway Stand Des 9:31–36 (in Chinese)
3. Cao DF, Wu ML, Zhan QC, Chu ZY, Chen JG, Yu YL, Cao XB (2007) Rail leakage resistance test in ballastless track experimental section of Sui-Yu line. Electr Railway 6:10–12 (in Chinese)
4. Hill RJ, Carpenter DC (1993) Rail track distributed transmission line impedance and admittance: theoretical modeling and experimental results. IEEE Trans Veh Technol 42:225–241
5. Hill RJ, Carpenter DC, Tasar T (1989) Railway track admittance, earth-leakage effects and track circuit operation. In: Proc.1989 IEEE/ASME Joint, pp 55–62
6. Mariscotti A, Pozzobon P, Vanti M (2005) Distribution of the traction, return current in AT electric railway systems. IEEE Trans Power Delivery 20:2119–2128
7. Silvester P (1972) Modal theory of skin effect in single and multiple turn coils. Trans Power Apparatus Syst (PAS) 91(1):29–34
8. Triantafyllidis DG, Papagiannis GK, Labridis DP (1999) Calculation of overhead transmission line impedances: a finite element approach. IEEE Trans Power Delivery 14(1):287–293
9. Lüttgen AAE, Bantin CC, Balmain KG (2010) Electric fields from RF tag interrogators underneath an urban rail train. In: Proceedings of the 2010 IEEE antennas and propagation society international symposium (APSURSI), pp 1–4
10. Liu ZY, Zhan ZF (2006) Research on electrical resistivity of concrete and its application in durability appreciation of reinforced concrete. Concrete 10:13–16 (in Chinese)
11. Hou ZY, Gao SB, Ruan Y (2009) Study on rail impedance in ballastless track. Electr Railway 4:48–50 (in Chinese)
12. Yang HN, Yuan JS, Zong W (2001) Determination of three-layer earth model from Wenner four-probe test data. IEEE Trans Magn 37(5):3684–3687

Part II
System Safety and Emergency Management

Chapter 23
Research on Transport Capacity of Urban Rail Transit Based on RailSys

Yanjie Wang and Xi Zhang

Abstract The paper establishes a simulation model to study transport capacity of urban rail transit based on RailSys software and illustrates the core principle and quantitative standard of the method "UIC 406" and the modeling process of dynamic simulation. A concrete example is analyzed by application of RailSys dynamic simulation model. Then find out an effective way to improve transport capacity and put forward the corresponding optimization. The results indicate that RailSys software which is applied to the analysis of dynamic simulation of transport capacity of urban rail transit provides feasible suggestions for the actual operation and management.

Keywords Urban rail transit · Transport capacity · RailSys simulation · UIC 406

23.1 Introduction

Urban rail transit is a mode of transportation which possesses the characteristics of large capacity, high speed, convenience, punctuality, safety, and environmental protection. Developing urban rail transit is the radical way to solve the problem of urban public transportation. China's population and economic growth lead to a surge of traffic volume, which presses for large- and medium-sized cities to develop rail transit as the backbone of public transport system [1].

The research of line transport capacity is helpful to promote the whole reasonable layout and balanced sustainable development of urban rail transit network. Transport capacity is the most important parameter of urban rail transit. All transport organizations should put the transport capacity as a starting point [2]. As

Y. Wang (✉) · X. Zhang
School of Traffic and Transportation, Beijing Jiaotong University, No.3, Shangyuancun, Haidian District, Beijing, People's Republic of China
e-mail: 12120845@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_23,
© Springer-Verlag Berlin Heidelberg 2014

far, there is no special software or system at home and abroad to determine the transport capacity of urban rail transit. There is also no special calculation method for a model. This is a new research field whose exploratory is very strong.

International Railway Union published UIC 406 ability manual which described the railway transport capacity and made railway transport capacity defined as "during a defined period of time, the total number of all possible train paths considering the conditions of actual mixed paths" in 2004 [3]. UIC 406 is a kind of calculation method based on timetable-compressing ability. By modifying the basic timetable, train paths will be compressed together as close as possible. And during the process of compression, it is not allowed to change about the running time in each section and the avoiding, overtaking, and stopping way. The vacant time of compression represents the largest free time that can be used to arrange extra train paths in theory. UIC 406 has the advantage of flexibility which makes it be an efficient reference frame in studying the transport capacity of urban rail transit.

Analysis framework of UIC 406 relies on the infrastructure constructed in advance (fixed equipment), train types (moving equipment) and the specific timetable (the specific method to use the fixed and moving equipment). Germany's RailSys, a microcosmic simulation system based on railway network which is jointly developed by University of Hannover and Co. RMCon, can construct the model of infrastructure, train types and timetable [4]. Therefore, RailSys simulation software may analyze the line transport capacity of urban railway transit combined with UIC 406, studies the timetable optimization scheme suited for marketing demands and then hopes to supply decision supports for the actual transport organization and management of urban rail transit [5].

23.2 Model for Analyzing Transport Capacity by Dynamic Simulation

23.2.1 Principle of Analyzing Transport Capacity

Analysis of transport capacity through UIC 406 is produced by the situation of infrastructure occupation during a specified period of time to measure, namely analysis through calculation of capacity consumption time of infrastructure. UIC 406 puts capacity consumption which is calculated by dividing infrastructure occupation time plus buffer time plus supplement for maintenance by chosen time window as the index measuring capacity consumption.

Capacity consumption may describe the degree of saturation of train timetable. Its value is greater, the train timetable is more saturated, and the space for improving transport capacity is smaller. The formula is described as follows [6]:

$$K = k \times \frac{100}{U} \tag{23.1}$$

$$k = A + B + C \tag{23.2}$$

Wherein K represents capacity consumption [%], k represents total consumption time [min], U represents chosen time window [min], A is infrastructure occupation time [min], B is buffer time [min], and C is supplement for maintenance [min].

23.2.1.1 Infrastructure Occupation Time

After a basic train timetable is established according to market demands, a compression strategy is taken, which puts the entire train paths together compression to the minimum theoretical interval according to the train sequence of timetable without considering any buffer time. The result of compression is namely the final infrastructure occupation time of train paths in the timetable.

A conclusion is drawn by UIC from the application of European railway network: the limiting factor of infrastructure occupation time comes from the setting of buffer time under the condition of considering stability demands; there is always the possibility drawing new train paths, but it must be appropriate because of the specific stability requirement. So, the above compression method is very difficult to give a standard accurate value. Then, the upper threshold (V) of infrastructure occupation is proposed by UIC 406 based on the actual railway management of infrastructure, as described in Table 23.1.

23.2.1.2 Buffer Time

Buffer time shall be time that is inserted between train paths in addition to the minimum interval between trains that arises depending on the signal systems. They serve to reduce transfer of delays from one train to the next. They shall be introduced either between each path, or globally, provided that their total amount is sufficient.

Buffer time has a main effect on absorbing propagation delay and preventing initial delay [7]. The effect of buffer time is as described in Fig. 23.1.

23.2.1.3 Supplement for Maintenance

The main contents of railway maintenance include the repair of works, contact nets and communication signals, etc. The supplement for maintenance may either be part of infrastructure occupation or may be shown as an additional supplement. But, it shall be a part of capacity consumption.

Table 23.1 Upper threshold (V) of infrastructure occupation

Type of line	Peak hour (%)	Daily period (%)	Comment
Dedicated suburban passenger traffic	85	70	The possibility to cancel some services allows for high levels of capacity consumption
Dedicated high-speed line	75	60	
Mixed-traffic lines	75	60	Can be higher when number of trains is low (smaller than 5 per hour) with strong heterogeneity

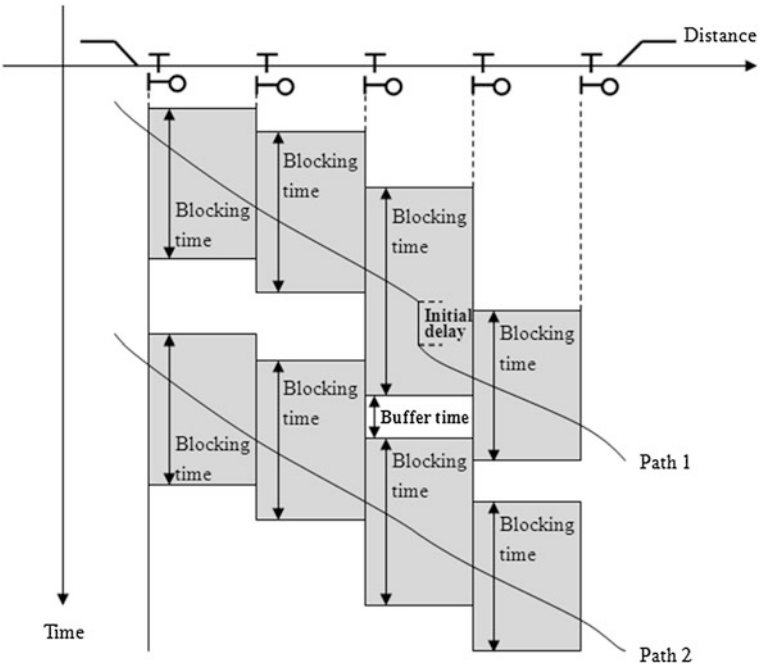


Fig. 23.1 Effect of buffer time

23.2.1.4 Analysis of Capacity Consumption

Capacity consumption shall be analyzed in several steps after compression of timetable:

1. Compress train paths of timetable and obtain capacity consumption (K).
2. The value of capacity consumption as a result of this compression must be compared with certain typical value (V) corresponding to the type of line.
3. If capacity consumption is higher than or equal to this certain typical value, the analyzed line section shall then be called congested infrastructure and no more additional train paths may be added to the timetable.

4. If capacity consumption is lower than this certain typical value, the capacity analysis must then be developed further. Attempts must be made to incorporate further additional train paths of the type corresponding to the particular area into the timetable concerned.
5. If this incorporation is not possible, the leftover capacity is lost capacity which cannot be used any longer.
6. If this incorporation is possible, a certain part of the leftover capacity shall then become usable capacity. In this case, further analysis must be carried out, beginning with the compression of train paths of timetable including the additional train paths.
7. This procedure shall be repeated until either the infrastructure occupation reaches congestion level (see paragraph 3) or no more train paths can be incorporated into the timetable (see paragraph 5).

23.2.2 Process of Dynamic Simulation

The operation of RailSys dynamic simulation model mainly relates to layout of infrastructure, formation of train types, construction and compression of timetable and analysis of capacity consumption about urban rail transit [8]. The concrete process of dynamic simulation is described in Fig. 23.2.

23.3 A Case About Analyzing Transport Capacity of Urban Rail Transit

23.3.1 Design of Simulation Environment

The line is established in the part of layout of infrastructure of RailSys software, which includes five stations from Tongji south Road to Yizhuang station (TJN-JHL-CQN-CQ-YZH) and four sections of Beijing Metro Yizhuang line. The block mode is set to moving-like block. The uplink and downlink input seven pairs of train paths from 6:00 to 8:00 am. The layout of infrastructure is shown in Fig. 23.3.

23.3.2 Analysis of Simulation Results

Through the simulation model which includes layout of infrastructure, formation of train types and construction of timetable, the timetable has been compressed from 6:00 to 8:00 am and the simulation data of capacity consumption has been gotten, as described in Table 23.2.

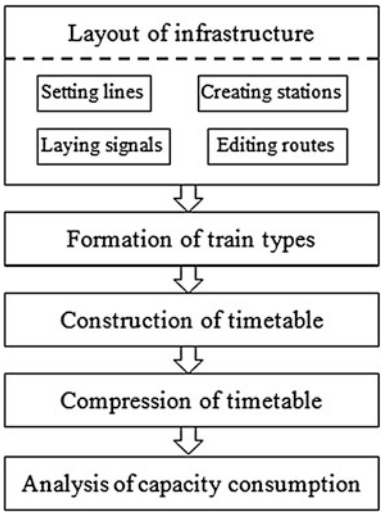


Fig. 23.2 Process of dynamic simulation

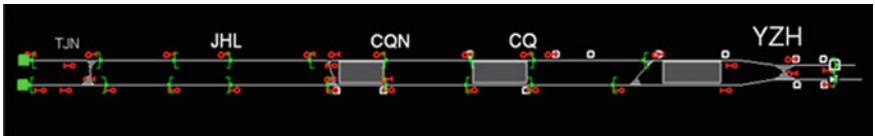


Fig. 23.3 Layout of infrastructure

Table 23.2 Simulation data of capacity consumption (K)

Downlink	Capacity consumption (K) (%)	Uplink	Capacity consumption (K) (%)
TJN → JHL	13.4	YZH → CQ	13.2
JHL → CQN	15.3	CQ → CQN	9.0
CQN → CQ	9.7	CQN → JHL	16.4
CQ → YZH	7.9	JHL → TJN	14.2
TJN → YZH	18.5	YZH → TJN	20.2

Through the train paths of timetable are compressed in the RailSys software, the capacity consumption of each section is calculated. The maximum values of uplink and downlink are 20.2 and 18.5 %, respectively. And the upper threshold (V) of infrastructure occupation about Yizhuang line uses the date (85 %) which is decided by peak hour and dedicated suburban passenger traffic from the recommendation value of UIC 406. Compared the capacity consumption (K) with the upper threshold (V) of infrastructure occupation, K is lower than V. This shows that transport capacity of uplink and downlink may be further improved. Attempts can be made to incorporate further additional train paths of the type corresponding to the particular

area into the timetable concerned. Combined with the market demand, the train paths mustn't be increased until either the infrastructure occupation reaches congestion level or no more train paths can be incorporated into the timetable.

23.4 Conclusion

This paper establishes a simulation model to study transport capacity of urban rail transit based on RailSys software and sets up the modeling process of dynamic simulation. Through the simulation based on five stations and four sections of Beijing Metro Yizhuang line, the transport capacity is analyzed and a practical method is proposed to improve the transport capacity. RailSys software may gradually change train paths until saturated state by quantified analysis on simulation data.

The results show that the application of RailSys simulation software and UIC 406 binding to analyze transport capacity is a very practical quantitative method. This method can provide a suitable timetable scheme based on marketing demands so as to supply decision supports for the actual transport organization and management of urban rail transit. However, the method of RailSys simulation combined with UIC 406 applies only to the calculation of section ability not to calculation of entire railway network. UIC 406 only calculates respectively the capacity consumption of each section which is divided from entire railway network or lines by RailSys software. The capacity consumption of entire railway network or lines is taken as the maximum value of all sections included. It needs to be further investigated that application of RailSys simulation model can provide one more opportunity and satisfactory optimization scheme or not for transport capacity of actual urban rail transit network in the future.

References

1. Li J (2010) Research on evaluation of operation organization of URT based on transport capacity. Southwest Jiaotong University, Chengdu, 5 (in Chinese)
2. Zhang Y (2009) Research on capacity of urban rail transit system bassed on network. Beijing Jiaotong University, Beijing, 6 (in Chinese)
3. Zhang J, Han B (2011) Consumption and utilization of railway capacity: a UIC 406 based analysis. Logistics Technol 30(5):31–33, 115 (in Chinese)
4. Xie S (2006) Development and prospect of simulation programs for rail transit operation. Urban Mass Transit 9(10):64–65 (in Chinese)
5. Zhang H (2011) Research on capacity of urban rail transit bassed on blocking time. Beijing Jiaotong University, Beijing, 12 (in Chinese)
6. UIC Capacity (2004) International Union of Railways, Paris, p 9
7. Tang Y (2013) Timetable stability analysis. Beijing Jiaotong University, Beijing, 4 (in Chinese)
8. Railsys manual (2010) Rail management consultants, Hanover, p 8

Chapter 24
Contribution Rates Calculation for Regulations About Urban Rail Vehicle Inspection and Repair Based on AMSAA Model

Jianwei Li, Minzheng Yuan, Yong Qin, Limin Jia and Zongyi Xing

Abstract Various regulations about urban rail vehicle inspection and repair can ensure the safety of vehicle. This paper proposes an effective calculation method for the contribution rates of regulations based on the decline of the average failure rate. First, the reliability growth mathematical model named AMSAA is introduced. Then the fault time series are used to estimate the model parameters and tests are given to estimate the goodness of fit. According to AMSAA model, the failure rate functions and the average failure rates are calculated. The declines of the average failure rates are used to calculate the contribution rates. Finally, the subway fault time series are used for experimentation. The results show that the method has a high feasibility.

Keywords AMSAA · Reliability · Contribution rate · Failure rate

24.1 Introduction

The urban rail vehicle inspection and repair which matters to the operation safety directly is an important part of the transport system. At the present the inspection and repair modes include "plan inspection," "fault inspection," and "condition repair," and each of them includes a plurality of regulations [1]. The "plan

J. Li · Z. Xing (✉)
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
e-mail: xingzongyi@163.com

M. Yuan
Engineering technology research center, Guangzhou Metro Corporation, Guangzhou 510030, China

Y. Qin · L. Jia
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_24,
© Springer-Verlag Berlin Heidelberg 2014

inspection" is divided into daily inspection, weekly inspection, triple-weekly inspection, monthly inspection, etc. The "fault inspection" mainly refers to repair after the failure, and the regulations in this mode are similarity with the regulations in the "plan repair." The "condition repair" is based on the condition of the vehicle. These regulations about urban rail vehicle Inspection and Repair (hereinafter to be referred as regulations) ensure the safe of the urban rail vehicle. However, the regulations in different transport unit differ. The regulations in the Guangzhou subway are divided into daily inspection, weekly inspection, Half-yearly inspection, yearly inspection, twice-yearly inspection, triple-yearly inspection, minor overhaul, and overhaul. And the regulations in the Nanjing subway are divided into daily inspection, condition inspection, intermediate inspection, and overhaul [2].

In urban rail transit, an accurate calculation method of the security contribution rates about different regulations has a high value. So far, some studies about vehicle regulations have been made. These researches have focused on calculating the reliability index, which can judge a regulation qualitatively. However, most of them just concentrate on one regulation without consideration of all regulations' contribution. For example, Vehicle Component Failures and Balanced-Maintenance Period written by Zuguo Cheng gives detailed instructions on a regulation named balanced-maintenance and describes the balanced-maintenance' contribution to reduce faults qualitatively [3].

In summary, there is almost no method to calculate the contribution rates of various regulations quantitatively at present. This paper proposes an effective calculation method based on the decline of vehicle average failure rate. First, the failure time series is preprocessed. Second, the parameters of AMSAA model are estimate according to failure time series and the goodness of fit is tested. Third, the failure rate functions and the average failure rates are calculated. Then the decreased values of the average failure rates are calculated while one of the regulations is removed. Finally, the decreased values are normalized to get the contribution rates of the regulations. To illustrate the feasibility of the method, a simulation example is given.

24.2 Reliability Growth

24.2.1 Basic Description for Reliability Growth

Reliability growth is a process in which the equipment reliability characteristic quantity improves gradually. Reliability growth ensures the complex system reliability's requirement, and takes place at all stages of the life cycle [4]. It has proved that failure rate of most equipment is a function of time and the typical failure curve is called the bathtub curve [5]. Reliability growth is used in the process of the new equipment development usually. In this phase, weak links are

eliminated which exist in the process of products designing and manufacturing. Appropriate corrective measures are taken constantly to promote improved the equipment reliability. At present urban rail vehicles in the initial failure period or in the random failure period has a series of faults. The faults are constantly exposed, and then corrective measures are taken to eliminate hidden dangers. The vehicle failure rate is reduced and vehicle reliability is improved. This procedure can also be analyzed by reliability growth theory.

In general, reliability growth models include many different models, such as the Duane model, the AMSAA model, the Gompertz model, etc. Among them, the Duane model and the Gompertz model just pay attention to the point estimation of the concerned reliability index while the AMSAA model can give both the point estimation and the interval estimation of the MTBF. Therefore, the AMSAA model is used more widely in engineering [6].

24.2.2 The AMSAA Model

AMSAA model can solve the statistical inference problem of reliability growth data systematically. It has two basic assumptions [7]:

(1) The failure number $N(t)$ of the repairable system in the development interval $(0, t]$ is a nonhomogeneous poisson process with the mean value function $EN(t) = v(t) = at^b$ and the instantaneous failure intensity $\lambda(t) = dEN(t)/dt = abt^{b-1}$:

$$P\{N(t) = n\} = \frac{[v(t)]^n}{n!} e^{-v(t)}, n = 0, 1\ldots \tag{24.1}$$

(2) After the setting moment T, the failure rate function $\lambda(t)$ follows the exponential distribution, that is $\lambda(t) = abt^{b-1}$, $t \geq T$. Let $0 = t_0 < t_1 < t_2 \ldots < t_n$ be the failure time of repairable system in development period $(0, T]$. With $0 < b < 1$, the fault interval $t_i - t_{i-1}$ increase randomly. Then, for the system, reliability is growing. In addition, when time-terminated test is conducted, a and b can be replaced by the unbiased estimators. The formula is as follows:

$$\begin{cases} \bar{a} = \frac{n}{T^b} \\ \bar{b} = \frac{n-1}{\sum_{k=1}^{n} \ln(T/t_k)} \end{cases} \tag{24.2}$$

According to the formula (24.2), we can obtain the fitting failure rate function.

Cramér-Von Misesas criterion is used to judge the fitting goodness of model, whose test statistic is

$$C_M^2 = \frac{1}{12M} + \sum_{i=1}^{M} \left(Z_i^{\bar{b}} - \frac{2i-1}{2M} \right)^2 \tag{24.3}$$

where $Z_i = t_i/T$, $M = n$. If $C_M^2 \le C_{M,\alpha}^2$, the reliability growth data accord with AMSAA model; if not, the experimental time will be extended to realize the fitting goodness test successfully [8].

24.3 The Contribution Rates Calculation for the Inspection and Repair Regulations

24.3.1 The Time-Terminated Test Data

In fact, for urban rail vehicle, the fault data primarily come from online operating failures and other fault data are found by inspection and repair. Generally, the more faults are found by inspection and repair, the less online operating failures appear. According to these time data when failures occurred, both the vehicle reliability and the regulations' contribution rates can be calculated.

The time series of failures occurrence should be counted. For online operating failures, within the time T ($T < 10$ years), the time series t_i $(i = 1, 2, \ldots m_0)$consist of various time between the vehicle starting moment and the faults occurrence time. For the other faults found by inspection and repair, within the same time T, $t_{ij}\left(i = 1, 2, \ldots m_j,\ j = 1, 2 \ldots n\right)$ also consist of different time between the vehicle starting moment and the moment when faults are found. Where n is the total number of regulations, $j = 1, 2 \ldots n$ refers to the various regulations, m_0 is the total number of online operating failures, m_j is the total number of vehicle failures found by inspection and repair under the guidance of the jth regulation, and i represents the ith failure.

24.3.2 Contribution Rate Calculation

In this paper, regulations' contribution is calculated by the failure data found by inspection and repair under the guidance of regulation. Obviously, if a certain regulation, the failures which are found in this regulation will be found in other regulation or even occur in the vehicles' travel. This paper assumes that the corresponding failures will all occur in the vehicles' travel if a certain regulation is removed.

As shown in Sect. 24.2.1, on the basis of the obtained failure data, the contribution of different regulations can be gained by the following calculation procedures.

(1) Fault time series preprocessing
The vehicle fault time series found under the guidance of the jth regulation are t_{ij} and the vehicle online operating fault time series is $t_{i0}(i = 1, 2, \ldots m_0)$. In order to obtain the vehicle fault time series while the jth regulations is removed, combine t_{ij} with t_{i0} and further rearrange them by the fault time. The new fault time series as follows can be considered as the vehicle fault time series statistics while the jth regulations is removed.

$$t_{ij}^{'}\left(i = 1, 2 \ldots \left(m_0 + m_j\right), j = 1, 2 \ldots n\right)$$

(2) Fault rate function fitting
According to the formula (24.2), we obtain the vehicle online operating fault rate function $\lambda_0(t) = \overline{a_0}\overline{b_0}T^{\overline{b_0}-1}, (0 < t \leq T)$. Then calculate the vehicle fault rate function $\lambda_j(t) = \overline{a_j}\overline{b_j}T^{\overline{b_j}-1}, (0 < t \leq T)$ when the jth regulation is removed on the basis of formula (24.2).

(3) Fitting goodness test
In this step, the vehicle online operating fault time series is used to calculate the fitting goodness test statistics through the formula (24.3). All the fault time series have to realize the fitting goodness test.

(4) The average failure rates calculation
In this step, the vehicle online operating average failure rates $\bar{\lambda}_0$ is obtained by the following formula.

$$\bar{\lambda}_0 = \frac{1}{T}\int_0^t \lambda_0(t)dt \tag{24.4}$$

While the jth regulation is removed, the vehicle average failure rates $\bar{\lambda}_j$ is calculated as follows:

$$\bar{\lambda}_j = \frac{1}{T}\int_0^t \lambda_j(t)dt \tag{24.5}$$

(5) The contribution calculation
The contribution rate of the jth regulation is obtained by calculating the difference between $\bar{\lambda}_0$ and $\bar{\lambda}_j$.

$$^{\Delta}\overline{\lambda}_j = \overline{\lambda}_j - \overline{\lambda}_0 \tag{24.6}$$

(6) The contribution rate calculation
$\Delta\bar{\lambda}_j$ is normalized to get the contribution rates of each regulation.

$$C_j = \Delta\overline{\lambda}_j / \sum_{j=1}^{n} \Delta\overline{\lambda}_j \tag{7}$$

Table 24.1 Simulation results

Item number	Faults number	a	b	C_M^2	$C_{M,\alpha}^2$
Fault rate with all regulations	20	0.0102	1.1503	0.067	0.172
Fault rate without the first regulation	25	0.0392	1.0686	0.158	0.172
Fault rate without the second regulation	17	0.0364	1.0503	0.147	0.170
Fault rate without the third regulation	17	0.0433	1.0238	0.140	0.170
Fault rate without the fourth regulation	19	0.0216	1.1371	0.046	0.171

24.4 Simulation and Results Analysis

24.4.1 Simulation Data Acquisition

With an example of the vehicle fault time in 2 years ($T = 730$ days) from one subway, four different regulations keep the subway safe. t_0 = [30, 150, 166, 196, 218, 276, 282, 288, 294, 304, 356, 402, 422, 486, 584, 586, 620, 626, 690, 720]; t_1 = [34, 40, 144, 150, 180, 184, 194, 200, 216, 224, 278, 278, 280, 300, 384, 408, 428, 436, 514, 532, 536, 536, 604, 640, 712]; t_2 = [38, 112, 138, 150, 190, 190, 208, 218, 286, 294, 344, 370, 476, 490, 560, 644, 710]; t_3 = [58, 132, 134, 136, 136, 138, 152, 152, 194, 362, 362, 364, 414, 564, 578, 708 710]; t_4 = [106, 106, 112, 112, 118, 122, 262, 262, 442, 442, 442, 442, 498, 498, 524, 580, 644, 660, 680].

24.4.2 Results Analysis of Simulation

According to Sect. 24.2.2, the fault time series should be preprocessed first to gain the new fault time series, which will be used to fit the vehicle fault rate function next. Finally, the fitting goodness test is conducted, whose results are shown in Table 24.1.

From Table 24.1, it can be seen that all pass the fitting goodness test. The corresponding faults curves are shown in Fig. 24.1.

As shown in Fig. 24.1,The black one stands for the vehicle online operating fault rate curve with all regulations. Taking the worst situation into consideration, the other four curves stand for the vehicle fault rate curve without one regulation respectively. It is obvious that the black fault rate curve has lower values than the others.

According to the formula in Sect. 24.2.2, the contribution rates of the four regulations to keep the vehicle safe are respectively $C_1 = 0.3205, C_2 = 0.2179$, $C_3 = 0.2178$ and $C_4 = 0.2437$. Therefore, for guaranteeing the safe operation of the vehicle, the first regulation's contribution is bigger and it eliminates more vehicle failures, while each the other three regulations' contribution is generally smaller.

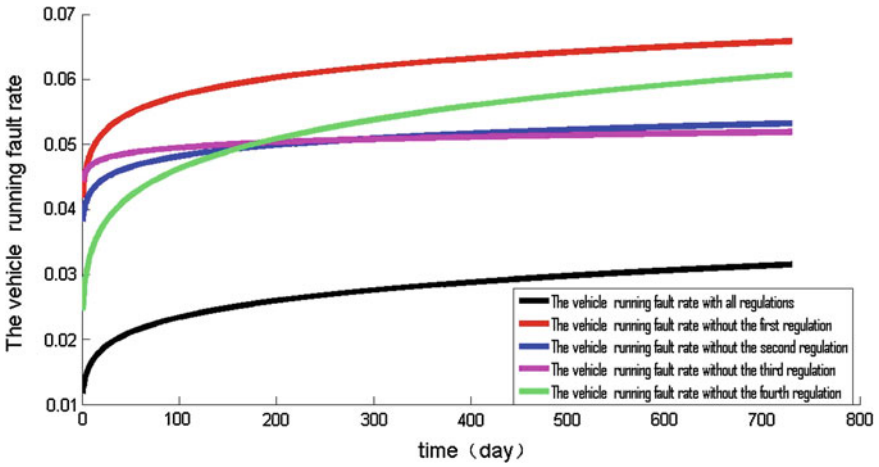


Fig. 24.1 The vehicle fault rate curves

24.5 Conclusion

This paper proposes AMSAA model based on the reliability growth and calculates the contribution rate by the decrease value of average failure. The simulation results show that, this method can calculate the contribution rate for regulations about urban rail vehicle inspection and repair objectively and quantitatively, it has a high practical value.

However, the severity of fault is not considered when choosing failure data. In order to make the contribution rate match the actual situation more, the next work is considering fault severity when using AMSAA model.

Acknowledgments This research was sponsored by National High-tech R&D Program of China (863 Program, No. 2011AA110501) and National Key Technology R&D Program of China (No. 2011BAG01B05). The supports are gratefully acknowledged.

References

1. Quanjun Zeng, Lin Gong (2002) The necessity and applicability of planned repair and “on condition” repair in inspection and repair operation of guangzhou metro cars. Railway Veh 40(7):27–29
2. Jingbing Wu (2008) Discussion on flange gradient (qR) of Nanjing subway passenger car. Railway Locomotive Car 28(1):49–50
3. Zuguo Chen, Jukuan Wang (2006) Vehicle component failures and balanced-maintenance period. Urban Mass Transit 1:46
4. Yamada S, Ohba M, Osaki S (1983) S-shaped reliability growth modeling for software error detection. IEEE Trans Reliab 32(5):475–484

5. Klutke GA, Kiessler PC, Wortman MA (2003) A critical look at the bathtub curve. IEEE Trans Reliab 52(1):125–129
6. Yuanwei M, Shan L (2011) Research on estimation of system reliability based on AMSAA model. Aeronaut Comput Tech 41(4):45–48
7. Zhao Y (2011) Reliability data analysis, vol 4. National Defense Industry Press, Beijing, pp 288–293
8. Lawless JF. Statistical models and methods for lifetime data. Wiley, Hoboken, pp 437–442

Chapter 25
Failure Mode Criticality Analysis of Metro Door System

Jun Xia, Li sha Pan, Xiao qing Cheng, Yong Qin and Zong yi Xing

Abstract Implementing failure mode criticality analysis on metro door system could help to find the failure modes which have great criticality on door system, and it can be helpful for perfecting door maintenance decision. A criticality analysis method which is based on fuzzy evidential reasoning and grey theory is proposed. First, fuzzy belief structure is employed to assess failure modes; second, grey theory is used to calculate the degrees of grey relation of failure modes, that the experiment results show that three failure modes, including "EDCU function is broken," "Limit switch S1 wears out," and "Nut component wears out," have great damage on door system. The results can be used for optimal design and maintenance of the metro door system.

Keywords Metro door · Criticality analysis · Fuzzy evidential reasoning · Grey theory

25.1 Introduction

The metro door system is the channel for passengers to enter and go out of the train, and its work performance is related to personal safety of passengers and safe operation of the metro, it is particularly important to make a critical analysis on door system to identify the weakness in design and maintenance.

J. Xia · Z. y. Xing (✉)
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, People's Republic of China
e-mail: xingzongyi@163.com

L. s. Pan
Engineering Technology Research Center, Guangzhou Metro Corporation, Guangzhou 510030, People's Republic of China

X. q. Cheng · Y. Qin
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, People's Republic of China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_25,
© Springer-Verlag Berlin Heidelberg 2014

Scholars have conducted many studies to find the failure modes, which have great criticality on products. Ahmet [1] has proposed a criticality analysis model which is based on fuzzy TOPSIS; Yang [2] has made failure modes and effects analysis on machining center, and the key components and the failure modes are found; Liu [3] has invented the method which is based on fuzzy evidential reasoning and grey theory.

Until now, any research literature which apply fuzzy evidential reasoning and grey theory to analyze failure modes criticality of metro door system have not been retrieved. Therefore, the method is applied to analyze criticality of door system. The weakness of metro door system can be found and the obtained result can be used for optimal design and maintenance decision.

25.2 Criticality Analysis Method

25.2.1 Fuzzy Evidential Reasoning Theory

Fuzzy evidential reasoning (FER) approach was proposed by Yang [4], and the approach extends the original evidential reasoning individual grades to fuzzy grades. Being combined with the fuzzy interval reasoning which was proposed by Guo [5], the method could deal with fault information which has both fuzzy and interval grade assessments more reasonable.

25.2.1.1 Assessment of Risk Factors Using Belief Structures

Occurrence (O), severity (S), and detection (D) are three risk factors which are evaluated in failure modes criticality analysis. Because of the lack of product failure data, it is difficult to evaluate three risk factors and their weights accurately. As such, we choose linguistic terms for the assessment of risk factors, and the individual evaluation grade set is defined as a fuzzy set H as follows:

$$\begin{aligned} H &= \{H_{11}, H_{22}, H_{33}, H_{44}, H_{55}\} \\ &= \{VeryLow, Low, Moderate, High, VeryHigh\} \end{aligned}$$

Based on the experts' opinions, we can approximate all the five individual assessment grades by trapezoidal fuzzy numbers, and their membership function values is given in Table 25.1 [6, 7].

25.2.1.2 Group Belief Structure

Suppose there are N analysts $(TM_1 \ldots TM_N)$ in the failure modes criticality team responsible for the assessment of K failure modes with respect to three factors

Table 25.1 Fuzzy numbers for assessment grades

Linguistic terms	Fuzzy number
VeryLow	(0, 0, 1, 2)
Low	(1, 2, 3, 4)
Moderate	(3, 4, 6, 7)
High	(6, 7, 8, 9)
VeryHigh	(8, 9, 10, 10)

(O, S, D). Each analyst has the same weight. Let $\tilde{w}_m^n$ be the fuzzy weight of risk factor RF_m given by the member TM_n. The weights of the risk factors are assessed using the linguistic terms in Table 25.2 [8].

Let $\left\{\left(H_{ij}, \alpha_{ij}^n(FM_k, RF_m)\right), i = 1, \ldots 5;\ j = 1, \ldots 5\right\}$ be the belief structure provided by TM_n on the assessment of FM_k with respect to RF_m, where H_{ii} are fuzzy assessment grades defined for risk assessment, H_{ij} are the intervals fuzzy assessment grades between H_{ii} and H_{jj}, and $\alpha_{ij}^n(FM_k, RF_m)$ are the belief degrees to which FM_k assessed on RF_m to the intervals H_{ij}.

The group assessment of all the criticality analysts for failure mode k with respect to risk factor RF_m is also a belief structure, called group belief structure, which is denoted as:

$$C_k^{'}(m) = \left\{\left(H_{ij}, \alpha_{ij}(FM_k, RF_m)\right)\right\}$$

where $m = 1, 2, 3$, corresponding to O, S, and D, respectively. Moreover, α_{ij} (FM_k, RF_m) is referred to as group belief degree and is determined by

$$\alpha_{ij}(FM_k, RF_m) = \frac{1}{N}\sum_{n=1}^{N} \alpha_{ij}^n(FM_k, RF_m) \tag{25.1}$$

In addition, the group belief structures for the K failure modes with respect to the three risk factors form a fuzzy belief decision matrix.

$$C' = \begin{bmatrix} FM_1 \\ FM_2 \\ \cdots \\ FM_K \end{bmatrix} = \begin{bmatrix} C_1^{'}(1) & C_1^{'}(2) & C_1^{'}(3) \\ C_2^{'}(1) & C_2^{'}(2) & C_2^{'}(3) \\ \vdots & \vdots & \vdots \\ C_K^{'}(1) & C_K^{'}(2) & C_K^{'}(3) \end{bmatrix}$$

25.2.1.3 Defuzzification

Chen [9] has proposed an easy defuzzification method for obtaining the crisp number of a fuzzy set, which is shown here

Table 25.2 Fuzzy weights for risk factors

Linguistic terms	Fuzzy number
VeryLow	(0, 0, 0.25)
Low	(0, 0.25, 0.5)
Moderate	(0.25, 0.5, 0.75)
High	(0.5, 0.75, 1.0)
VeryHigh	(0.75, 1.0, 1.0)

$$h_{ij} = \frac{\sum_{i=0}^{n} (b_i - c)}{\sum_{i=0}^{n} (b_i - c) - \sum_{i=0}^{n} (a_i - d)} \tag{25.2}$$

where h_{ij} is the defuzzified crisp number of H_{ij}. After defuzzification, the group assessment of three risk factors is also a crisp number, called group belief structure.

$$C_k(m) = \sum_{i=1}^{5} \sum_{j=1}^{5} h_{ij} \alpha_{ij} (FM_k, RF_m) \tag{25.3}$$

25.2.2 Grey Theory

Grey theory is one of the most used decision-making methods. Its basic idea is to confirm the actual optimal scheme and the criticality sequence of all schemes through calculating the degree of grey relation between each scheme and the ideal optimal scheme on the basis of actual background [10].

25.2.2.1 Building the Comparative Matrix and the Standard Matrix

In the failure modes criticality analysis, the comparative matrix is generated based on the crisp belief decision matrix, that is to say, the comparative matrix $C_c = C$.

The standard matrix can be the lowest level of fuzzy linguistic terms:

$$C_0 = \{C_0(m)\} = \begin{bmatrix} h_{11} & h_{11} & h_{11} \\ \vdots & \vdots & \vdots \\ h_{11} & h_{11} & h_{11} \end{bmatrix}$$

25.2.2.2 Calculating Grey Relation Coefficient

According to the grey theory, the grey relation coefficient between the comparative matrix and the standard matrix is calculated using Eq. (25.4).

$$\lambda(C_0(m), C_k(m)) = \frac{A + \varepsilon B}{|C_0(m) - C_k(m)| + \varepsilon B} \tag{25.4}$$

$$A = \min_k \min_m |C_0(m) - C_k(m)|$$

$$B = \max_k \max_m |C_0(m) - C_k(m)|$$

where ε only affects the relative value. Generally, ε can be 0.5.

25.2.2.3 Calculating the Degree of Grey Relation

According to grey theory, the degree of grey relation $R(C_0, C_k)$ based upon the $\lambda(C_0(m), C_k(m))$ and the group weights $\overline{w}_i$. Firstly, $\overline{w}_i$ could be converted to a crisp value using Eq. (25.2), and then be normalized using Eq. (25.5).

$$\overline{w}_m = \frac{w_m}{\sum_{m=1}^{3} w_m} \tag{25.5}$$

$$R(C_0, C_k) = \sum_{m=1}^{3} \overline{w}_m \lambda(C_0(m), C_k(m)) \tag{25.6}$$

25.2.2.4 Criticality Sequence

The degree of grey relation denotes the relationship between the criticality of each failure mode and the ideal optimal value. The higher of the value obtained reflects the smaller criticality of the identified failure modes.

25.3 An Illustrative Example

According to the statistical analysis of door fault information, seven common failure modes are selected for criticality analysis, as given in Table 25.3.

The group weights of three risk factors and the group belief structures could be calculated by Eqs. (25.1) and (25.2) respectively, as presented in Table 25.4.

Table 25.3 Common failure modes of door system

No.	Failure modes
FM1	Pinch roller becomes loose
FM2	Trolley-guiding wears out
FM3	Long guide pillar has a poor lubrication
FM4	EDCU function is broken
FM5	Limit switch S1wears out
FM6	Nut component wears out
FM7	Nut component becomes loose

Table 25.4 Group assessment on the failure modes and the group weights of three risk factors

Failure modes	Occurrence	Severity	Detection
FM1	{(H_{22}, 0.60), (H_{23}, 0.40)}	{(H_{22}, 0.80), (H_{23}, 0.20)}	{(H_{23}, 0.60), (H_{33}, 0.40)}
FM2	{(H_{22}, 0.80), (H_{23}, 0.20)}	{(H_{22}, 0.40), (H_{23}, 0.40), (H_{33}, 0.20)}	{(H_{22}, 0.2), (H_{23}, 0.40), (H_{33}, 0.40)}
FM3	{(H_{22}, 0.60), (H_{23}, 0.40)}	{(H_{22}, 0.80), (H_{23}, 0.20)}	{(H_{33}, 0.60), (H_{34}, 0.40)}
FM4	{(H_{33}, 0.60), (H_{34}, 0.40)}	{(H_{44}, 0.60), (H_{45}, 0.40)}	{(H_{34}, 0.60), (H_{44}, 0.40)}
FM5	{(H_{33}, 1.0)}	{(H_{34}, 0.40), (H_{44}, 0.60)}	{(H_{33}, 0.60), (H_{34}, 0.40)}
FM6	{(H_{33}, 0.60), (H_{34}, 0.40)}	{(H_{34}, 0.40), (H_{44}, 0.60)}	{(H_{23}, 0.20), (H_{33}, 0.80)}
FM7	{(H_{12}, 0.40), (H_{22}, 0.60)}	{(H_{33}, 0.80), (H_{34}, 0.20)}	{(H_{23}, 0.20), (H_{33}, 0.80)}
Group weights	(0.40, 0.65, 0.85)	(0.55, 0.80, 0.95)	(0.10,0.35,0.60)

Table 25.5 Degrees of grey relation and criticality sequence

Failure modes	Degrees of grey relation	Criticality sequence
FM1	0.890	7
FM2	0.872	5
FM3	0.876	6
FM4	0.652	1
FM5	0.691	2
FM6	0.691	2
FM7	0.831	4

The crisp belief decision matrix could be attained by Eq. (25.3). When the linguistic term VeryLow is defuzzified, the crisp number is 0.130, but the ideal level should be 0, so all the elements of standard matrix should be 0 [11].

The group weights of O, S, and D are respectively 0.35, 0.42, and 0.23 by Eqs. (25.2) and (25.5). The degrees of grey relation and criticality sequence for failure modes could be obtained and presented in Table 25.5 by Eqs. (25.4) and (25.6).

It is clear that the degrees of grey relation of 4 (EDCU function is broken), 5 (limit switch S1 wears out), and 6 (nut component wears out) are far less than other

failure modes, and they should be regarded as the focus of attention in the design improvement and routine maintenance. In addition, reasonable improvement measures which are combined with the actual maintenance should be given to failure mode 7 (nut component becomes loose), 2 (trolley-guiding wears out), 3 (long guide pillar has a poor lubrication), and 1 (pinch roller becomes loose).

25.4 Conclusions

The criticality analysis method takes the relative weights of risk factors as well as uncertainty of assessment into account, so it has a good rationality and applicability. Through criticality analysis, it can be concluded that three failure modes "EDCU function is broken," "Limit switch S1 wears out" and "Nut component wears out" have great criticality on door system and should be the focus of maintenance department. The result is consistent with the field engineers, and then will provide the decision-making reference for the door maintenance.

Acknowledgments This research is supported by National High-tech R&D program of China (863 Program, No.2011AA110501) and National Technology R&D Program of China (No. BAG01B05). Zong yi Xing is the corresponding author.

References

1. Ahmet CK, Mehmet C (2012) Fuzzy failure modes and effects analysis by using fuzzy TOPSIS based fuzzy AHP. Expert Syst Appl 39:61–67
2. Yang ZW, Ren GC, Meng BM (2011) Reliability analysis of failure mode for machining center. Modular Mach Autom Manuf Tech 10:10–16 (in Chinese)
3. Liu L, Liu HC, Lin QL (2011) FMEA method based on fuzzy evidential reasoning approach and grey relational theory. Fuzzy Syst Math 25(2): 71–80 (in Chinese)
4. Yang JB, Wang YM, Xu DL (2006) The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. Eur J Oper Res 17:309–343
5. Guo M, Yang JB, Chin KS (2009) Evidential reasoning approach for multiattribute decision analysis under both fuzzy and interval uncertainty. IEEE Trans Fuzzy Syst 17(3):683–698
6. Feng M (2008) FMEA method based upon fuzzy set theory and grey relational theory. Ind Eng J 11(4):109–112 (in Chinese)
7. Liu H-C, Liu L, Liu N (2012) Risk evaluation in failure mode and effects analysis with extended VIKOR. Expert Syst Appl 39:129126–129134
8. Zhang Z-F, Chu X-N (2011) Risk prioritization in failure mode and effects analysis under uncertainty. Expert Syst Appl 38:206–214
9. Chen CB, Klien CM (1997) A simple approach to ranking a group of aggregated fuzzy utilities. IEEE Trans Syst Man Cybern Part B 27(1):26–35
10. Luo D, Liu S-F (2005) Grey relational decision-making method research. Chin J Manage Sci 13(1):101–106 (in Chinese)
11. Pillay A, Wang J (2003) Modified failure mode and effects analysis using approximate reasoning. Reliab Eng Syst Saf 79:69–85

Chapter 26
Reliability Analysis of the Sliding Plug Door System Based on Bayesian Network

Lingli Mao, Zhaoyi Su, Jing Long, Limin Jia and Zongyi Xing

Abstract The metro door system is one of the high failure rate systems of metro vehicles, and the failure of driver motor is the major cause of the door failure, so the Bayesian network (BN) is applied to analyze the reliability of drive motor. First, the BN of the drive motor is established according to its fault tree, and the failure probability of drive motor is calculated by the method of bucket elimination. Then the fault diagnosis of the drive motor is made through the posterior probability. Finally, the key links of door system are affirmed based on the sensitivity of each basic event, which can provide support and reference for the maintenance of door system.

Keywords Reliability · Metro door · Bayesian network · Fault diagnosis

26.1 Introduction

The metro door system is one of the important components of the subway. Its reliability not only directly impacts on passengers' comfort and convenience, but also is important to the safety of vehicle operation and vehicle's overall performance index. So, it is of great significance to make reliability analysis of the metro

L. Mao · Z. Xing (✉)
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
e-mail: xingzongyi@163.com

Z. Su · J. Long
Engineering Technology Research Center, Guangzhou Metro Corporation, Guangzhou 510030, China

L. Jia
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_26,
© Springer-Verlag Berlin Heidelberg 2014

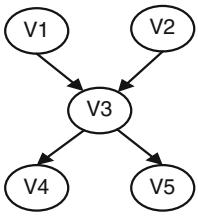


Fig. 26.1 Structure of BN

door system and identify its weakness for the maintenance and optimal design of the metro door system.

Bayesian network can make good use of graph and probability theory to intuitively reveal the structure of the problem, and also reduce the complexity of the reasoning. So BN is widely applied in many fields. Su aoxue presents a reliability analysis method for distribution system based on Dynamic BN [1]; Yu lei studies effectiveness evaluation for ballistic missile defense systems based on battle simulation and experiments and BN analysis theory [2]; Shanghui introduces BN into safety analysis of rail transit system and analyzes the reason and consequence of the hazard source of platform screen door's clipping passengers [3].

This paper will apply BN to the metro door system's motion performance research. Because of the complexity of the door system, the inaction of drive motor is chosen as the research object. The remainder of this paper is organized as follows: Sect. 26.2 shows the general theories of BN and the mapping algorithm from Fault Tree to BN. Section 26.3 discusses a case study, in which BN is applied in the analysis of metro door. Finally, some conclusions are offered in Sect. 26.4.

26.2 Analysis of Bayesian Networks

A Bayesian network (BN) is a directed acyclic graph (DAG) in which the nodes represent the system variables and the arcs symbolize the dependencies or the cause–effect relationships among the variables. Each basic node has a conditional probability table (or function) associated with it, given the state (or value) of the parent nodes. An example of a BN is shown in Fig. 26.1, only the qualitative part is given.

26.2.1 BN Modeling

The construction of a BN includes the determination of structure and the assessment of conditional probability. BN is very similar to Fault Tree Analysis (FTA) from reasoning mechanism and failure description, so in order to make a

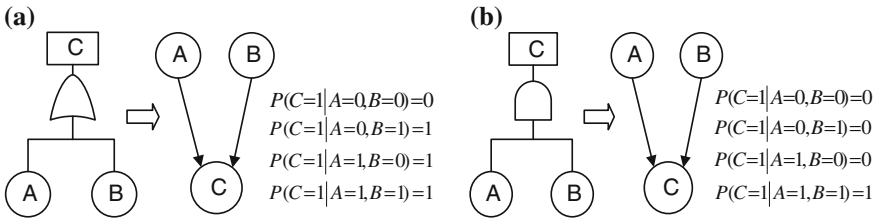


Fig. 26.2 BN expression for basic relations in FTA. **a** OR Gate, **b** AND Gate

convenient research, BN can be mapped from the mature FTA, and the BN has binary events (working/not-working).

26.2.1.1 Mapping of Logic Gate

Figure 26.2 shows the conversion of an OR and an AND gate into equivalent nodes in a BN [4]. Parent nodes A and B are assigned prior probabilities, and child node C is assigned its CPT. Since, the OR and AND gates represent deterministic causal relationships, all the entries of the corresponding CPT are either 0 or 1 s.

26.2.1.2 Mapping Algorithm

The conversion algorithm from FTA to BN proceeds along the following steps [5]:

1) For each leaf node (i.e., primary event or system component) of the FT, create a root node in the BN; If more leaves of the FT represent the same primary event (i.e., the same component), create only one root node in the BN;
2) Assign to root nodes in the BN the prior probability of the corresponding leaf node in the FT (computed at a given mission time t);
3) For each gate of the FT, create a corresponding node in the BN. Connect nodes in the BN as corresponding gates connected in the FT. For each gate in the FT, assign the equivalent CPT to the corresponding node in the BN.

26.2.2 BN Probabilistic Reasoning

The quantitative analysis of a BN may proceed along two lines [6]. A forward (or predictive) analysis concerns the occurrence probability of any node of the network. The calculation of the probability is based on the prior probabilities of the root nodes and the conditional dependence of each node. A more standard backward (diagnostic) analysis that concerns the computation of the posterior probability, with any given set of variables which have given some observation (the evidence).

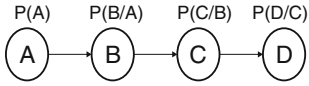


Fig. 26.3 The chain Bayesian network

26.2.2.1 Forward (Predictive) Analysis

According to the conditional independence and the chain rule, BN represents the joint probability distribution $P(U)$ of variables $U = \{A1,\ldots, An\}$ included in the network as (26.1):

$$P(U) = \prod_{i=1}^{n} P(A_i|Pa(A_i)) \tag{26.1}$$

where $Pa(A_i)$ are the parents of A_i in the BN, and $P(U)$ reflects the properties of the BN.

The superiority of BN in inference calculation is that it represents the multi-dimensional distribution in a cost-efficient manner (through its factorized representation), and minimizes the number of parameters required to make the acquisition and expression of knowledge a bit more simply. This paper applies Bucket Elimination [7] which reduces the complexity of reasoning by localized calculation. Using Bucket Elimination, the joint probability corresponding to the BN in Fig. 26.3 is factorized as (26.2):

$$P(D) = \sum_{C} P(D|C) \sum_{B} P(C|B) \sum_{A} P(A)P(B|A) \tag{26.2}$$

26.2.2.2 Backward (Diagnostic) Analysis

A BN's main application is as an inference engine to calculate beliefs of events with given observation of other events, called evidence, according to Bayes rule as (26.3).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{26.3}$$

According to Bayes rule, the sensitivity of the malfunction causes can also be received [8], which can represent the cause event's influence degree of the result event. The calculation formula is as (26.4), X_i means cause event, and X_N means result event. The events have two states: 0 as working, 1 as not working.

$$S_{Xi} = \left| \frac{P(X_N = 1|X_i = 1) - P(X_N = 1|X_i = 0)}{P(X_N = 1|X_i = 0)} \right| \tag{26.4}$$

26.3 Case Study

Electric door and pneumatic door are two kinds of doors applied in subway. The electric door is more widely used in newly built vehicles because of its simple structure, easy control, less maintenance, etc. Electric door is driven by motor. EDCU controls motor's movements when it receives instruction, then the motor rotates to make the locking wheel away from the locking bar and drives screw-nut system to open the door. It is totally opposite to close door. So, the drive motor's action is the basis of door's motion performance. So it is analyzed that why drive motor does not function with BN.

26.3.1 Establishment of Fault Tree and Bayesian Network

According to the failure records analysis of X metro corporation, the basic events causing the inaction of drive motor can be obtained, including motor failure, loose of attachment plug, the internal drive module failure of motor, the loose of EDCU's RS485 plug, etc. Figure 26.4 shows the fault tree of drive motor which is established according to the obtained failure modes. Figure 26.5 shows the BN which is mapped from the tree showed in Fig. 26.4 without the display of conditional probability.

26.3.2 Probabilistic Reasoning

According to the fault record data of the subway, the failure probability of the each basic event which causes the inaction of drive motor is received as shown in Table 26.1.

The probability of root node (inaction of drive motor) can be calculated by Bucket Elimination:

$$P(t=1) = \sum_{b_1,\ldots,b_9,m_1,\ldots,m_4} P(b_1,\ldots,b_9,m_1,\ldots,m_4,t) = 0.2077 \times 10^{-3}$$

The occurrence probability of internal drive module failure of motor ($b_3 = 1$) can be judged when a known evidence is the inactive of drive motor ($t = 1$):

$$P(b_3 = 1|t = 1) = \frac{P(t = 1|b_3 = 1) \cdot P(b_3 = 1)}{P(t = 1)} = 0.5490$$

Table 26.2 shows the other basic events' failure probability with the known inactive of drive motor ($t = 1$). Seen from Table 26.2, when the drive motor does not operate, internal drive module failure of motor (B3) is of a maximum

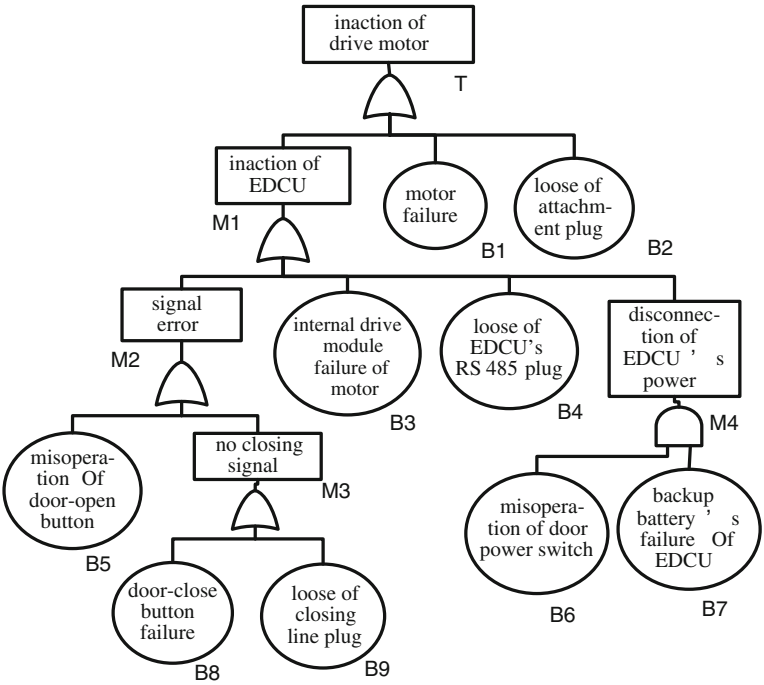


Fig. 26.4 Fault tree of driver motor

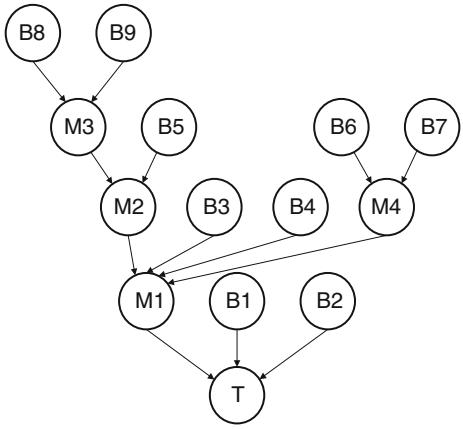


Fig. 26.5 Bayesian model of driver motor

occurrence probability, 0.5490, so it can have the priority of maintenance troubleshooting and exclusion.

Bayesian network can make diagnosis based on several evidence variables. When the drive motor does not operate ($t = 1$), and the connection of EDCU's

Table 26.1 Failure rate of each element of driver motor

Event code	Event description	Fault probability $10e^{-3}$/h
B1	Motor failure	0.0228
B2	Loose of attachment plug	0.0057
B3	Internal drive module failure of motor	0.1140
B4	Loose of EDCU's RS485 plug	0.0342
B5	Misoperation Of open button	0.0057
B6	Misoperation of door power switch	0.0171
B7	Backup battery's failure Of EDCU	0.0114
B8	Close button failure	0.0057
B9	Loose of closing line plug	0.0342

Table 26.2 Failure rate of each element when driver motor in failure

i	1	2	3	4	5	6	7	8	9
$P(b_i = 1 \mid t = 1)$	0.1105	0.0274	0.5490	0.1674	0.0274	0.0237	0.1095	0.0274	0.1647

Table 26.3 Sensitivity of each element of driver motor

i	1	2	3	4	5	6	7	8	9
S_{bi}	4.2882	3.9237	8.4607	8.2678	3.9237	3.8473	3.8473	3.9237	4.5679

power has been confirmed well ($m_4 = 0$), the occurrence probability of internal drive module failure of motor ($b_3 = 1$) can be judged:

$$P(b_3 = 1 \mid t = 1, m_4 = 0) = \frac{P(t = 1, m_4 = 0 \mid b_3 = 1) \cdot P(b_3 = 1)}{P(t = 1, m_4 = 0)} = 0.5527$$

Sensitivity of the leaf nodes can be calculated by Eq. (26.4) to get the key cause failure of the drive motor's inaction. Table 26.3 shows the sensitivity of each basic event. Internal drive module failure of motor (B3) and loose of EDCU's RS485 plug (B4) have the highest sensitivity, 8.4607 and 8.2678, so they should be ensured high reliability.

26.4 Conclusion

This paper applies BN to reliability analysis of metro door system. A BN is established to describe why the drive motor does not operate, then bidirectional inference is made to get the failure probability of child node and parent node, finally the sensitivity of each basic event is calculated. We can get that the most likely cause of the inaction of drive motor is internal drive module failure of motor, and the events having the highest sensitivity are internal drive module failure of motor and loose of EDCU's RS485 plug.

Acknowledgments This research is supported by National High-tech R&D program of China (863 Program, No.2011AA110501) and National Technology R&D Program of China (No. BAG01B05 The supports are gratefully acknowledged.

References

1. Aoxue S, Mingtian F, Zhonglai L (2012) Reliability analysis of distribution system based on dynamic Bayesian Network. East China Electr Power 40(11):1912–1916 (in Chinese)
2. Lei Y, Shuo T (2012) Research on missile defense effectiveness evaluation based on Bayesian network model. J Spacecr TT&C Technol 31(50):89–94 (in Chinese)
3. Hui S, Fei Y, Kaicheng L (2011) The application of Bayesian network in safety assessment of the platform screen door system of rail transit system. Urban Mass Transit 14(11):47–52 (in Chinese)
4. Simon Ch, Weber Ph, Levrat E (2007) Bayesian networks and evidence theory to model complex systems reliability. J Comput 2(1):33–43
5. Bobbio A, Portinale L, Minichino M (2001) Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. Reliab Eng Syst Saf 71:249–260
6. Mahadevan S, Zhang R, Smith N (2001) Bayesian networks for system reliability reassessment. Struct Saf 23:231–251
7. Dechter R (1996) Bucket elimination: a unifying framework for probabilistic inference. In: Proceedings of 12th conference on uncertency in artificial intellience. Morgan Kaufmann, San Francisco, pp 211–219
8. Xiaowei Y (2012) Elements importance and sensitivity analysis based on Bayesian network. J Shenyang Inst Eng 8(3):262–265 (in Chinese)

Chapter 27
The Prediction of Derailment Coefficient Based on Neural Networks

Xiulian Yu, Guangwu Liu, Yong Qin, Yuan Zhang and Zongyi Xing

Abstract Derailment coefficient is an important criterion to evaluate the operating safety of rail vehicles. A derailment coefficient prediction method based on neural network is proposed in this paper. First, the basic concepts of derailment coefficient are briefly discussed. Then the principle of BP and NARX networks and their related learning rules are presented. BP network is compared to NARX network and their disadvantages are outlined. Finally, BP and NARX neural networks are established to analyze their prediction performances. The experimental result shows that, compared with BP neural network, NARX neural network offers better predictive performance of the derailment coefficient.

Keywords Prediction · Derailment coefficient · BP neural network · NARX neural network

27.1 Introduction

Derailment coefficient is an important criterion to evaluate the safety of rail vehicle derailment [1, 2]. In 1896, Nadal gave a simple but most conservative criterion for wheel climb derailment, i.e., Nadal's limit. In 1984, a scientist from

X. Yu · Z. Xing (✉)
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
e-mail: xingzongyi@163.com

G. Liu
Engineering Technology Research Center, Guangzhou Metro Corporation, Guangzhou 510030, China

Y. Qin · Y. Zhang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_27,
© Springer-Verlag Berlin Heidelberg 2014

the National Transportation Test Center called Dr. H. Weinstock proposed a more involved approach to discuss the same problem by introducing the sum of derailment coefficient of two coaxial wheels as derailment criterion. Later on, the scientists from Japanese Railway Technical Research Institute carried out many fruitful researches in the field of derailment mechanism and developed JNR derailment safety standards [3].

The most popular derailment coefficient prediction methods are instrumented wheelset measurement, multibody system dynamics simulation modeling and mechanism modeling. In the instrumented wheelset measurement method the wheels are used as the components to transmit derailment information. But its disadvantages such as excessive costs and high failure rate, limit its promotion [4–6]. The derailment coefficient is got through calculating pharmacokinetic parameters of rolling stock, track, and wheel/rail force. But it has the disadvantages of high purchase price and poor real-time performance [7]. The mechanism modeling methods are often used to get train dynamics response neural networks are computational tools that have found extensive utilization in solving many complex real-world problems. They have been widely used because of their remarkable information processing characteristics [8].

A derailment coefficient prediction model using track irregularity as input is introduced in this paper. The number of hidden layer nodes is determined with experts experiences algorithm and traversal algorithm. Bayesian regularization (BR) algorithm is used to train the neural network. The raw data is provided by the infrastructure testing center of Ministry of Railways. The experimental results show that this method can accurately predict derailment coefficient and wheel load reduction rate.

27.2 Basic Concepts

27.2.1 Track Irregularities

Track irregularities refer to the deviation between the current state and the standard state of track geometry, size and spatial location. There are six types of track irregularities, including longitudinal irregularity, alignment irregularity, cross-level irregularity, twist rail irregularity, gauge irregularity, and complex irregularity [9]. Longitudinal irregularity has obvious influence on vertical vibration, it causes the change of the vertical force and load reduction rate. Alignment irregularity brings about the lateral vibration and the change of derailment coefficient. Cross-level irregularity has a certain effect on both the vertical and horizontal force. In which, cross-level irregularity and gauge irregularity can be calculated from longitudinal irregularity and alignment irregularity.

In this paper, longitudinal and alignment irregularity are used as the inputs of neural network, including left longitudinal irregularity, right longitudinal

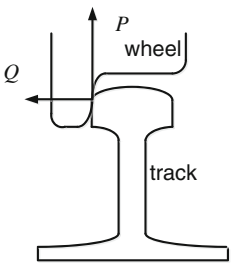


Fig. 27.1 The force diagram on the wheel flange

irregularity, left alignment irregularity and right alignment irregularity, while the outputs are derailment coefficient and wheel load reduction rate.

27.2.2 Derailment Coefficient

Derailment coefficient is used to assess the possibility that the wheels climb or slide up the rail. In general, derailment coefficient is calculated based on the wheel weight and the lateral force data measured from the running experiment. Equation 27.1 is derailment coefficient formula.

$$d = Q/P \tag{27.1}$$

Where Q is lateral wheel-rail forces, P is vertical wheel-rail force, d is derailment coefficient (Fig. 27.1).

27.3 Neural Network

27.3.1 The Structure of BP Neural Network

BP neural network is a multilayered network consisting of an input layer with nodes representing input variables to the problem, an output layer with nodes representing the dependent variables, and one or more hidden layers containing nodes to help capture the nonlinearity in the data. The term back-propagation refers to the way the error computed at the output side is propagated backward from the output layer, to the hidden layer, and finally to the input layer [10].

Consider a BP neural network with three layers. The number of nodes in input layer, hidden layer and output layer are M, N and P, respectively. The output of the hidden layer is shown in Eq. 27.2, and the output of the output layer is shown in Eq. 27.3.

$$o = f_{B_1}\left(\sum_{i=1}^{M} w_{B_1} x_i + b_{B_1}\right) \tag{27.2}$$

$$y = f_{B_2}\left(\sum_{j=1}^{N} w_{B_2} o_j + b_{B_2}\right) \tag{27.3}$$

where the input vector x is $[x_1, x_2, \ldots, x_M]$, the output vector y is$[y_1,y_2, \ldots, y_P]$, the hidden layer nodes vector o is $[o_1, o_2, \ldots, o_N]$, w_{B1} is the connection weight vector between input layer nodes and hidden layer nodes, w_{B2} is the connection weight vector between output layer nodes and hidden layer nodes, f_{B1} is the activation function of hidden layer nodes, f_{B2} is the activation function of output layer nodes, i and j represent the i-th input and the j-th hidden layer output.

27.3.2 LM Algorithm

Levenberg–Marquardt (LM) is used to train the BP neural network. LM algorithm is a unconstrained nonlinear least squares optimization algorithm with second-order convergence rate, and its nature is the modified form of the Gauss—Newton method. It both has the local convergence properties of the Gauss—Newton method and the global convergence properties of the gradient descent method [11].

Assuming that w is weight vector, w^k is the weight vector of k-th iteration, w^{k+1} is the weight vector of $(k + 1)$-th iteration, then:

$$w^{k+1} = w^k + \Delta w \tag{27.4}$$

$$\Delta w = [J^T J + \mu I]^{-1} J^T e_l \tag{27.5}$$

where I is unit matrix, μ is a constant coefficient, J is Jacobian matrix.

27.3.3 The Structure of NARX Neural Network

NARX neural network is a nonlinear regression neural network with external input, it is also called nonlinear autoregressive filter. NARX network consists of input layer, delay layer, hidden layer, and output layer.

Assuming that the number of input delay steps is R, the number of output delay steps is L, the outputs of the j-th hidden layer nodes have the following form

$$x(k) = [x(k), x(k-1), \ldots x(k-R), y(k-1), y(k-2), \ldots y(k-L)] \tag{27.6}$$

$$o(k) = f_{N_1}(x(k) w_{N_1} + b_{N_1}) \tag{27.7}$$

$$y(k) = f_{N_2}(o(k)w_{N_2} + b_{N_2}) \tag{27.8}$$

where x is a P-dimensional input vector, y is a Q-dimensional output node vector,o is N-dimensional hidden layer node vector, R is delay order, L is output delay order, b_{N1} is the threshold of the input layer,b_N2 is the threshold of the hidden layer, w_{N1} is the connection weight between delay layer and hidden layer, w_{N2} is the connection weight between output layer and hidden layer, f_{N1} is the transfer function of the hidden layer nodes, f_{N2} is the activation function of the output nodes.

27.3.4 Bayesian Regularization Algorithm

Bayesian regularization algorithm can be used to modify neural network training performance and improve the generalization capability. The performance evaluation function is the weighted sum of squared weights and the squared errors. The BR algorithm is used to train NARX neural network.

Bayesian Regularization algorithm can adjust α and β adaptively during training. According to the Bayes rule, the posterior probability density function of weights are determined after the learning set is given.

$$p(w|D, \alpha, \beta, M) = \frac{p(D|w, \beta, M)p(w|\alpha, M)}{p(D|\alpha, \beta, M)} \tag{27.9}$$

where D is learning set data, M is the neural network model, $p(D|\alpha, \beta, M)$ is normalization factor. $p(w|\alpha, M)$ is the priori probability density function of weight vector, $p(D|w, \beta, M)$ is the probability density function of the output.

27.4 Simulation Analysis

The experimental train from a Metro Corporation is chosen as study object to verify the effectiveness of the method. According to treatment measures of running gear, the hazard state is divided into three levels, in which level I represents the running gear can work properly, level II represents the running gear can work but needs attention, and level III represents some measures must be adopted to prevent the occurrence of worse failure.

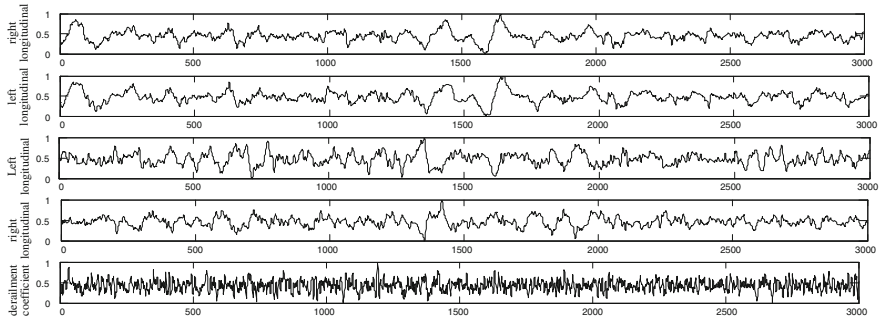


Fig. 27.2 The normalized data

27.4.1 Data Preprocessing

The experimental data in this paper are provided by the Ministry of Railways Infrastructure Testing Center. The total number of datasets is 3,000, in which the first 2,500 groups of data are used to train neural network, while the other 500 groups are used to test the obtained model.

When the input data take large values, it will cause larger numbers overriding smaller ones, and premature saturation of hidden nodes, which impedes the learning process. In order to improve the learning efficiency of neural network and to accelerate the speed of convergence, the collected data are linear normalized. The normalization equation is given below.

$$x_i^{\text{scal}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{27.10}$$

where x_i is the raw data. $x_{\max}$ is the maximum value of the variable x and $x_{\min}$ is the minimum value of the variable x.

The normalized data are shown in Fig. 27.2. The figures from top to bottom are left alignment irregularity, right alignment irregularity, left longitudinal irregularity, right longitudinal irregularity, derailment coefficient, and wheel load reduction rate, respectively.

27.4.2 Simulation Results and Its Analysis

The structural parameters of BP neural network and NARX neural network are determined with trial and error method and experience traversal. The structural parameters are shown in Table 27.1. The performance indicators of each model after simulation are shown in Table 27.2.

In summary, both the training and testing RMSE of BP neural network are less than those of NARX neural network, and the correlation coefficient R is greater

Table 27.1 Structure parameters of neural network

	Neural network model	The number of hidden nodes	The number of iterations	The number of delay steps
Derailment coefficient	BP	11	70	–
	NARX	9	20	40

Table 27.2 Performance indicators

	Neural Network Model	Training RMSE	Training R	Testing RMSE	Testing R
Derailment coefficient	BP	0.1213	0.4021	0.1280	0.3219
	NARX	0.0935	0.8039	0.0788	0.7298

than that of NARX neural network. According to the experiment results, the performance of NARX is better than BP. NARX neural network can be used to analyze track irregularity, derailment coefficient, and wheel load reduction.

27.5 Conclusion

To achieve the accurate prediction of derailment coefficient, BP neural network and NARX neural network model are used in this paper. BP neural network is a static neural network that lacks delay and feedback loops and it can't achieve the accurate prediction. As the input of NARX neural network is delayed and contains the feedback of output, the performance of NARX neural network is better than BP neural network. The result shows that the NARX neural network based on BR algorithm can achieve accurate prediction of the derailment coefficient. The future work is to optimize the training algorithm and to improve the structure and prediction accuracy of NARX neural network.

Acknowledgments This research was sponsored by National High-tech R&D Program of China (863 Program, No.2011AA110501) and National Key Technology R&D Program of China (No. 2011BAG01B05). The supports are gratefully acknowledged.

References

1. Zeng J, Guan Q (2007) Wheelset climb derailment criteria for evaluation of railway vehicle running safety. J Traffic Transp Eng7(6):1–5 (in Chinese)
2. Zhai W, Chen G (2001) Method and criteria for evaluation of wheel derailment based on wheel vertical rise. J China Railw Soc 23(2):17–26 (in Chinese)
3. Guan Q (2006) Study on the derailment mechanism and running safety of trains. Southwest Jiaotong University, Beijing

4. Chang C, Wang C, Li L (2007) Study on numerical test of instrumented wheelset for continuous measurement. Railw Locomot Car 27(5):50–54 (in Chinese)
5. Christoph, W (2010) State-of-the-art railway vehicle design with multi-body simulation. J Mech Syst Transp Logist 3(1):12–26
6. Huo W (2007) Vehicle-track coupling dynamics, 2nd edn. China Railway Publishing House, Beijing (in Chinese)
7. Luber B (2009) Railway track quality assessment method based on vehicle system identification. Elektrotechnik und Informationstechnik 126(5):180–185
8. Fujie X, Colin C, Peter W (2007) An inverse railway wagon model and its applications.Veh Syst Dyn 45(6):583–605 (in Chinese)
9. Gualano L, Iwnicki SD, Ponnapalli PVS, Allen PD (2006) Prediction of wheel-rail forces, derailment and passenger comfort using artificial neural networks. In: Proceedings of the EURNEX-ZEL conference, Zilina, Slovakia 30th–31st May 2006
10. Wong CX, Worden K (2007) Generalised NARX shunting neural network modelling of friction. Mech Syst Signal Process 21(1):553–572
11. Zhang S, Kang X, Liu X (2008) Characteristic analysis of the power spectral density (PSD) of track irregularity on Beijing–Tianjin inter-city railway. China Railw Sci 29(5):25–29

Chapter 28
Research on Hazard Evaluation of Urban Rail Train Based on the Extension Theory

Guiling Liao, Yong Qin, Yuan Zhang, Xiaoqing Cheng and Zongyi Xing

Abstract Running gear, as the key equipment of metro vehicles, its hazard evaluation is important for metro operation. But the most existing evaluation methods lack entire evaluation and over-reliance on subjective experience. Therefore, a complete hazard evaluation index system of running gear was established based on the inspection data and expert assessment. And for considering both quantitative and qualitative aspects, the hazard state of running gear was assessed by extension theory and entropy weight method. And specific application to a running gear at different time proved that the proposed method is practicable.

Keywords Running gear · Extension theory · Hazard evaluation · Urban rail train

28.1 Introduction

Due to the rapid urbanization, urban rail transit has been developed fast with its features of large volume, secure, reliable, and fast delivery. Therefore, high quality and efficient service are required. Running gear can ensure the train running and stopping safety. For the running gear, the accurate hazard assessment can effectively prevent accidents and improve the proactive prevention capabilities of urban rail train. Meanwhile, the assessment results can also support trains maintenance, fault prevention, dispatching for operations management department.

G. Liao · Y. Qin · Y. Zhang (✉) · X. Cheng
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian, Beijing, China
e-mail: zhangyuan_111@126.com

School of Traffic and Transportation, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian, Beijing, China

Z. Xing
School of Mechanical Engineering, Nanjing University of Science and Technology, 200 Xiao Ling Wei Street, Nanjing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_28,
© Springer-Verlag Berlin Heidelberg 2014

At present, researches on running gear include two major types: (a) diagnosis and monitoring of components, such as wheels fault diagnosis [1] and on-line monitoring system [2], bearing fault diagnosis [3] and evaluation [4, 5], the suspension device fault diagnosis method [6–8], etc.; (b) evaluating the state of the running gear's subcomponents according to experts' experience [4, 9].

Studies above have achieved effective results, while the majority still lack overall hazard assessment or rely on subjective factors too much. In response to these problems, hazard evaluation index system of running gear has been established in this paper, and then the extension method by using detection data and experts' experience were used to assess the hazard state of running gear.

28.2 Hazard Evaluation Index System of Running Gear

Running gear generally includes wheelset, axle box, bogie, traction drive device, foundation brake, grounding devices, and the suspension system. Wheelset can ensure train running safely with its power of transferring train load, traction, and braking force. Xu et al. [10] optimized wheel repair plan based on wheel diameter and rim thickness value. Axle box is the joint which connects bogie and wheelset. RMS, crest factor, and kurtosis of bearing vibration signal are often selected to diagnosis the state of bearing [11]. Bogie is the carrier for all devices of running gear. Traction drive device is the power source of motor car. Currently, fault diagnosis of it is based on the vibration signal, temperature, and current [12]. Foundation brake and grounding device can guarantee train suspending safety, they are easily worn parts, If the thickness of brake shoe is less than 15 mm, the brake shoe must be replaced. Suspension system can maintain the balance of train, faults such as air spring breakage, material aging, orifice blockage always happen in suspension system.

The running gear hazard evaluation index system established can be seen in Fig. 28.1. It is on the basis of the above analysis and principles of indicators establishment.

28.3 Research on Running Gear Hazard Evaluation Based on Extension Theory

Extension theory is a method of solving real-world contradictive problems developed in recent years. The theory system is a science studying matter-element and its transformations [13]. Entropy method is an objective weighting method, depending on the degree variation of each index. This part describes extension theory and entropy method, and its usage to hazard evaluation.

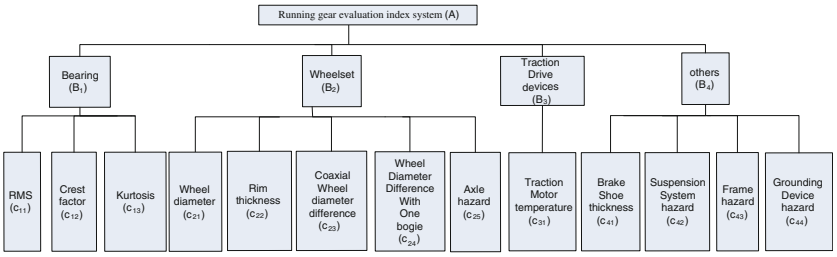


Fig. 28.1 Running gear evaluation index system

28.3.1 Hazard Evaluation Modeling Based on Extension Theory

Classical field is index range of corresponding levels, segment field is the value of all hazard state. Matter-element is the state index value of evaluation category. Assuming the number of hazard level is m, the evaluation set can be got as follows: $M = \{M_1, M_2,\ldots,M_m\} = \{$hazard level I, hazard level II,…, hazard level m$\}$. If R represents the matter, c represents the characteristic, v represents the measure of the characteristic, the classical field matter-element, segment field matter-element, and matter-element can be established as formulas (28.1), (28.2), and (28.3).

$$R_0 = \begin{bmatrix} M & M_{01} & M_{02} & \cdots & M_{0m} \\ c & v_{01} & v_{02} & \cdots & v_{0n} \end{bmatrix} = \begin{bmatrix} M & M_{01} & M_{02} & \cdots & M_{0m} \\ c_1 & \langle a_{11}, b_{11}\rangle & \langle a_{12}, b_{12}\rangle & \cdots & \langle a_{1m}, b_{1m}\rangle \\ c_2 & \langle a_{21}, b_{21}\rangle & \langle a_{22}, b_{22}\rangle & \cdots & \langle a_{2m}, b_{2m}\rangle \\ \vdots & \vdots & \vdots & \vdots & \\ c_n & \langle a_{n1}, b_{n1}\rangle & \langle a_{n2}, b_{n2}\rangle & \cdots & \langle a_{nm}, b_{nm}\rangle \end{bmatrix} \tag{28.1}$$

$$R_P = (M_P, c, v_P) = \begin{bmatrix} M_P & c_1 & v_{P1} \\ & c_2 & v_{P2} \\ & \vdots & \vdots \\ & c_n & v_{Pn} \end{bmatrix} = \begin{bmatrix} M_P & c_1 & \langle a_{P1}, b_{P1}\rangle \\ & c_2 & \langle a_{P2}, b_{P2}\rangle \\ & \vdots & \vdots \\ & c_n & \langle a_{Pn}, b_{Pn}\rangle \end{bmatrix} \tag{28.2}$$

$$R = (M, c, v) = \begin{bmatrix} M & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} \tag{28.3}$$

where: R_0 represents all the status of the evaluation categories; c_i represents the ith hazard evaluation indicator; n represents the number of hazard index; v_{0n} is the value for the range of the c_i;; M_{0m} represents the m-level of the evaluation categories; $<a_{nm}, b_{nm}>$ is the classical field; M_P are all the elements of the comment set; v_{pn} represents the value for the range of c_n., $<a_{pm}, b_{pm}>$ is segment field; M is object to be evaluated; v_i is the value of c_i.

Association degree expresses the degree of an element to its property quantitatively and objectively. Association degree K_j (v_i) between the ith index and the jth hazard state can be calculated as formula (28.4).

$$K_j(v_i) = \begin{cases} \dfrac{\rho(v_i, v_{0ji})}{\rho(v_i, v_{Pi}) - \rho(v_i, v_{0ji})} & v_i \notin v_{0ji} \\ -\dfrac{\rho(v_i, v_{0ji})}{|v_{0ji}|} & v_i \in v_{0ji} \text{ and } \rho(v_i, v_{Pj}) = \rho(v_i, v_{0ji}) \end{cases} \tag{28.4}$$

$$\rho(v_i, v_{0ji}) = \left| v_i - \frac{a_{0ji} + b_{0ji}}{2} \right| - \frac{b_{0ji} - a_{0ji}}{2} \quad (i = 1, 2, \ldots, n) \tag{28.5}$$

$$\rho(v_i, v_{Pi}) = \left| v_i - \frac{a_{Pi} + b_{Pi}}{2} \right| - \frac{b_{Pi} - a_{Pi}}{2} \quad (i = 1, 2, \ldots, n) \tag{28.6}$$

where $\rho(v_i,v_{0ji})$ represents the membership grade between V_i and V_{0ji},, $\rho(v_i,v_{0ji})$ represents the membership grade between V_i and V_{pi}.

In order to calculate the overall association degree the evaluation category has to its different state level, the weight of each index must be proper. Entropy weight method is used to calculate the index weight, it includes four steps. First, build judgment matrix, namely the association degree matrix $K = (k_{ij})_{n\times m}$. Second, get matrix $S = (s_{ij})_{n\times m}$ through normalizing the association degree matrix, where $s_{ij} = k_{ij} / \sum_{j=1}^{m} k_{ij}$. Third, calculate the entropy value of indicator through formula (28.7). If $s_{ij} = 0$, then $s_{ij} \ln s_{ij} = 0$, so $0 \le E_i \le 1$. Finally, calculate the entropy weight coefficient of indicator through Eq. (28.8). After the calculation above, weight coefficient of each index can be acquired. Formula (28.9) calculates the association degree of an evaluated matter on grade j using the connection Function $k_j(v_i)$ and the weight coefficient of the indexes w_i. Then $\max\{K_j(M) j = 1, 2, \ldots, m\}$ is the final evaluation value which presents the grade of the evaluated matter.

$$E_i = -\frac{1}{\ln} \sum_{j=1}^{m} s_{ij} \ln s_{ij} (i = 1, 2, 3, \cdots, n; j = 1, 2, 3, \cdots, m) \tag{28.7}$$

Table 28.1 Index standards and four groups sample data

Category	Level I	Level II	Level III	Sample 1	Sample 2	Sample 3	Sample 4
c_{11}	[0.05, 0.85]	(0.085, 0.1]	(0.1, 1)	0.0644	0.0741	0.0917	0.0566
c_{12}	[2.5, 3.2]	(3.2, 3.5]	(3.5, 30)	3.2241	3.3622	3.5621	3.3541
c_{13}	[2.5, 3.5]	(3.5, 4]	(4, 30)	2.7301	2.9901	2.7782	3.0803
c_{21}	(780, 840]	(775, 780]	[770, 775]	791.5	786.2	782.3	782.9
c_{22}	(29, 32]	(28, 29]	[26, 28]	31.6	29.1	27.2	27.3
c_{23}	[0, 1.5)	[1.5, 2)	[2, 10)	0.9	1.8	0.8	1.7
c_{24}	[0, 3)	[3, 4)	[4, 15)	2.6	3.9	3.85	1.7
c_{25}	[0, 50)	[50, 75)	[75, 100)	20	26	15	29
c_{31}	[0, 70)	[70, 100)	[100, 300)	93	61.6	92.8	52
c_{41}	[20, 55]	[15, 20)	[0, 15)	40	35	17	19
c_{42}	[0, 50)	[50, 75)	[75, 100)	30	24	26	23
c_{43}	[0, 50)	[50, 75)	[75, 100)	15	18	29	20
c_{44}	[0, 50)	[50, 75)	[75, 100)	51	40	55	47

$$\omega_i = \frac{1 - E_i}{\sum_{i=1}^{n} (1 - E_i)} \tag{28.8}$$

$$K_j(M) = \sum_{i=1}^{n} w_i K_j(v_i) \tag{28.9}$$

$$K = \omega K_j(M) \tag{28.10}$$

28.4 Example of Running Gear Evaluation

The experimental train from a Metro Corporation is chosen as study object to verify the effectiveness of the method. According to treatment measures of running gear, the hazard state is divided into three levels, in which level I represents the running gear can work properly, level II represents the running gear can work but needs attention, level III represents some measures must be adopted to prevent the occurrence of worse failure.

28.4.1 Evaluation Data

The index standard of each level is divided in Table 28.1, which is based on the data obtained, maintenance standards in a metro corporation, and experts' advice. The data of four groups selected sample are also showed in Table 28.1.

28.4.2 Classical Field and Segment Field Matter-Element

Classical field and segment field matter-element of running gear hazard evaluation can be acquired as R_0 and R_P following the above analysis.

$$R_0 = \begin{array}{c} \\ c_{11} \\ c_{12} \\ c_{13} \\ c_{21} \\ c_{22} \\ c_{23} \\ c_{24} \\ c_{25} \\ c_{31} \\ c_{41} \\ c_{42} \\ c_{43} \\ c_{44} \end{array} \begin{array}{c} \begin{array}{ccc} M_{01} & M_{02} & M_{03} \end{array} \\ \begin{bmatrix} \langle 0.05, 0.085\rangle & \langle 0.085, 0.1\rangle & \langle 0.1, 1\rangle \\ \langle 2.5, 3.2\rangle & \langle 3.2, 3.5\rangle & \langle 3.5, 30\rangle \\ \langle 2.5, 3.5\rangle & \langle 3.5, 4\rangle & \langle 4, 30\rangle \\ \langle 780, 840\rangle & \langle 775, 780\rangle & \langle 770, 775\rangle \\ \langle 29, 32\rangle & \langle 28, 29\rangle & \langle 26, 28\rangle \\ \langle 0, 1.5\rangle & \langle 1.5, 2\rangle & \langle 2, 10\rangle \\ \langle 0, 3\rangle & \langle 3, 4\rangle & \langle 4, 15\rangle \\ \langle 0, 50\rangle & \langle 50, 75\rangle & \langle 75, 100\rangle \\ \langle 0, 70\rangle & \langle 70, 100\rangle & \langle 100, 300\rangle \\ \langle 20, 55\rangle & \langle 15, 20\rangle & \langle 0, 15\rangle \\ \langle 0, 50\rangle & \langle 50, 75\rangle & \langle 75, 100\rangle \\ \langle 0, 50\rangle & \langle 50, 75\rangle & \langle 75, 100\rangle \\ \langle 0, 50\rangle & \langle 50, 75\rangle & \langle 75, 100\rangle \end{bmatrix} \end{array}$$

$$R_P = \begin{bmatrix} M_P & c_{11} & \langle 0.05, 1\rangle \\ & c_{12} & \langle 2.5, 30\rangle \\ & c_{13} & \langle 2.5, 30\rangle \\ & c_{21} & \langle 770, 840\rangle \\ & c_{22} & \langle 26, 32\rangle \\ & c_{23} & \langle 0, 10\rangle \\ & c_{24} & \langle 0, 15\rangle \\ & c_{25} & \langle 0, 100\rangle \\ & c_{31} & \langle 0, 300\rangle \\ & c_{41} & \langle 0, 55\rangle \\ & c_{42} & \langle 0, 100\rangle \\ & c_{43} & \langle 0, 100\rangle \\ & c_{44} & \langle 0, 100\rangle \end{bmatrix}$$

Association Degree matrix of the indicators bellow B_1, B_2, B_3, B_4 can be calculated following the formula (28.4)–(28.6), whose results are K_{11}, K_{12}, K_{13}, K_{14}. Then the weights of indexes as follows are calculated by using the association degree matrix K_{11}, K_{12}, K_{13}, K_{14} as input of Eqs. (28.7)–(28.8), the results are that $w_1 = (w(c_{11}), w(c_{12}), w(c_{13})) = (0.2557, 0.5585, 0.1859), w_2 = (w(c_{21}), w(c_{22}), w(c_{23}), w(c_{24}),\ w(c_{25})) = (0.1742, 0.1349, 0.2695, 0.2272, 0.1943),\ w_3 = 1, w_4 = (w(c_{41}), w(c_{42}), w(c_{43}), w(c_{44})) = (0.2338, 0.2723, 0.1690, 0.3249)$. Association degree of sample 1 to its hazard level has been calculated by using the formula (28.9)–(28.10), In the equation, w is the weight of B_1, B_2, B_3, B_4. Their values are defined as $w = (w(B_1), w(B_2), w(B_3)) = (0.1163, 0.4528, 0.2404, 0.1851)$ $w = (w(B_1), w(B_2), w(B_3)) = (0.1163, 0.4528, 0.2404, 0.1851)$.

The result is that $K_1 = (0.1352, -0.2309, -0.4349)$, so the sample 1 is at hazard level I. The same method has been used to sample 2, sample 3, and sample 4 respectively, according to the results that $K_2 = (0.0798, -0.0251, -0.3200)$, $K_3 = (-0.0974, 0.0679, -0.1623)$, $K_4 = (0.0503, 0.0173, -0.2909)$, the state of sample 2, sample 4 is hazard levelI, the state of sample 3 is hazard level II. The evaluation results satisfied the relevant experts highly.

$$K_{11} = \begin{bmatrix} 0.4114 & -0.5886 & -0.7120 \\ -0.0560 & 0.1470 & -0.2559 \\ 0.2301 & -0.7699 & -0.8466 \end{bmatrix}$$

$$K_{12} = \begin{bmatrix} 0.1917 & -0.3485 & -0.4342 \\ 0.1333 & -0.8667 & -0.9000 \\ 0.4000 & -0.4000 & -0.5500 \\ 0.1333 & -0.1333 & -0.3500 \\ 0.4000 & -0.6000 & -0.7333 \end{bmatrix}$$

$$K_{13} = \begin{bmatrix} -0.1983 & 0.2333 & -0.0700 \end{bmatrix}$$

$$K_{14} = \begin{bmatrix} 0.4286 & -0.5714 & -0.6250 \\ 0.4000 & -0.4000 & -0.6000 \\ 0.3000 & -0.7000 & -0.8000 \\ -0.0200 & 0.0400 & -0.3288 \end{bmatrix}$$

28.5 Conclusion

Running gear is the key equipment of train, hazard evaluation of which will prevent the occurrence of failure, reduce parts damage and improve the usage of urban rail trains effectively. This paper has built a running gear hazard evaluation index system, and adopted extension theory to assess its hazard level. The assessment results of different samples show that this method can assess the state of running gear accurately and effectively.

Each index has been evaluated to its level in the proposed method, single index and multiple indicators can all be assessed. But classical field and segment field must be chosen. At present, the field value delineated in this paper mainly depends on the sample training data and common standards, but the method is still unclear and uneffectively then researches in the future should take the level domain determination into consideration. At the same time, for the index system modification and improvement should also be studied.

Acknowledgments This research was sponsored by National Hightech R&D Program of China (863 Program, No. 2011AA110501). The supports are gratefully acknowledged.

References

1. Xiao N, Xie J, Zhou S (2010) Method to assess fatigue of tread breaking of metro vehicle wheel and its application. Eng Mech 27(9):234–239 (in chinese)
2. Su D, Wang F, Chi B (2005) The monitoring system of metro vehicle running gear. Mech Eng 2007(7):92–93 (in chinese)
3. Pan L, Chen H, Qin Y et al (2012) Fault diagnosis method for rolling bearing of railway vehicle based on wavelet packet and RBF neural network. Railway Comput Appl 21(7):8–11 (in chinese)
4. Li X (2011) Research on urban rail transit vehicle running gear safety assessment methodology. Doctoral Dissertation of Beijing Jiaotong University, 2011:17–63 (in chinese)
5. Zhang Y, Qin Y, Xing Z et al (2012) Roller bearing safety region estimation and state identification based on LMD-PCA-LSSVM. Measurement (Accepted 2012)
6. Wei X, Jia L, Liu H A (2013) comparative study on fault detection methods of rail vehicle suspension systems based on acceleration measurements. Vehicle Syst Dyn 51(5):700–720
7. Mei TX, Ding XJ (2008) A model-less technique for the fault detection of rail vehicle suspensions. Vehicle Syst Dyn 46:277–287
8. Wei X, Liu H, Qin Y (2011) Fault diagnosis of rail vehicle suspension systems by using GLRT. In: 2011 23rd Chinese control and decision conference, pp 1932–1937
9. Fu X (2007) Research on urban rail transit vehicle failure risk assessment methodology. Master degree thesis of Beijing Jiaotong university, pp 11–72 (in chinese)
10. 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):88–94 (in chinese)
11. Wu M (2009) Research on rolling bearing fault inspection system. Master degree thesis of North China electric power university (in chinese)
12. Fu Y (2008) Safety evaluation monitoring system of train equipment. Master degree thesis of Beijing Jiaotong university, pp 52–71 (in chinese)
13. Yang C, Cai W (2007) Extension engineering. Science Press, Beijing, pp 1–111 (in chinese)

Chapter 29
Reliability Analysis of Metro Door System Based on Fuzzy Reasoning Petri Net

Ping Liu, Xiaoqing Cheng, Yong Qin, Yuan Zhang and Zongyi Xing

Abstract Door system greatly affects the vehicle's safety and operational performance, so it is necessary and urgent to carry out reliability analysis of door system. The methodology proposed in this paper involves FTA and FRPN (Fuzzy Reasoning Petri Net) method. First, fault tree model of door system is established on the basis of failure causality analysis; second, the FRPN model is converted from a fault tree model through graphical transformation rule, and the reasoning algorithm is used to calculate the failure rate of target event. Finally, the important degree of initial events is calculated to evaluate the influence of different faults. The results show that the EDCU failure, DLS breakage, Screw/nut clamping are the most critical faults. The result is helpful for the technical personnel to assess the systems' behavior and to improve their performance by adopting suitable maintenance strategies.

Keywords Sliding plug door · Reliability analysis · Fuzzy reasoning petri net · FTA

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

X. Cheng · Y. Qin · Y. Zhang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

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

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_29,
© Springer-Verlag Berlin Heidelberg 2014

29.1 Introduction

Urban metro vehicle door system is a complex combination of mechanical and electrical subsystems. The door system opens and closes at high frequency due to the short distance between neighboring subway stations. It may easily cause mechanical and electrical failures and finally leads to passenger door system breakdown and resection (electrical isolation). The reliability of door system touches on the vehicle's safety. So doing reliability analysis of the door system is important for keeping urban rail traffic safe.

Currently, researches were conducted on the system door reliability by domestic and overseas scholars, Zhu [1] and Zhou [2] analyze the door system of subway, through the establishment of the fault tree, conclude the weak links led to the doors failure. Zhou Liming [3] proposes a reliability analysis method based on Monte-Carlo simulation and fault-tree analysis method. Due to the lack of effective means of mathematical expression, it is not easy for FTA to conduct further quantitative analysis automatically [4]. The method combining FRPN with FTA is proposed in this paper. This efficient formal and mathematical algorithm is used to utilize intelligent inference and control the spread of the fault accurately [5].

This paper focuses on the application of formal reasoning algorithm based on FTA and FRPN method in the reliability analysis of metro door system. The FRPN model of the system is built on the basis of FTA through the graphical transformation rules. Then qualitative and quantitative analysis is carried out.

29.2 Introduction to Fuzzy Reasoning Petri Net Model

29.2.1 Composition of FRPN

FRPN is a formal and mathematical modeling method, which is widely used in modeling and analyzing discrete event systems such as computer network, transportation, and manufacturing systems [6]. In the case of FRPN, a token incorporated with a place is associated with a real value between 0 and 1; a transition is associated with a CF between 0 and 1. If the statement represented by the FRPN place is true, a token is positioned. And the rule reasoning processes can be expressed by means of firing FRPN transitions. An FRPN model can be defined as an 8-tuple model [5] as shown below:

$$FRPN = (P, T, I, O, H, \gamma, C) \tag{29.1}$$

In which $P = \{p_1, p_2, \ldots, p_n\}$ is a limited set of places; $T = \{t_1, t_2, \ldots, t_n\}$ is a limited set of transitions; $I : P \times T \rightarrow \{0, 1\}$ is the input matrix, presents directed arcs propositions to rules; $O : P \times T \rightarrow \{0, 1\}$ is the output matrix, presents directed arcs rules to propositions; $H : P \times T \rightarrow \{0, 1\}$ is the Complementary arcs from

positions to rules; θ is the truth degree of a proposition's states, $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}^T, \theta \in [0, 1]; \gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}^T :\rightarrow \{0, 1\}$ is the marking vector, if the element $\gamma_i = 1$, the place p_i will get the token; $C = diag\{c_1, c_2, \cdots, c_m\}, c_i \in [0, 1]$ is the confidence matrix, representing the confidence degree of the fuzzy rule t_j.

29.2.2 Modeling of FRPN

In the case of FRPN in knowledge representation, the framework of a net stands for the syntax structure of rule-based knowledge, and FTA is a kind of graphic deductive method, which represents the causal interaction of system failure modes with a tree diagram [7]. From the point of reasoning mechanism and fault description, a FRPN is similar with Fault Tree [8]. So, in this paper, the FRPN model is mapped from the Fault Tree for convenience.

29.2.2.1 Structural Transformation Between FT and FRPN

In general the transfer rules between a fault tree and a Petri net are as follows:

(1) Events of a fault tree are transferred to places of a Petri net, and places of basic events in a fault tree are called initial places and place of top event is called object place for convenience; (2) Consequences between events of a fault tree are transferred to transitions of a Petri net; (3) Logical gates in a fault tree, AND gates and OR gates, can be transformed into symbols in Petri net, Fig. 29.1; especially, a fault tree with other types of logical gates should be transferred to a fault tree only with AND- or OR-gates for simplicity.

29.2.2.2 Algorithmic Transformation Between FT and FRPN

When transferring AND/OR gates into FRPN symbols, the corresponding algorithm is updated, shown in Fig. 29.1. The following are the main rules:

(1) If one transition is fired, the token will be sent to the upper place. (2) If there is a AND gate-like structure, the upper truth value will be the minimum; if there OR gate-like structure, the upper truth value will be the maximum.

29.2.3 Reasoning Algorithm of Fuzzy Reasoning Petri Net

To make positive directional reasoning, MYCIN confidence method is adopted. The max/min algebra is shown as below:

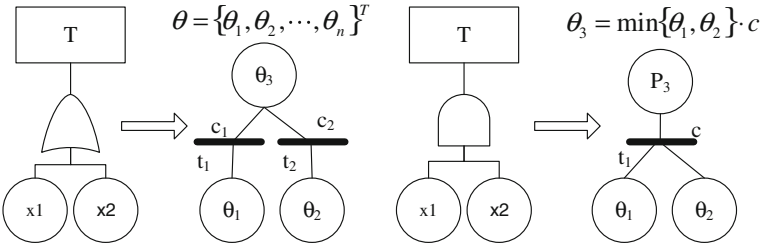


Fig. 29.1 Transformation of the AND gate and the OR gate

(1) $\oplus$, $A \oplus B = D$, where A, B, and D are all $m \times n-$dimensional matrices, such that $d_{ij} = \max\{a_{ij}, b_{ij}\}$;

(2) $\otimes$, $A \otimes B = D$, where A, B and D are $m \times p$, $p \times n$ and $m \times n-$dimensional matrices, respectively, such that $d_{ij=} \underset{1<k<p}{\max} \{a_{ik} \cdot b_{kj}\}$

The firing and control vectors are stated as follows:

$$\begin{aligned} \mu_{m\times 1}^{k} &= 1_{m\times 1} - (I + H)^{T} \otimes \bar{\gamma}^{k} \\ \rho_{m\times 1}^{k} &= 1_{m\times 1} - \left(I^{T} \otimes \left(\bar{\gamma}^{k} \oplus \bar{\theta}^{k}\right)\right) \oplus \left(H^{T} \otimes \left(\bar{\gamma}^{k} \oplus \theta^{k}\right)\right) \end{aligned} \tag{29.2}$$

In which, operator "neg" is adopted in order to get the minimal truth degree value of the input places,

$$\begin{aligned} neg\ \theta^{k} &= \bar{\theta}^{k} = 1_{m\times 1} - \theta^{k} \\ neg\ \gamma^{k} &= \bar{\gamma}^{k} = 1_{m\times 1} - \gamma^{k} \end{aligned} \tag{29.3}$$

The marking and truth degree vectors can be obtained by

$$\begin{aligned} \theta^{k+1} &= \theta^{k} \oplus \left[(O \cdot C) \otimes \rho^{k}\right] \\ \gamma^{k+1} &= \gamma^{k} \oplus \left[O \otimes \mu^{k}\right] \end{aligned} \tag{29.4}$$

Therefore, the reasoning algorithm is given as follows:

Step 1. Set k = 0;

Step 2. Calculate $\theta^{k+1}$from θ^{k} by formulas (29.2) (29.3) (29.4);

Step 3. If $\theta^{k+1} \neq \theta^{k}$, then $k = k + 1$, repeat step2; if $\theta^{k+1} = \theta^{k}$, then terminate the reasoning process.

29.3 Reliability Analysis of Metro Door

29.3.1 FTA Model of Metro Door System

Metro door system is composed of door leaf, mechanism, and control system as shown in Fig. 29.2. The door leaf includes doorplate, sealing joint strip, block pin, and other parts. The mechanism consists of driving/locking device, bearing orientation device, and operating switches. The control system is composed of electronic control (EDCU), executive device, and detection devices.

Failure mode is considered as the manifestation of failure. Passenger door faults are mainly classified into electrical fault and mechanical failures. Electrical fault consists of EDCU breakdown, door motor failure, relay contacts sticking, inching switch translocation, etc., and mechanical failure mainly include deformed door size, unlock unit, mechanical transmission, screw nut wear/cracking, loose screws, etc. Fault tree of metro door system is established as shown in Fig. 29.3.

29.3.2 FRPN Model of Metro Door System

FRPN model of Metro door system is transferred from the FTA model in the last section, shown in Fig. 29.4. The initial places are P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}, P_{11}, P_{14}, P_{15}, P_{16}, P_{17}, P_{18}. Table 29.1 represents the events of FRPN. The FRPN model owns 23 places and 22 transitions, so let $m = 23$, $n = 22$.

29.3.2.1 Probability Analysis of Top Event

The following part of this section gives the description and operation method of the FRPN model according to the main rules. The initial variables are shown as follows:

(1) The input matrix $I_{22 \times 23}$ and output matrix $O_{22 \times 23}$ is defined as

$$I = \begin{bmatrix} I_{7\times7} & A \\ 0 & I_{15\times15} \\ 0 & 0 \end{bmatrix}, O = \begin{bmatrix} 0_{11\times7} & 0_{11\times15} \\ B1 & 0_{7\times15} \\ 0_{5\times7} & B2 \end{bmatrix}$$

In which, $A = \begin{bmatrix} 0_{6\times6} & 0_{6\times1} & 0_{6\times8} \\ 0_{1\times6} & 1 & 0_{1\times8} \end{bmatrix}$,

$$B1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ & & & 0_{5\times7} & & & \end{bmatrix}, \quad B2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

.

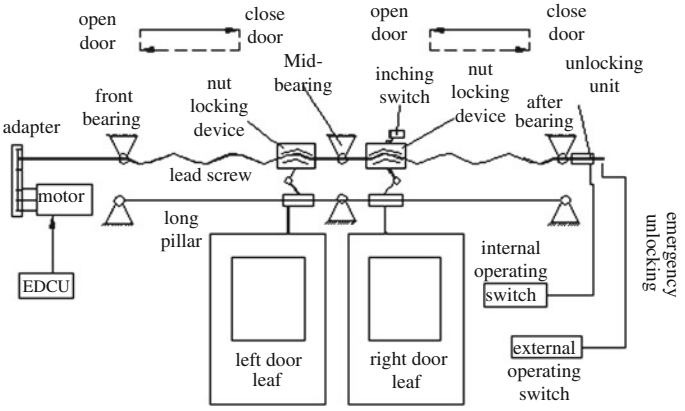


Fig. 29.2 Metro door system diagram

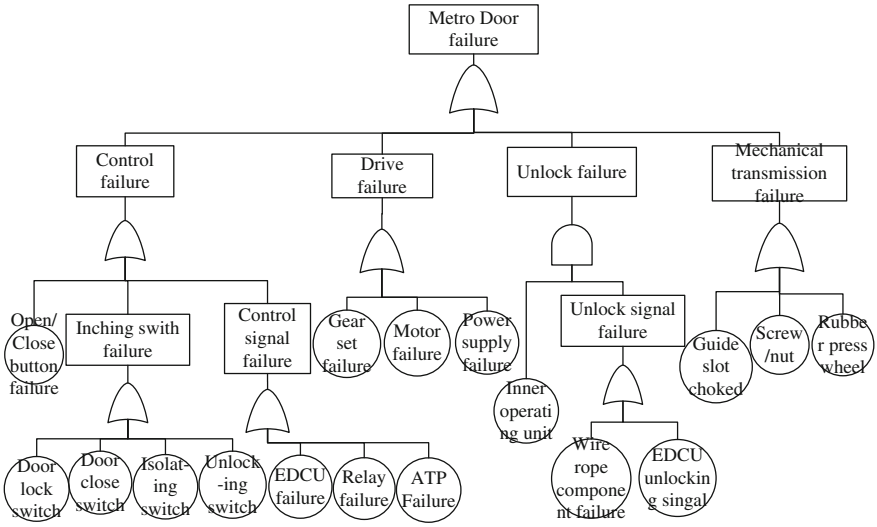


Fig. 29.3 FTA model

(2) The matrix H is described as $H_{22 \times 23} = 0$, as there is no complementary arc from positions to rules.

(3) The matrix C that shows the confidence of transitions can be defined as $C_{22 \times 22} = \text{diag}(0.9, 0.8, 0.7, 0.7, 0.95, 0.7, 0.9, 0.8, 0.85, 0.7, 0.7, 0.85, 0.95, 0.8, 0.7, 0.7, 0.8, 0.75, 0.8, 0.9, 0.85, 0.75)$.

(4) The initial marking vector is $\gamma^0 = [1, 1, \ldots, 1]_{23\times1}$. The truth degree vector is $\theta^0 = (i = 1, \ldots, 11, 14, \ldots, 18) = (0.8, 0.7, 0.4, 0.7, 0.3, 0.3, 0.9, 0.4, 0.6, 0.6, 0.4, 0.1, 0.6, 0.6, 0.7, 0.6)$.

According to the reasoning algorithm, we can obtain the following results:

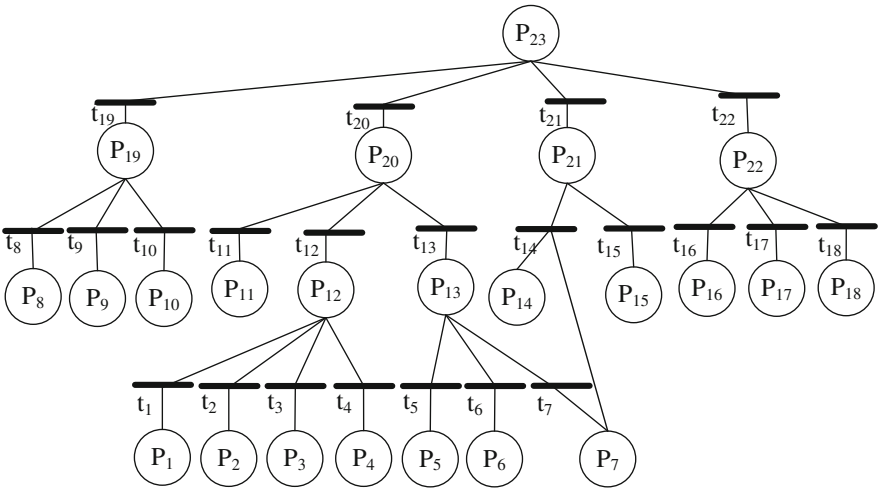


Fig. 29.4 FRPN model

Table 29.1 Events of petri net

Markers	Places	Markers	Places
P_1	Door lock switch	P_{13}	Control signal failure
P_2	Door close switch	P_{14}	Inner operating unit
P_3	Isolating switch	P_{15}	Unlock unit failure
P_4	Unlocking switch	P_{16}	Guide slot choked
P_5	ATP Failure	P_{17}	Screw/nut clamping
P_6	Relay failure	P_{18}	Rubber press wheel
P_7	EDCU failure	P_{19}	Drive failure
P_8	Gear set failure	P_{20}	Control failure
P_9	Motor failure	P_{21}	Unlock failure
P_{10}	Power supply failure	P_{22}	Mechanical transmission failure
P_{11}	Open/close button failure	P_{23}	Metro door failure
P_{12}	Inching switch failure		

$$\theta^1(\theta_i, i = 12, 13, 19, 20, 21, 22, 23) = (0.72,\ 0.81,\ 0.51,\ 0.28,\ 0.42,\ 0.56,\ 0)$$
$$\theta^2(\theta_i, i = 12, 13, 19, 20, 21, 22, 23) = (0.72,\ 0.81,\ 0.51, 0.7695,\ 0.42,\ 0.56,\ 0.42)$$
$$\theta^3(\theta_i, i = 12, 13, 19, 20, 21, 22, 23) = (0.72,\ 0.82,\ 0.51,\ 0.28,\ 0.42,\ 0.56,\ 0.6925)$$
$$\theta^3(\theta_i, i = 12, 13, 19, 20, 21, 22, 23) = (0.72,\ 0.82,\ 0.51,\ 0.28,\ 0.42,\ 0.56,\ 0.6925)$$

With $\theta^4 = \theta^3$, the reasoning process comes to an end. Thus, the maximum confidence of the object event's occurrence is 0.69255 under the initial condition θ^0.

Table 29.2 Index of importance degree of initial places

Places	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
I_{otd}	0.5508	0.4284	0.2142	0.3749	0.2437	0.1796	0.6926	0.256
Places	P_9	P_{10}	P_{11}	P_{14}	P_{15}	P_{16}	P_{17}	P_{18}
I_{otd}	0.408	0.336	0.252	0	0.357	0.315	0.42	0.3375

29.3.2.2 Importance Determination of Basic Events

In order to evaluate the importance and influence of different basic faults, the importance degree I_{otd} associated with the truth degree of the object places is calculated. Consequently, I_{otd} demonstrates the risk degree of the faults.

In I_{otd} calculation, only single fault is considered and all of the transitions are fired. When p_1 gets the token, the initial marking vector is $\gamma = [1, 0, \ldots, 0]_{23\times1}$, then $I_{otd}(1) = 0.5508$ is obtained. Calculation results are listed in Table 29.2. If the I_{otd} is equal to 0, namely, the truth degree of the object place is 0, as in the case of p_{14}, the failure of the system cannot happen. The initial places which result in higher I_{otd} are more dangerous than other places.

The results show that places p_7, p_1, p_2, p_{17} get the highest I_{otd}, which means the EDCU failure, Door lock switch, Door close switch, Screw/nut clamping are the most important events. Technical personnel should pay more emphasis on these components so as to improve the whole system performance.

29.4 Conclusion

In the reliability analysis of the metro door system, the FRPN model of door system model is established based on fault tree model and transformation rules. Probability of the top event and importance degree of basic events are obtained using formal reasoning algorithm. The failure probability of door system is calculated as 0.69255. The importance degree calculation demonstrates the sorting of failure modes. Modeling and reasoning algorithm of FRPN is rooted strictly in mathematical theory, so it has intelligent reasoning ability, and is convenient for computer programming and calculation.

Acknowledgments This research was sponsored by National High-tech R&D Program of China (863 Program, No. 2011AA110501) and National Key Technology R&D Program of China (No. 2011BAG01B05). The supports are gratefully acknowledged.

References

1. Zhu XJ, Wang JB, Yin ZM (2006) On the reliability of passenger compartment door in Shanghai Metro Trains. Urban Mass Transit (3):31–34. (in Chinese)
2. Zhou JL, Wang JB (2006) Reliability and safety research for passenger compartment door of Shanghai Metro Vehicles. Electr Locomotive Mass Transit Veh 29(4):7–10. (in Chinese)
3. Zhou LM, Yang JW, Cai GQ (2010) Monte-Carlo simulation based on FTA in reliability analysis of door system. ICCAE, 2010. The 2nd International Conference on Computer and Automation Engineering, Vol 5. 26–28 Feb 2010, pp. 713–717
4. Wu JN, Yan SZ, Xie LY (2011) Reliability analysis method of a solar array by using fault tree analysis and fuzzy reasoning petri net. Acta Astronautica 69:960–968
5. Ma XQ, Li M, Bao HY (2003) Tape transportation system fault diagnose based on fuzzy reasoning petri nets. J Southeast Univ 33:127–129. (in Chinese)
6. Yuan JM (2011) Petri net modeling method for complex systems in reliability and intelligent analysis. National Defence Industry Press, Beijing. (in Chinese)
7. Quan D, Sun XF, Wang M (2012) Study on reliability of identity module based on fault tree analysis. Mod Manuf Eng 4:122–125, 93. (in Chinese)
8. Lu Q, Huang GQ, Zhu HP (2007) Fuzzy analysis of accidents diagnosis based on fuzzy petri net. Int J Syst Cont 2(3):228–236

Chapter 30
Fault Criticality Evaluation of Metro Door Based on WLSM and FWGM

Jin Bao Ren, Jing Long, Yong Qin, Zong Yi Xing and Jun Xia

Abstract The reliability of metro door is directly related to the safety of passengers. The failure mode and effects analysis (FMEA) method based on the weighted least squares model (WLSM) and the fuzzy weighted geometric mean (FWGM) is proposed to evaluate metro door fault criticality. First, WLSM is used to collect individual judgments to form the final aggregated judgment. Then, fuzzy risk priority numbers (FRPNs) are defined as FWGM of the fuzzy ratings for the three risk factors for prioritization of failure modes. Finally, the FRPNs are defuzzified using the centroid defuzzification method. The experiment results show that three failure modes including ''EDCU function is broken,'' "Limit switch S1wears out," and "Nut component wears out" have greater damage on door than other failure modes.

Keywords Metro door · FMEA · FRPN · Weighted least squares model · Fuzzy weighted geometric mean

J. B. Ren · Z. Y. Xing (✉) · J. Xia
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
e-mail: xingzongyi@163.com

J. Long
Engineering Technology Research Center, Guangzhou Metro Corporation, Guangzhou 510030, China

Y. Qin
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiao Tong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_30,
© Springer-Verlag Berlin Heidelberg 2014

30.1 Introduction

With the rapid development of underground railway traffic, the metro plays an important role in people's life. However, short distances between metro stations and frequently opened metro doors could result in positive line operations door glitches, which is a great threat to the safety of passengers [1]. Therefore, the reliability analysis method of door system was introduced to reduce the failure rate of the metro door.

Failure mode and effects analysis (FMEA) is the widely used reliability analysis technique. Xu et al. [2] constructed a fuzzy assessment system using the expert knowledge and expertise for FMEA. Pillay and Wang [3] presented a new approach using a fuzzy rule base and gray relation theory for FMEA. Sharma et al. [4] used a fuzzy rule-based inference method and the gray theory for prioritizing failure modes. The researches have made great contributions to reliability of metro doors and ensure safety of train operation, but all of them have the shortcoming of fuzziness and uncertainty in traditional FMECA.

In this paper, a new fuzzy-RPNs approach based on the weighted least squares model and the fuzzy weighted geometric mean is proposed for FMEA under uncertainty. A key purpose of FMEA is to identify, evaluate, and rank the potential failure modes by using risk priority numbers. The RPN is employed for the final ranking of failure modes, which is a product of the risk factors occurrence (O), severity (S), and detection (D). The weakness of metro door system can be found through the ranking of failure modes, and the result can be used for optimal design and maintenance decision.

30.2 FMEA Method Based on WLSM and FWGM

The first step in FMEA is to identify all possible potential failure modes of the product or system. After that, critical analysis is performed on these failure modes considering the following three risk factors: O, S, and D. Generally, risk degree of a failure can be represented by a RPN that is defined as the product of the scores of O, S and D, [5]

$$\mathrm{RPN} = O \bullet S \bullet D$$

In order to overcome the drawbacks of fuzziness and uncertainty in traditional FMECA, the risk factors like O, S, and D of each failure mode are used for fuzzy linguistic terms and fuzzy ratings. FRPNs are proposed for prioritization of failure modes. The concrete steps of failure mode decision method based on WLSM and FWGM are as follows.

30.2.1 WLSM are Used Aggregate Multi-Granularity Judgments

A linguistic variable, defined as a variable whose values are expressed in natural or artificial languages, is useful for providing approximate characterization in uncertain environment. The fuzzy RPNs differ from the traditional RPNs, because the input data are linguistic variables, which can be quantified through trapezoidal fuzzy numbers. Assume the positive trapezoidal fuzzy numbers $\tilde{F} = (a_1, b_2, c_3, d_4)$. The decision makers (DM's) give their judgments on failure modes using different linguistic term sets. Let $\tilde{a}_i = (a_{il}, a_{ig}, a_{ih}, a_{iu})$ be the ith positive trapezoidal fuzzy number in a sequence of fuzzy numbers. The distance $d(\tilde{a}_i^k, \tilde{a}_i)$ can be used for measuring the deviation degree between the kth judgment a_i^k and the aggregated judgment $\tilde{a}_i$ of the ith risk factor.

$$d(\tilde{a}_i^k, \tilde{a}_i) = \frac{\left|a_{il}^k - a_{il}\right| + \left|a_{ig}^k - a_{ig}\right| + \left|a_{ih}^k - a_{ih}\right| + \left|a_{iu}^k - a_{iu}\right|}{4} \tag{30.1}$$

Generally, the importance weights of DM's should be taken into consideration. Let $V = (v_1, v_2, \ldots, v_k)$ be a weight vector of DMs, with $\sum_{i=1}^{k} v_i = 1$. In order to ensure that the aggregated assessment reflects all DM's opinions as much as possible, the total aggregated deviation degree for the ith risk factor can be determined by using the fuzzy WLSM as following [6],

$$\begin{aligned} \min D_i &= \sum_{k=1}^{K} v_k d^2(\tilde{a}_i^k, \tilde{a}_i) \\ &= \frac{1}{4}\sum_{k=1}^{K} v_k[(a_{il}^k - a_{il})^2 + (a_{ig}^k - a_{ig})^2 + (a_{ih}^k - a_{ih})^2 + (a_{iu}^k - a_{iu})^2] \end{aligned} \tag{30.2}$$

s.t. $0 \le a_{il}^k \le a_{ig}^k \le a_{ih}^k \le a_{iu}^k \le 10$

All the group judgments on evaluating the failure modes can be determined through solving the optimization problem.

Suppose there are n failure modes, FM_i ($i = 1,\ldots,n$), to be evaluated and prioritized by a FMEA team consisting of m DMs, TM_j($j = 1,\ldots,m$). Let $\tilde{R}_{ij}^O = \left(R_{ijL}^O, R_{ijM_1}^O, R_{ijM2}^O, R_{ijU}^O\right)$, $\tilde{R}_{ij}^S = \left(R_{ijL}^S, R_{ijM}^S, R_{ijU}^S\right)$ and $\tilde{R}_{ij}^D = \left(R_{ijL}^D, R_{ijM}^D, R_{ijU}^D\right)$ be the fuzzy ratings of the ith failure mode on the risk factors O, S, and D, $w_j^O = \left(w_{jL}^O, w_{jM}^O, w_{jU}^O\right)$, $w_j^S = \left(w_{jL}^S, w_{jM}^S, w_{jU}^S\right)$and $w_j^D = \left(w_{jL}^D, w_{jM}^D, w_{jU}^D\right)$ be the fuzzy weights of the three risk factors provided by the jth FMEA team member (TMj), and h_j ($j = 1,\ldots,m$) be the relative importance weights of the m team members, satisfying $\sum_{j=1}^{m} h_j = 1$and $h_j > 0$ for $j = 1,\ldots,m$.

Define the fuzzy risk priority number (FRPN) of each failure mode as [7]

$$\mathrm{FRPN}_i = (\tilde{R}_i^O)^{\frac{\tilde{w}^O}{\tilde{w}^O+\tilde{w}^S+\tilde{w}^D}} \times (\tilde{R}_i^S)^{\frac{\tilde{w}^S}{\tilde{w}^O+\tilde{w}^S+\tilde{w}^D}} \times (\tilde{R}_i^D)^{\frac{\tilde{w}^D}{\tilde{w}^O+\tilde{w}^S+\tilde{w}^D}} \tag{30.3}$$

Let $\tilde{W} = (\tilde{w}_1, \tilde{w}_2, \tilde{w}_3)$ be a fuzzy weight set that can express the importance weights of O, S, and D. Assume $\tilde{E} = (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3)$ be a fuzzy risk score set for a failure mode with O, S, and D. The fuzzy RPNs can be formulated as

$$\tilde{R} = \prod_{i=1}^{3} (\tilde{a}_i)^{\frac{\tilde{w}_i}{\sum_{j=1}^{3} \tilde{w}_j}} \tag{30.4}$$

30.2.2 FWGM are Used to Calculate the α Level Set of FRPN

Fuzzy sets can be represented by intervals, which are called α-level sets. Let Ã be a fuzzy set on the universe of discourse X. Then the α-level sets of Ã are defined as

$$A_\alpha = \{x \in X | u_{\tilde{A}}(x) \geq \alpha\} = [\min\{x \in X | u_{\tilde{A}}(x) \geq \alpha\}, \max\{x \in X | u_{\tilde{A}}(x) \geq \alpha\}] \tag{30.5}$$

Let $(\tilde{R})_\alpha = [R_\alpha^L, R_\alpha^U]$be a α-cut set of $\tilde{R}$. These intervals indicate risk change ranges of each failure mode at the possibility level α. Generally, the mathematical models for calculating the fuzzy RPNs are formulated as follows [8]:

$$R_\alpha^L = \mathrm{Min}\ \exp\left(\frac{\sum_{i=1}^{3} w_i \ln(a_i)_\alpha^L}{\sum_{i=1}^{3} w_i}\right) \tag{30.6}$$

$$R_\alpha^U = \mathrm{Max}\ \exp\left(\frac{\sum_{i=1}^{3} w_i \ln(a_i)_\alpha^U}{\sum_{i=1}^{3} w_i}\right) \tag{30.7}$$

s.t. $(w_i)_\alpha^L \leq w_i \leq (w_i)_\alpha^U$

After solving the above-mentioned models (30.6) and (30.7) with different values of α, the α-cut sets of $\tilde{R}$ can be determined using

$$R_\alpha^L = \exp(y_\alpha^L) \tag{30.8}$$

$$R_\alpha^U = \exp(y_\alpha^U) \tag{30.9}$$

30.2.3 Defuzzify the FRPNs by the Centroid Defuzzification Method

Fuzzy numbers need to be transformed into crisp numbers for comparison or ranking. Such a transformation process is called defuzzification, which can be conducted in many different ways. The most extensively used defuzzification approach is the centroid defuzzification, which is also known as the center of gravity or center of area defuzzification. The centroid method of defuzzification is given by the following algebraic expression.

$$\bar{x}_0(\tilde{A}) = \frac{\int_a^d x u_{\tilde{A}}(x)dx}{\int_a^d u_{\tilde{A}}(x)dx} \tag{30.10}$$

where $\bar{x}_0(\tilde{A})$is the defuzzified value, $u_{\tilde{A}}(x)$is the explicit membership function of a fuzzy number $\tilde{A}$.

When the explicit membership function of a fuzzy number $\tilde{A}$ is not known, but the α-level sets are available, then its defuzzified centroid can be determined by (30.9) and (30.10), [9]

$$\int_a^d u_{\tilde{A}}(x)dx = \frac{1}{2n}[((x)_{\alpha_0}^U - (x)_{\alpha_0}^L) + ((x)_{\alpha_n}^U - (x)_{\alpha_n}^L) + 2\sum_{i=1}^{n-1}((x)_{\alpha_i}^U - (x)_{\alpha_i}^L)] \tag{30.11}$$

$$\int_a^d x u_{\tilde{A}}(x)dx = \frac{1}{6n}[((x)_{\alpha_0}^{2U} - (x)_{\alpha_0}^{2L}) + ((x)_{\alpha_n}^{2U} - (x)_{\alpha_n}^{2L}) + 2\sum_{i=1}^{n-1}((x)_{\alpha_i}^{2U} - (x)_{\alpha_i}^{2L})] + \frac{1}{6n}\sum_{i=0}^{n-1}((x)_{\alpha_i}^U \bullet (x)_{\alpha_{i+1}}^U - (x)_{\alpha_i}^L \bullet (x)_{\alpha_{i+1}}^L) \tag{30.12}$$

where $\alpha_i = i/n$, $i = 0,1,\ldots n$.

30.2.4 Criticality Sequence

The failure modes are prioritized by the defuzzified centroid values of their respective FRPN in the manner. The higher the defuzzified centroid value, the higher would be the overall risk, and the risk priority.

Table 30.1 Failure modes of door system

Number	Failure mode
FM1	Loose roller
FM2	Wheel wear
FM3	EDCU function is broken
FM4	Limit switch S1 wears out
FM5	Screw poor lubrication
FM6	Nut component wears out

Table 30.2 Fuzzy assessment information for six failure modes

Failure mode	Occurrence (O)	Severity (S)	Detection (D)
FM1	(1.7,2.7,4.1,5.1)	(1.4,2.4,3.6,4.6)	(2.3,3.3,5.0,6.0)
FM2	(1.4,2.4,3.6,4.6)	(2.1,3.1,4.7,5.7)	(2.1,3.1,4.7,5.7)
FM3	(3.6,4.6,6.4,7.4)	(6.4,7.4,8.4,9.2)	(3.8,4.8,6.5,7.5)
FM4	(3.0,4.0,6.0,7.0)	(5.6,6.6,7.7,8.7)	(3.5,4.5,6.3,7.3)
FM5	(1.4,2.4,3.6,4.6)	(2.3,3.3,5.0,6.0)	(3.6,4.6,6.4,7.4)
FM6	(3.5,4.5,6.3,7.3)	(5.4,6.4,7.6,8.6)	(2.7,3.7,5.6,6.6)
Weights	(0.37,0.62,0.83)	(0.56,0.81,0.96)	(0.11,0.36,0.61)

30.3 Case Analysis

According to the statistical analysis of door fault information, six failure modes are selected for criticality analysis, as shown in Table 30.1.

The assessment information of the six failure modes on each risk factor and the risk factor weights were provided by the five team members. Five members were technical department manager (TM1), Chief Engineer (TM2), engineer (TM3), squad leader (TM4), and technician (TM5), respectively. The five team members are assigned the following relative weights: 30, 25, 20, 15, and 10 %.

The five team members' assessment information was provided in Table 30.2. Since the risk ratings and the risk factor weights are all fuzzy numbers, the overall risk of each failure mode will be a fuzzy number either. The fuzzy risk priority numbers are defined by Eq. (30.3).

To calculate the fuzzy risk priority numbers of the six failure modes, the nonlinear programming models (30.6) and (30.7) are used to calculate all α-levels. For the α-levels are set as 0, 0.1, 0.2, …,1.0, the results are shown in Table 30.3.

Since the unit interval [0, 1] is equally divided by eleven α-levels into 11 subintervals, Eqs. (30.11) and (30.12) are therefore used to compute the centroids of the seven FRPNs. The results are shown in the second column of Table 30.4.

From Table 30.4, it can be observed that FM1 is with the least overall risk, whereas FM3 is with the maximum overall risk. FM3 should be given the top risk priority, followed by FM4, FM6, FM5, and FM2, respectively. The result is perfectly consistent with the ranking achieved by our former intuitive analysis.

Table 30.3 α-level sets of the FRPN for the six failure modes

α	Failure mode					
	FM1	FM2	FM3	FM4	FM5	FM6
0	[1.528,5.238]	[1.677,5.469]	[4.297,8.568]	[3.747,8.117]	[1.806,6.117]	[3.651,8.085]
0.1	[1.644,5.114]	[1.795,5.355]	[4.437,8.432]	[3.881,7.981]	[1.945,5.966]	[3.789,7.933]
0.2	[1.759,4.988]	[1.914,5.238]	[4.581,8.298]	[4.015,7.838]	[2.081,5.818]	[3.928,7.784]
0.3	[1.876,4.869]	[2.032,5.119]	[4.721,8.166]	[4.150,7.693]	[2.221,5.675]	[4.067,7.645]
0.4	[1.989,4.745]	[2.149,4.998]	[4.865,8.021]	[4.289,7.561]	[2.356,5.535]	[4.208,7.508]
0.5	[2.102,4.623]	[2.266,4.884]	[5.013,7.893]	[4.428,7.426]	[2.492,5.398]	[4.349,7.374]
0.6	[2.217,4.499]	[2.379,4.764]	[5.155,7.768]	[4.568,7.301]	[2.627,5.254]	[4.486,7.228]
0.7	[2.328,4.384]	[2.494,4.641]	[5.302,7.637]	[4.707,7.171]	[2.762,5.114]	[4.627,7.092]
0.8	[2.440,4.263]	[2.609,4.522]	[5.452,7.516]	[4.850,7.043]	[2.895,4.973]	[4.783,6.959]
0.9	[2.555,4.143]	[2.724,4.406]	[5.601,7.389]	[4.993,6.910]	[3.028,4.831]	[4.928,6.828]
1.0	[2.665,4.023]	[2.835,4.289]	[5.755,7.257]	[5.135,6.787]	[3.158,4.688]	[5.073,6.669]

Table 30.4 The priority ranking of the six failure modes

Failure mode	FRPN	Sequence
FM1	3.366	6
FM2	3.572	5
FM3	6.454	1
FM4	5.933	2
FM5	3.945	4
FM6	5.862	3

30.4 Conclusions

The FMEA method which is based on the weighted least squares model and the fuzzy weighted geometric mean was used to assess metro door fault criticality. The approach overcomes certain disadvantages associated with the traditional RPN based on FMEA. Through FMEA analysis, the failure mode that EDCU function is broken is given the top priority for correction, followed by "Limit switch S1wears out," "Nut component wears out," "Screw poor lubrication," "Wheel wear" and "Loose roller." The result is consistent with the field engineers, and can provide the decision-making reference for the door design improvement and maintenance.

Acknowledgments This research was sponsored by National High-tech R&D Program of China (863 Program, No.2011AA110501) and National Key Technology R&D Program of China (No. 2011BAG01B05). The supports are gratefully acknowledged.

References

1. Zhao Y, Liu F (2009) Application of FMECA in designing railway vehicle parts. Railw Veh 47(2):11–15 (in Chinese)
2. Xu K, Tang TC, Xie M, Ho SL, Zhu ML (2002) Fuzzy assessment of FMEA for engine systems. Reliab Eng Syst Saf 75:17–29
3. Pillay A, Wang J (2003) Modified failure mode and effects analysis using approximate reasoning. Reliab Eng Syst Saf 79:69–85
4. Sharma RK, Kumar D, Kumar P Fuzzy (2008) modeling of system behavior for risk and reliability analysis. Int J Syst Sci 39(6):563–581
5. Liu, H-C, Liu L, Liu N (2013) Risk evaluation approaches in failure mode and effects analysis: a literature review. Expert Syst Appl 40:828–838
6. Zhang Z, Chu X (2011) Risk prioritization in failure mode and effects analysis under uncertainty. Expert Syst Appl 38:206–214
7. Wang Y, Chin K, Gary KKP et al (2009) Risk evaluation in failure mode and effects analysis using fuzzy weighted geometric. Experts Syst Appl 36:1195–1207
8. Chen L, Ko W (2009) Fuzzy linear programming models for new product design using QFD with FMEA. Appl Math Model 33:633–647
9. Gargama H (2011) Sanjay Kumar Chaturvedi. Criticality assessment models for failure mode effects and criticality analysis using fuzzy logic. IEEE Trans Reliab 60(1):102–110

Chapter 31
Risk-Based Maintenance Optimization of Metro Vehicle Door System

Xiang Guo, Lisha Pan, Limin Jia, Yuan Zhang and Zongyi Xing

Abstract Traditional maintenance of door system has not taken the risk of operating failure into consideration. So a risk-based maintenance strategy is developed and applied to maintenance decision of door system. First, the EDCU unit is identified as the key subsystem of door based on the FMEA method. Then, risk assessment of the EDCU is carried out in accordance with cost. Finally, the rule of age reduction and increasing failure rate is applied to establish the model of the maintenance cycle optimization. The experiment result shows that optimal maintenance cycle of the door system is 57 days, which is verified by expert engineers.

Keywords Maintenance cycle optimization · Door system · Risk-based maintenance · Age reduction factor

31.1 Introduction

The door system is an important part of metro vehicle, and it is a complex mechatronic product containing mechanical parts and electrical components. The efficient operation of rail vehicles is closely related to the safety and reliability of

X. Guo · Z. Xing (✉)
School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
e-mail: xingzongyi@163.com

L. Pan
Engineering Technology Research Center, Guangzhou Metro Corporation, Guangzhou 510030, China

L. Jia · Y. Zhang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_31,
© Springer-Verlag Berlin Heidelberg 2014

the door system. According to the research in the domestic subway operating company, the door system of the metro vehicle has a high failure rate and serious fault hazard, which will lead to delays or service disruptions in the operation of the metro.

Major maintenance strategies of mechatronic products include preventive maintenance (PM) strategy, condition-based maintenance (CBM) strategy, reliability-centered maintenance (RCM) strategy, and the developing strategy recently: risk-based maintenance (RBM) [1]. At present, the major domestic maintenance strategy of the door system is PM with a fixed cycle, which may result in excessive maintenance. To solve this problem, some scholars have researched on the door system by using RCM [2, 3], while research on maintenance which considers operational risk caused by door failure has not been found. RBM strategy achieves rapid development in recent years, Yang improves the RBM strategy, and both objective and subjective methods are given to determine the reduction of equipment failure probability to keep subsystem risk under acceptable risk criterion [1]. Dong suggests a frame of RBM. In the frame, the risk which integrates reliability and fault effect is used as maintenance optimization aim, and a fault tree algorithm is used to realize the optimization of maintenance interval [4].

An RBM decision is applied to metro vehicle door system. The risk of door system's failure is evaluated, the risk cost is calculated in consideration of reliability, and a PM cycle optimization model of the door (Electronic Door Controlling Unit) EDCU unit is established to optimize the maintenance cycle.

31.2 Risk-Based Maintenance Procedures

Typical procedures of RBM decision include risk identification, risk assessment, and maintenance decision.

31.2.1 Risk Identification

Risk identification means the scope determination of maintenance decision. First, the system to be analyzed should be confirmed. Second, the system's major failure should be determined. Finally, the system should be divided into several subsystems for further research.

Failure Mode and Effects Analysis (FMEA) is an important technique which is used to identify and eliminate known or potential failures to enhance the reliability and safety of complex systems. FMEA is intended to provide information for making risk management decisions [5]. According to FMEA, analysis of the system's main failure mode is made, and key system unit which is the specific object of the maintenance decision is identified.

31.2.2 Risk Assessment

Risk grade (RF) is generally used to make a comprehensive risk assessment of failure [5]. Three factors are mainly considered: the severity of failure's consequences, failure probability of occurrence, and discovery probability of failure, which are, respectively, measured by the risk grade relies on severity rating (YZ), failure probability level (GL), and detection rate (JC). The risk grade (RF) can be calculated as Eq.(31.1):

$$\mathrm{RF} = \mathrm{YZ} \times \mathrm{GL} \times \mathrm{JC} \tag{31.1}$$

According to the risk level of failure, the risk is judged whether it can be accepted by ALARP (As Low As Reasonably Practicable). ALARP means that the risk can not be eliminated and risk grade should be controlled within an acceptable range. However, the risk grade (RF) is a semiquantitative method, and severity rating (YZ), failure probability level (GL), and detection rate (JC) could be subjective and uncertain.

To solve the problem, risk assessment based on cost is applied. The loss resulted from the equipment's failure is divided into two parts: the direct economic loss L_D and the indirect economic loss L_I [6]. The total loss resulted from the key system's failure is:

$$L = L_D + L_I \tag{31.2}$$

In order to simplify the calculation, the ratio factor γ of the indirect economic loss to the direct economic loss is introduced, and the ratio factor usually is 4:1. The indirect economic loss is:

$$L_I = \gamma * L_D \tag{31.3}$$

The risk of key system's failure should be assessed in two aspects: failure probability and failure consequence. The reliability parameters of the key system are estimated according to the maintenance record. The consequence of failure is quantified from the economic loss caused by failure.

31.2.3 Maintenance Decision

According to the assessment results of the unit's failure risk, if the risk exceeds the acceptable risk level, the maintenance strategy or maintenance cycle of the unit should be adjusted. So that the reliability index of the unit is improved and failure risk of the unit is reduced to an acceptable range.

Periodic PM is adopted for equipments at maintenance site. However, the determination of maintenance cycle lacks reasonable and scientific guidance. It will result in excessive maintenance and increasing maintenance cost. An

assessment of equipment's maintenance cost and risk loss based on reliability is carried out to help optimizing the maintenance cycle, decreasing the maintenance cost and controlling the risk of equipment's failure.

Preventive maintenance cannot restore equipments as new. With the increase of maintenance times and service age, equipment's failure rate is increasing. Periodic PM cannot stop the decrease of equipments' reliability. The rule of age reduction and increasing failure rate are introduced to describe the failure rate of equipments in periodic PM [7].

According to the rule of age reduction, the equipment's failure rate after *i*th PM is:

$$\lambda_{i+1}(t) = \lambda\left(\sum_{k=1}^{i} a_k T_k + t\right) \quad t \in (0, T_{i+1}) \tag{31.4}$$

where, $0 < a_k < 1$ is age reduction factor, T_i is time interval between $(i-1)$th PM and *i*th PM, t is the equipment's operating time after *i*th PM, $\lambda(t)$ is the equipment's failure rate function.

In the rule of increasing failure rate, after *i*th PM, the equipment's failure rate is:

$$\lambda_{i+1}(t) = b_i \lambda_i(t) = \left(\prod_{k=1}^{i} b_k\right)\lambda(t) \tag{31.5}$$

where, $1 < b_1 \le b_2 \le \cdots \le b_i$ is increasing failure rate factor, $\lambda(t)$ is the equipment's failure rate function.

In engineering practice PM not only decreases the equipment's service age, but also changes slope of the equipment's failure rate function. With the two rules, the equipment's failure rate function in PM is [8]:

$$\lambda_{i+1}(t) = \left(\prod_{k=1}^{i} b_k\right)\lambda\left(\sum_{k=1}^{i} a_k T_k + t\right) \quad t \in (0, T_{i+1}) \tag{31.6}$$

where, $a_k = \dfrac{k}{3k+7}$, $b_k = \dfrac{12k+1}{11k+1}$.

A periodic PM of the equipment costs M Yuan, a replacement of the equipment costs F Yuan, the risk indicator of once equipment's failure is L Yuan. In a life cycle of the equipment its maintenance interval number is N. The equipment has $N-1$ PM, maintenance cycle is T and the equipment will be replaced in the last maintenance cycle. An equipment's life cycle maintenance cost is:

$$C = L * \int_0^T \left(\sum_{i=1}^{N} \lambda_i(t)\right) dt + (N-1) * M + F \tag{31.7}$$

where, the first part is risk loss of the equipment's failures in a life cycle, the second part is periodic PM's cost and the third part is replacement's cost.

The optimization model of the equipment's life cycle maintenance cost is:

$$\min c = \frac{L * \int_0^T \left(\sum_{i=1}^{N} \lambda_i(t) \right) dt + (N-1) * M + F}{NT} \tag{31.8}$$

This model could be used to calculate minimum of average life cycle maintenance cost.

31.3 The Application of the Maintenance Based on Risk to the Maintenance Optimization of Door System

The metro door system is one of the important components of the metro vehicle. The reliability of door system has direct impact on the quality of the metro vehicle's operation. RBM is applied to the maintenance optimization of door system, which can guarantee the safety, reliability, and economical efficiency of the door system.

31.3.1 The Maintenance Identification of Door System's Subsystems

Door system is composed of leaves, external bearing oriented equipments, internal driving and locking equipments, interface units, internal operating equipments, external operating equipments, and EDCU.

The FMEA of door system have drawn some conclusions. The criticality sequence of door system's subsystems is: the switch S1, EDCU, the Stopper, the switch S4, the Electromagnetic brakes, and the Drive Motor. The severity of EDCU and the switch S4 is higher than those of the Stopper, the Electromagnetic brakes, and the Drive Motor. The EDCU is identified as one of the key subsystems of door system, so that maintenance decision of EDCU is studied in this paper [9].

31.3.2 The Risk Assessment of the EDCU

The failure modes of EDCU are: Plug Loose, Function Failure, and low Software Version. Plug Loose needs to check the plug online and fasten the plug. Function Failure leads to the consequence that the door cannot switch or switch slowly. What is worse, the door must be cut off when it is serious. Low Software Version leads to that the door cannot switch or the DDU displays abnormal. All these three failure modes result in a decrease of the service quality or a delay of operation.

Table 31.1 EDCU's optimal maintenance cycle

N	1	2	3	4	5	6	7	8	9	10
T(Day)	143	81	57	44	36	30	26	22	20	17
minc(Yuan/Day)	355	329	325	329	338	349	363	378	395	413

The direct economic loss resulted from the EDCU's failure is the maintenance cost, including labor cost, spare parts consumption cost, and logistical costs. The indirect economic loss is the decrease of the service quality or the delay of operation resulted from the EDCU's failure. Generally, the indirect economic loss is invisible, because it does not generate cash outflows.

After field research, the maintenance direct economic cost of EDCU's failure is $L_D = 400$ Yuan. According to the ratio factor γ of the indirect economic loss to the direct economic loss, the indirect economic loss of EDCU's failure is $L_{\mathrm{I}} = 1600$ Yuan. The total economic loss, which is the risk indicator of EDCU's failure, is $L = L_{\mathrm{I}} + L_D = 2000$ Yuan.

31.3.3 Maintenance Cycle Optimization of the EDCU

On the basis of field experience, EDCU's failure is subject to Weibull distribution, its failure rate function is:

$$\lambda(t) = \frac{m}{\eta^m} t^{m-1} \tag{31.9}$$

where, $m = 1.275$, $\eta = 13.67$. A periodic PM of EDCU costs $M = 500$ Yuan, a replacement of EDCU costs F = 11000 Yuan. In a life cycle of an EDCU, the maintenance interval number is N. The EDCU has $N - 1$ PM, each maintenance cycle is T and the EDCU will be replaced in the last maintenance cycle.

There are two decision variables in the above model: N and T. The variable N is a positive integer so that there is one optimization corresponding to each N. Optimization algorithm is applied to calculate an optimal solution with different N. Optimal solution of the model is determined by comparing different optimal solutions with different N.

The results are in Table 31.1. As can be seen from Table 31.1, when $N = 3$, EDCU's minimum of cost per life cycle unit is 325.26 Yuan/Day and maintenance cycle T is 57 Day. With the increase of maintenance interval number N an EDCU's life cycle maintenance cost first decrease then increase.

31.4 Conclusions

Risk-based maintenance is applied to door system of metro vehicles. Operating risk of door system's failure is taken into consideration. The rule of age reduction and increasing failure rate are introduced to describe the reliability model of EDCU in periodic PM. The model of the periodic PM cycle optimization is established based on the principle that EDCU's life cycle maintenance cost should be minimal. An example from engineering practice is given with the optimization model, the result shows that the minimal EDCU's life cycle cost is 325 Yuan/Day and optimal maintenance cycle T is 57 Days.

Acknowledgments This research is supported by National High-tech R&D program of China (863 Program, No.2011AA110501) and National Technology R&D Program of China (No. 2011BAG01B05). The supports are gratefully acknowledged.

References

1. Yang Q, Hu H, Wang M (2013) Inspection/maintenance optimization for petrochemical system by risk based maintenance strategy. J Xi'an Jiaotong Univ 47(6):30–136 (in Chinese)
2. Zeng C, Su J (2011) Application of reliability-centered maintenance of door system of metro vehicles. Electric Drive Locomotives 11:61–63 (in Chinese)
3. Mao Y (2008) Planning and implementation of RCM of Nanjing metro equipment. Master Dissertation, Nanjing University of Science and Technology, Nanjin, 5 (in Chinese)
4. Dong Y, Wang X, Gu Y (2008) Risk-based maintenance decision and its application to power plant equipment. Mod Electr Power 25(2):52–56 (in Chinese)
5. Dong X (2009) Reliability, availability, maintainability and safety of rail train (RAMS). China Railway Publishing House, Beijing (in Chinese)
6. Zhang H, Wang E, Liu Z (2010) Research on prediction and risk management of mine disaster accident loss. Coal Eng 1:120–123 (in Chinese)
7. Zhou X, Xi L, Lee J (2005) A reliability-based sequential preventive maintenance model. J Shanghai Jiaotong Univ 39(12):2044–2047 (in Chinese)
8. Wang L (2011) Research on reliability-centered maintenance decision and support system for high speed train equipment. Beijing Jiaotong University Doctoral Dissertation, Beijing, pp 54–57 (in Chinese)
9. Liu J (2010) FMECA analysis and application research on door system of Shanghai metro line 5. Undergr Eng Tunnels 4:39–42 (in Chinese)

Chapter 32
Using Asynchronous Hot Standby Spare in Time-Stamped, Fault-Tolerant, Real-Time System

Mahmoud Jannesari Ladani and Ahmad Kezemi Gazanchaei

Abstract Failure in a real-time system has different definition from failure in other general purpose systems so that in addition to functional failures may have some time-dependent failures. In a real-time system an output must be produced after threshold and before deadline. So missing the deadline is a kind of failure. In [3] by using artificial neural networks (ANN), a special TMR model is proposed to provide fault tolerance in real-time systems. Inclusion of time for decision mechanism in this model makes it useful for real-time systems. According to the mentioned model, system continues to operate until all nodes fail; but since missing the deadline is considered as a failure in a real-time system and there is probability of missing the deadline by all of three nodes simultaneously, we have a scenario in which system will fail just because of MISSING the Deadline. In this paper, a more comprehensive Method is proposed by which a real-time system is capable to tolerate such a scenario. This model is derived from the proposed model in [3] by adding an asynchronous hot standby spare for covering time-dependent faults in a real-time system.

Keywords Real time · Fault tolerant · Time stamped · Standby spare

32.1 Previous Works

In [1] author discusses the key issues in the design of fault-tolerant real-time systems and presents a systematic methodology for the design of an application if the target architecture is time triggered.

M. J. Ladani (✉) · A. K. Gazanchaei
Department of Electrical Engineering, Pooyesh University , Qom, Iran
e-mail: jannesari.2006@gmail.com

A. K. Gazanchaei
e-mail: ahmad.kazemi20@gmail.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_32,
© Springer-Verlag Berlin Heidelberg 2014

In [2] the real-time recoverable action (RTR-Action) as a programming structure is designed to provide fault tolerance for real-time systems. By using an adaptable scheduler, timing errors are detected.

In [3] a special kind of hybrid (software, hardware) redundancy is proposed by using ANN. In my opinion, the proposed model in [3] is a smartly designed model to provide fault tolerance for real-time systems. So I decided to extend this concept, so proposed model can tolerate more failure scenarios.

32.2 Terminologies

32.2.1 Process Time

The time will be taken by a node's processor from getting the environment data to producing corresponding result out of it. We will call it PT.

32.2.2 Hot Standby Spare Latency

The latency by which the hot standby spare node starts to operation. Let us call it HSL.

32.3 Introduction

In a real-time system, generally, the system is processing some input data from external world and producing result periodically. This means that a real-time system is subjected to serve proper output at regular points of time. So whenever we need, system's output must be been prepared. If the desired output is not ready, system fails due to timing issues.

In Fig. 32.1, needs for the system's output are determined on output need line. In case (F), you can see the result of Deadline Missing. Note that increasing the PT is not the only reason for missing deadline, but also it may be caused by some other issues. Assuming that in case (F), instead of needed output, we use the last recently produced output by real-time system; thus if we are currently at time T3, then we have to perform our further processing over the output corresponding to T2. For E.g. when missing a video frame while streaming it over network, the receiver senses video freeze due to this deadline miss. Now assume that we have an output produced at time $T2 + \Delta t$; in such a case, the result of processing this output will be more acceptable than the result of processing T2′s output since $T2 + \Delta t$ is nearer to T3 than T2.

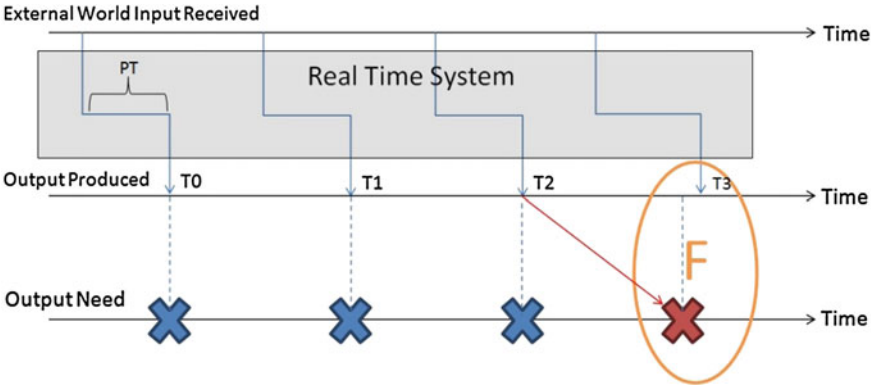


Fig. 32.1 Periodic result production in real-time system

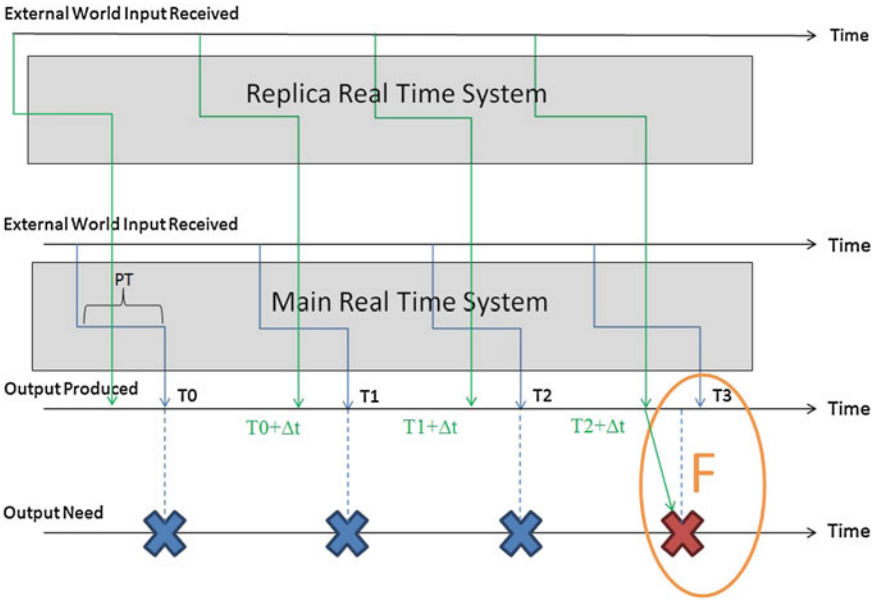


Fig. 32.2 Periodic result production in real-time system with the hot standby spare

32.4 Proposed Method

In this paper, a method is introduced by which, in case of time-dependent system failures, the system is capable of tolerating the issue.

Consider that we add a real-time system as a hardware replica to the design demonstrated in Fig. 32.1. Since the new replica starts before the main real-time system, it produces outputs between the regular time periods. Now you can compare condition (F) in Figs. 32.1 and 32.2.

The output used for forth need is very close to T3.

Note that the main real-time system can be designed using the model proposed in [3].

32.5 Conclusion

By adding the hot standby spare replica to the model proposed in [3], we introduced a method that helps the real-time system to tolerate the time-dependent failure conditions.

Because of time limitation, I could not perform all analyses on the proposed method or simulate it and could not explain clearly the relationship between TMR members and newly added replica.

As a future work I am going to complete the introduction of this method.

References

1. Kopets H (2009) The design of fault-tolerance real-time systems. Vienna
2. Malik S, Rehman MJ (2005) Time stamped fault tolerance in real time systems
3. Moran CE (1996) Designing adaptable real-time fault-tolerant parallel systems. Sao Carlos

Chapter 33
Formal Modeling and Analysis of Radio Block Center Handover

Xiaohui Hu, Jun Ma, Yong Chen and Xin Li

Abstract Radio Blocking Center (RBC) is one of the main factors that affect the control system of CTCS-4 level series of trains. An in-depth analysis has been about the differences of RBC handover process between one onboard mobile terminal and two onboard mobile terminals. Based on this particular analysis, a timed CPN formalizing model has been set by using CPN tools, which are the supporting tools of Colored Petri Net. Also emulation and verification have been made on this model. Through the comparison and analysis between these two handover modes, people can get a more profound reorganization on how to improve the RBC handover equipments and how to perfect the RBC handover agreements.

Keywords Radio blocking center (RBC) · Formal modeling · RBC handover

33.1 Introduction

Radio Blocking Center (RBC) is the core equipment of operation system of the CTCS-4 level train. According to the features of the track (e.g., gradient and speed limit), track occupancy, and interlock route status, it will send the "speed-distance" monitoring information to the vehicle equipment through radio communication system, to ensure the safe operation of trains in the region it monitors. As the monitoring region of one RBC is limited, there comes a switch between one RBC and another when the train is in operation. That is the so-called RBC handover. In order to eliminate the influence to the train, two independent vehicle stations are needed on the vehicle equipment. In this paper, an in-depth analysis

X. Hu (✉) · J. Ma · Y. Chen · X. Li
School of Electronic and Information Engineering, Lanzhou Jiaotong University, No.88 Anning West Road, Lanzhou, Gansu Province, China
e-mail: 928720752@qq.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_33,
© Springer-Verlag Berlin Heidelberg 2014

has been demonstrated about the specific processes and implementation methods of two vehicle stations and one vehicle station, when the train is running across the border of two RBC monitoring regions.

33.2 RBC Handover Management

When the train is running from one RBC monitoring region into another RBC monitoring region, the two RBC will handover the control rights of the train. The handover RBC will send switching notice information (e.g., vehicle equipment ID, transponder unit ID in the RBC region border, train data etc.), access requirement information, switching notice information, switching conforming information, and switching canceling information to the takeover RBC. The takeover RBC will send access information and information of the takeover train to the handover RBC [1]. In normal cases, it would not need the intervention from the diver and will not lead to braking. When the handover RBC has detected the train is approaching the switching region border, it will send train access information to the takeover RBC, after getting the necessary ground information of extending MA, through the communication with the takeover RBC, to achieve the seamless switching between two RBC monitoring regions [2].

The RBC handover includes the handover based on one vehicle station and the handover based on two vehicle stations. In order to improve the efficiency of the train operation, the switching mode based on two vehicle stations should be used. However, the switching mode based on one vehicle station should be used, when one of the vehicle stations breaks down. If both of the vehicle stations break down, it should downgrade to cross the RBC monitoring region border.

33.3 Formalized Modeling CPN

33.3.1 Colored Petri Nets

Colored Petri nets is a formalized method to simulate the complicated system and evaluate the performance. Compared with the classic Petri nets, the Colored Petri nets has three new conceptions: color, time, and hierarchy. This gives Petri nests a much stronger describing ability, which can model the dynamic complicated system [3].

33.3.2 Timed Colored Petri Nets [4]

When time extension is used, the normal method is that the mark will take timestamp as a value. This timestamp shows when the mark is available. For example, a mark with a t value timestamp can only be changed and consumed after time t. The precondition for change is when and only when the timestamp of the will-be consumed mark is earlier than current time. In other words, the ready time for change is the earliest time that can be used as mark in its input storage. The marks are consumed in first input first out rule. The mark with the earliest timestamp will be consumed first. Besides, the change with the earliest ready time will be implemented first. If one change is implemented and creates a mark, the mark will be given a timestamp which is later than the implemented timestamp. With the implementation of the change, the delay time of the creating of the mark can be fixed. The timestamp of the created mark equals to the implementation time plus delay time. The length of the delay time depends on the mark value that it consumes. The time delay could be a fixed value, e.g., 10 s, also could be a random value among a certain region. According to the fixed value and region value, timed Petri nets can be divided into fixed delay time nets and non-fixed delay time nets.

33.4 RBC Handover Simulation Model

CPN Tools is a powerful tool to establish, simulate, and analyze CPN model [5] which supports status space [6] analysis, time simulation, function analysis, and performance analysis [7] of the model. The users can compile their own program block in the simulation process to get the data needed in performance evaluating to get a more precise simulation model, according to their own needs. The feedback function can conveniently provide error information of context and the dependent relationships among the elements in the model. This tool also provides grammar detecting function and program block compiling function in the established model. This means to simulate and analyze the right part in the model when the whole model has not been finished yet or to neglect some errors becomes possible. Also, we can adjust the structure of the model in the simulation process to check the performance and functional changes. Meanwhile, this tool can also deal with timed and non-timed nets structure, and will create all or part status space, then to analyze the space automatically, and output a status space analysis report, such as boundedness and activity. Through some simple inquiry words, we can monitor some special features in the system.

In this paper, we use the simulation tool, CPN Tools, to model the RBC handover process based on two vehicle stations. First, we will establish a brief model to describe the switching process. Then, we will extend and detail the brief model, adding some changes, e.g., the handover RBC, the vehicle equipment, and

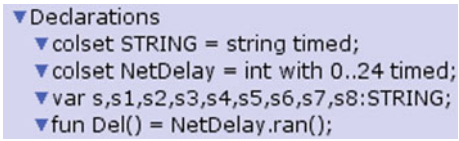


Fig. 33.1 Time simulation model of RBC handover based on two vehicle stations variable declaration

the takeover RBC to describe the information transmission in the switching process (Figs. 33.1 and 33.2).

33.4.1 Time Simulation Model of RBC Handover Based on Two Vehicle Stations

This model is of the RBC switching based on two vehicle stations. This model shows that the handover RBC will establish communication initiatively with the takeover RBC after sending the handover notice. After the communication is successfully established, the vehicle equipment will send location information to the two RBC simultaneously. Meanwhile, the handover RBC will ask for access information from the takeover RBC, to extend the vehicle MA to the region of the takeover RBC. Then, when the maximum safe point of the train has passed the border transponder unit, both of the two RBCs will receive the location information of the train. The handover RBC will give feedback to the takeover RBC after receiving the location information. The takeover RBC will send takeover confirmation after receiving the feedback. The handover RBC will send communication terminating order to the train after the minimum safe end point has passed border transponder unit. After the communication between the train and the handover RBC is over, RBC switching process is completed (Figs. 33.3 and 33.4).

33.4.2 Time Simulation Model of RBC Handover Based on One Vehicle Station

This model is of the RBC handover based on one vehicle station. This model shows that the vehicle equipment can establish communication with the takeover RBC only after the communication between the vehicle equipment and the handover RBC has finished, as the train only has one vehicle station and cannot communicate with two RBCs simultaneously. Before this, the handover RBC will forward the location information of the train to the takeover RBC. The takeover RBC will formally takeover the train to complete the whole takeover process, after it has established communication with the vehicle equipment and received the location information of the train.

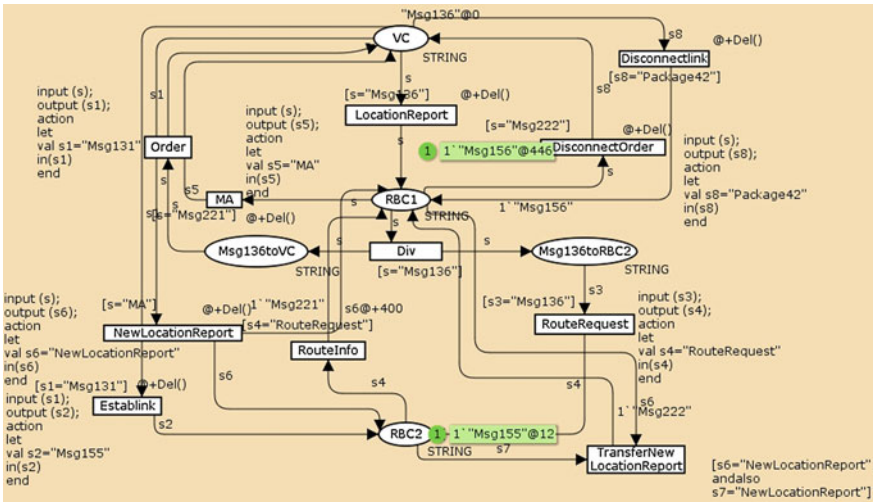


Fig. 33.2 Time simulation model of RBC handover based on two vehicle stations

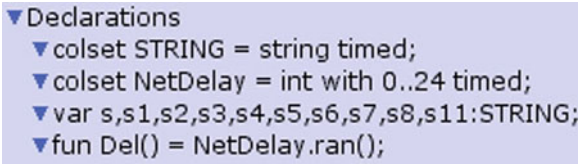


Fig. 33.3 Time simulation model of RBC handover based on one vehicle station variable declaration

33.4.3 Comparison and Analysis Between the Two RBC Handover Models

In order to get the time distribution of the two RBC handover, we use "time simulation model of RBC Handover based on two vehicle stations" and "time simulation model of RBC switching based on one vehicle station" to simulate 10,000 times switching processes and get 10,000 RBC switching time, respectively. Then, we sort out these data and use matlab to protract the switching time distribution figure, as Fig. 33.5.

For the "time simulation model of RBC Handover based on two vehicle stations", we use 0.1 s as the unit time. The time consumed in the processes of the vehicle equipment sending location information to the handover RBC, the handover RBC sending switching order to the vehicle equipment, the vehicle equipment establishing communication with the takeover RBC, and the handover RBC sending access permit to vehicle equipment are all random figures (0–24 unit time). The time consumed by the processes of the vehicle equipment sending train data to the handover RBC and takeover RBC, respectively, and receiving the train data to confirm is 40 s, 400 unit time to describe in the graph. The time consumed

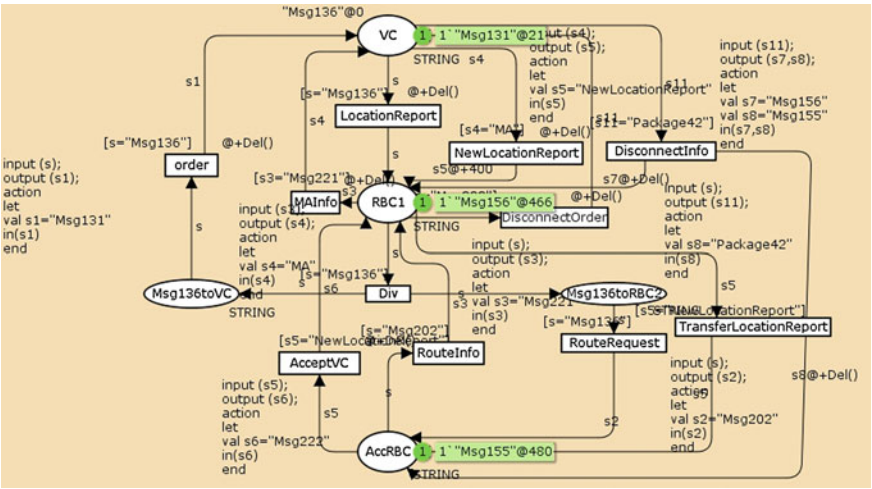


Fig. 33.4 Time simulation model of RBC handover based on one vehicle station

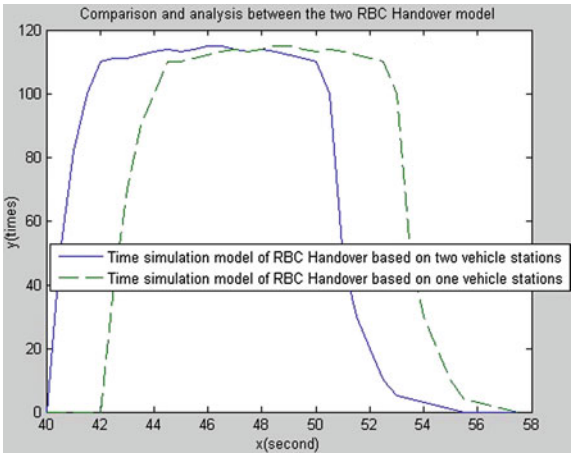


Fig. 33.5 Comparison and analysis between the two RBC handover models

in the processes of the handover RBC sending terminating order to the vehicle equipment and the vehicle equipment sending terminating information to the handover RBC are all random figures too. The total time is just the switching time based on two vehicle stations.

For the "time simulation model of switching based on one vehicle station", we also use 0.1 s as the unit time. The time consumed in the processes of the vehicle equipment sending location information to the handover RBC, the handover RBC sending switching order to the vehicle equipment, and the handover RBC sending access permit to the vehicle equipment are all random figures. The time consumed by the processes of the vehicle equipment sending train data to the handover RBC

and receiving train data is 40 s, 400 unit time to describe in the graph. The time consumed in the processes of the takeover RBC sending train data to the handover RBC, the handover RBC sending terminating order to the vehicle equipment, the vehicle equipment sending terminating information to the handover RBC, and the vehicle equipment establishing communication with takeover RBC are all random figures too. The total time is just the switching time based on one vehicle station.

Form the figure, we can see that the RBC switching time based on two vehicle stations is between 41 and 51 s, and the RBC switching time based on one vehicle station is between 45 and 54 s. The reason mainly lies in that vehicle equipment based on two vehicle stations can use the other station to establish communication with RBC2, while one station is keeping communication with RBC1; however, the vehicle equipment based on one vehicle station can only establish communication with RBC2 after the communication with RBC1 is terminated. Thus, the RBC switching time based on one vehicle station is a little more than the RBC switching time of two vehicle stations.

33.5 Conclusion

In this paper, we use the modeling tool, CPN Tools, to model the RBC switching based on one vehicle station and two vehicle stations. We get the conclusion as follows through the analysis of the time consumed in the two CPN models in a certain track condition.

The RBC switching time based on two vehicle stations is shorter than the time based on one vehicle station. In other words, the RBC switching based on two vehicle stations can improve the operation efficiency of the train. CTCS-4 level has two vehicle stations in the switching, using redundancy and backup way. The use of CTCS-4 level should coordinate with the lower level. When the situation can achieve CTCS-4 level, it will use CTCS-4 level. When the situation cannot achieve CTCS-4 level, it will use CTCS-3 level. From the reliability point of view, this can not only ensure the safe operation of the train, but also can ensure the train to be operated at a certain speed.

Acknowledgments This work was financially supported by the National Natural Science Foundation (61163009) and by the Lanzhou University Qing-lan works projects.

References

1. Guo Y-F, Liu J-D, Zhao J-G (2009) Radio block center (RBC) handover problem analysis. Railw signal commun 45(3):41–42 (in Chinese)
2. Liu Z-T, Sun W-l (2011) Research on functional requirements for radio block center in CTCS-3. Railw Comput Appl 20(6):5 (in Chinese)
3. Roubtsova EE (2004) Property specification for coloured petri nets. Syst Man Cybern 3:10–13

4. Wang Y-N, Wang M-Z (2005) Timed colored petri nets in modeling and simulation. Measur Control Tech 24(1):50 (in Chinese)
5. Jensen K (1998) A brief introduction to colored petri nets. In: Workshop on the applicability of formal models
6. Jensen K, Christensen S, Kristensen LM (2002) CPN tools occurrence graph manual. Version0.1 [EB/OL], University of Aarhus
7. Jensen K (1991) Coloured petei nets. A high level language for system design and analysis. Advances in Petri Nets. Springer, Berlin, pp 342–416

Chapter 34
Fault Diagnosis for Rail Vehicle Suspension Systems Based on Fisher Discriminant Analysis

Xiukun Wei, Sheng Wu, Jianlong Ding, Limin Jia, Qu Sun and Minzhen Yuan

Abstract In this paper, fisher discriminant analysis (FDA) is used for fault isolation and diagnosis in rail vehicle suspension systems. The suspension systems are equipped with acceleration sensors in the corners of the car body and the two bogies. The faults considered are the lateral damper faults and the lateral spring faults in suspension systems. FDA provides an optimal projection space on the basis of the training data including the fault data and normal data to classify the test data. A vehicle model is built by SIMPACK/MATLAB software with real parameters to obtain the simulation data and the effectiveness of the proposed method is demonstrated by simulation.

Keywords Rail vehicle suspension systems · Fault diagnosis · Fisher discriminant analysis · Fault isolation

34.1 Introduction

Suspension systems for rail vehicles are to support the car body and bogie, to isolate the forces generated by the track unevenness at the wheels, and to control the attitude of the car body with respect to the track surface for improving ride

X. Wei (✉) · L. Jia
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
e-mail: wwxxkk@gmail.com

S. Wu (✉) · Q. Sun
School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China
e-mail: wusheng6250@sina.com

J. Ding · M. Yuan
Guangzhou Metro Company, Guangzhou 510030, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_34,
© Springer-Verlag Berlin Heidelberg 2014

comfort [1]. So, the reliability of the suspension systems is directly related to the rail vehicles' safety. As is known to all, some small but key components of suspension systems, such as springs and dampers, the performance significantly degrades in performance after several years. Once faults occur in the rail vehicle's suspension system, which are even very small and insignificant, they may lead to serious accidents while the train is moving. Therefore, it is imperative to immediately detect and diagnose the faults of rail vehicle's suspension systems.

Fault diagnosis for rail vehicle suspension systems has received many attentions in recent years. Some studies on the condition monitoring of rail vehicle suspension systems have been concerned in [2–4, 6] and the references therein. In [5], the fault detection and isolation issue of the rail vehicle's lateral suspension systems is discussed, and the method proposed is efficient and has rapid response to the abrupt faults. However, this paper assumed that the vehicle parameters should be known precisely and the simulation is still based on a simple linear model. In [7], the authors proposed a fault detection method for the rail vehicle's suspension systems based on Kalman filter. The fault isolation is realized by using similarity measurements. However, it is unable to diagnose different faults effectively. In [8], a data-driven fault detection method of rail vehicle suspension systems was provided. The dynamical principle component analysis (DPCA) and canonical variate analysis (CVA)-based approaches are applied for the fault detection problems. In general, among these papers, the methods of determining the existence of faults have been considered. However, there are few papers pointing on fault diagnosis of rail vehicle suspension systems.

In rail vehicle suspension systems, when a fault occurs, the fault feature information can be obtained from the sensors equipped on the rail vehicles. However, information from a single sensor is incomplete, uncertain, and imprecise. So it is very necessary to find an approach to extract features from the information.

In this paper, fault diagnosis issue of rail vehicle suspension systems based on FDA is investigated. The considered faults are mainly the lateral damper faults and the lateral spring faults of the suspension systems. First, acceleration sensors collect data including the fault data selected for this paper. Then, the time domain and the frequency domain features are extracted from the connected data as the training data matrix. Finally, a projection space is obtained by applying FDA and the diagnosis of the test data is completed by using the criteria of Mahalanobis distance. The feasibility and effectiveness of the proposed method are shown by simulation for the considered rail vehicle suspension systems.

34.2 Rail Vehicle Suspension Systems

The subway served in Line 1 and Line 8 in Guangzhou, manufactured by CSR Zhuzhou Electric Locomotive Co. Ltd is used for the study in this paper. To build a more accurate model of the suspension systems, three degrees of freedom (yaw, shake, and roll) are considered for both car body and bogies. The lateral

suspension system of the subway vehicle is depicted in Fig. 34. 1. The equations describing the dynamic behavior of the rail vehicle are based on Newton's laws of motion. The dynamical equations of the suspension system for a rail vehicle moving on a straight track are derived as follows:

First, for the car body, the three degrees of freedom (DOFs) equations are described as

$$\begin{aligned} M\ddot{y} = &-2C_2_y\dot{y} - 2C_2_yh_2\dot{\theta} - 2K_2_yy - 2K_2_yh_1\theta + C_2_y\dot{y}_{FB} - C_2_yh_5\dot{\theta}_{FB} + K_2_yy_{FB} - K_2_yh_3\theta_{FB} \\ &+ C_2_y\dot{y}_{RB} - C_2_yh_5\dot{\theta}_{RB} + K_2_yy_{RB} - K_2_yh_3\theta_{RB} \end{aligned} \tag{34.1}$$

$$\begin{aligned} J_\psi\ddot{\psi} = &-2C_2_yl^2\dot{\psi} - 2K_2_yl^2\psi - C_2_yl\dot{y}_{FB} + C_2_ylh_5\dot{\theta}_{FB} - K_2_yly_{FB} + K_2_ylh_3\theta_{FB} + C_2_yl\dot{y}_{RB} \\ &- C_2_ylh_5\dot{\theta}_{RB} + K_2_yly_{RB} - K_2_ylh_3\theta_{RB} \end{aligned} \tag{34.2}$$

$$\begin{aligned} J_\theta\ddot{\theta} = &-2C_2_yh_2\dot{y} - 2(C_2_yh_2^2 + C_2_zb^2)\dot{\theta} - 2K_2_yh_1y - 2(K_2_yh_1^2 + K_2_zb^2 + K_\theta)\theta \\ &+ C_2_yh_2\dot{y}_{FB} + K_2_yh_1y_{FB} + (C_2_zb^2 - C_2_yh_2h_5)\dot{\theta}_{FB} - (K_2_yh_1h_3 + K_\theta - K_2_zb^2)\theta_{FB} \\ &+ C_2_yh_2\dot{y}_{RB} + (C_2_zb^2 - C_2_yh_2h_5)\dot{\theta}_{RB} + K_2_yh_1y_{RB} - (K_2_yh_1h_3 + K_\theta - K_2_zb^2)\theta_{RB} \end{aligned} \tag{34.3}$$

where y, y_{FB}, and y_{RB} are the lateral displacement of the car body, the front bogie, and the tail bogie, respectively. Letter ψ represents the pitch angle of the center of gravity, while θ is the roll angle of the center of gravity for the masses, and their subscripts have the same meaning with the subscripts of y.

For the bogies, here we take the front bogie for an example, the three DOFs equations are described as

$$\begin{aligned} M_B\ddot{y}_{FB} = &-2C_2_y\dot{y}_{FB} + C_2_yh_5\dot{\theta}_{FB} - (2K_1_y + K_2_y)y_{FB} - (2K_1_yh_4 - K_2_yh_3)\theta_{FB} + C_2_y\dot{y} \\ &- K_2_yl\psi + K_1_yy_{W1} + K_2_yh_1\theta + K_1_yy_{W2} - C_2_yl\dot{\psi} + K_2_yy + C_2_yh_2\dot{\theta} \end{aligned} \tag{34.4}$$

$$\begin{aligned} J_{B\psi}\ddot{\psi}_{FB} = &-C_2_\psi\dot{\psi}_{FB} - 2K_1_yl_1^2\psi_{FB} - 2K_1_\psi\psi_{FB} - K_1_yl_1y_{W1} \\ &+ K_1_\psi\psi_{W1} + K_1_yl_1y_{W2} - K_1_\psi\psi_{W2} \end{aligned} \tag{34.5}$$

$$\begin{aligned} J_{B\theta}\ddot{\theta}_{FB} = &-(C_2_yh_5^2 + C_2_zb^2 + 2C_1_zb_1^2)\dot{\theta}_{FB} + K_2_yh_3y_{FB} - (2K_1_yh_4^2 + K_2_yh_3^2 + K_2_zb^2 + 2K_1_zb_1^2 + K_\theta)\theta_{FB} \\ &- C_2_yh_5\dot{y} - C_2_yh_5\dot{y}_{FB} + (C_2_zb^2 - C_2_yh_2h_5)\dot{\theta} - K_2_yh_3y + K_1_yh_4y_{W1} - 2K_1_yh_4y_{FB} + K_1_yh_4y_{W2} \\ &- (K_2_yh_1h_3 + K_\theta - K_2_zb^2)\theta \end{aligned} \tag{34.6}$$

where y_{W1} represents the lateral displacement of the right wheel in the front wheel set under the front bogie, while y_{W2} means the lateral displacement of the left wheel in the tail wheel set under the front bogie. The meanings of the other symbols are similar to these. In the same way, the model of the tail bogie can be constructed.

The state space form of the rail vehicle dynamical model can be derived as

$$\dot{x} = Ax + Bu \tag{34.7}$$

$$y = Cx + Du \tag{34.8}$$

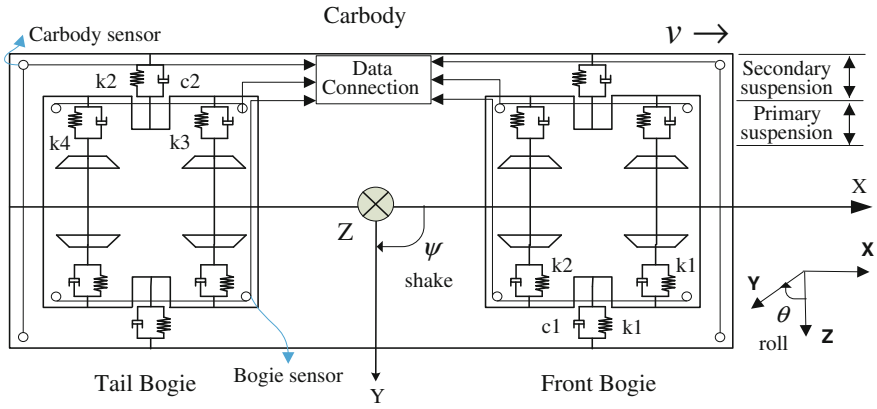


Fig. 34.1 *Top* view of the rail vehicle's lateral suspension system

where

$$x = [\dot{y} \quad \dot{\psi} \quad \dot{\theta} \quad y \quad \psi \quad \theta \quad \dot{y}_{FB} \quad \dot{\psi}_{FB} \quad \dot{\theta}_{FB} \quad y_{FB} \quad \psi_{FB} \quad \theta_{FB}$$
$$\dot{y}_{RB} \quad \dot{\psi}_{RB} \quad \dot{\theta}_{RB} \quad y_{RB} \quad \psi_{RB} \quad \theta_{RB} \quad \dot{y}_{W1} \quad \dot{\psi}_{W1} \quad y_{W1} \quad \psi_{W1} \quad \dot{y}_{W2} \quad \dot{\psi}_{W2}$$
$$y_{W2} \quad \psi_{W2} \quad \dot{y}_{W3} \quad \dot{\psi}_{W3} \quad y_{W3} \quad \psi_{W3} \quad \dot{y}_{W4} \quad \dot{\psi}_{W4} \quad y_{W4} \quad \psi_{W4}]^T$$

$$u = [y_{t1} \quad y_{t2} \quad y_{t3} \quad y_{t4}]^T$$

$$y = [\ddot{y} \quad \ddot{\psi} \quad \ddot{\theta} \quad y \quad \ddot{\psi} \quad \theta \quad \ddot{y}_{FB} \quad \ddot{\psi}_{FB} \quad \ddot{\theta}_{FB}$$
$$y_{FB} \quad \psi_{FB} \quad \theta_{FB} \quad \ddot{y}_{RB} \quad \ddot{\psi}_{RB} \quad \ddot{\theta}_{RB} \quad y_{RB} \quad \psi_{RB} \quad \theta_{RB}]$$

where the matrixes A, B, C, and D can be derived from the above equations. u is the lateral track velocity and displacement due to track lateral irregularities.

As the space is limited, the other parameters and the more details of the state space can be found in [9]. The problem considered in this paper is how to classify different faults in damping and spring of the lateral rail vehicle suspension system.

34.3 Fault Diagnosis

34.3.1 Fisher Discriminant Analysis

FDA is a linear dimensionality reduction technique widely used in the field of pattern classification [10]. The aim of FDA is to find the Fisher optimal discriminant vectors such that the Fisher criterion function is maximized [11]. Then, let the higher dimensional feature space project onto the already obtained optimal discriminant vectors to construct a low dimensional feature space. The different

class data can be separated in a lower dimensional Fisher space mostly. A brief mathematical description follows.

Suppose the training data stacked into a matrix $X \in R^{n\times m}$, where n is the number of observations and m is the number of the measurement variables. Define k as the number of class, j as the number of the class that an observation belongs to, and n_j as the number of rows in the jth class. After representing the ith row of X with column vector x_i, we can get the total scatter matrix

$$S_t = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T \tag{34.9}$$

where $\bar{x}$ represents the total mean vector whose elements correspond to the means of the columns of X, and its exact formula is

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{34.10}$$

We can define matrix X_j as the set of vectors x_i of the class j. Then, the within-scatter matrix for class j can be given as follows:

$$S_j = \sum_{x_i \in X_j} (x_i - \bar{x}_j)(x_i - \bar{x}_j)^T \tag{34.11}$$

where $\bar{x}_j$ is the mean vector of the class j, and its exact formula is

$$\bar{x}_j = \frac{1}{n_j}\sum_{x_i \in X_j} x_i. \tag{34.12}$$

Then, the within-class scatter matrix is:

$$S_w = \sum_{j=1}^{k} S_j \tag{34.13}$$

and the between-class scatter matrix is given by

$$S_b = \sum_{j=1}^{k} n_j(\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})^T. \tag{34.14}$$

The total scatter matrix can be concluded to the sum of the between-scatter matrix and the within-scatter matrix in the following:

$$S_t = S_b + S_w. \tag{34.15}$$

The next thing is to maximize the scatter between classes; in the mean time, minimize the scatter within classes. We can construct a criterion formula as follows, and the FDA method is to find a solution to maximize the value of the formula,

$$J(w) = \frac{w^T S_b w}{w^T S_w w} \tag{34.16}$$

where w is a FDA vector which can maximize $J(w)$. Because the direction and the magnitude of w are variable, the norm is usually set to be $||w|| = 1$. We can find $\min((k-1), m)$ vectors at most. If the first FDA vector w_1 cannot be enough to classify all classes, we can find the second FDA vector w_2 for help, and so on. The vectors are equal to the eigenvectors ϕ of the following equation:

$$S_b \phi = \lambda S_w \phi. \tag{34.17}$$

When matrix S_w is nonsingular, the Eq. 34.17 can convert the generalized eigenvalue problem to the following form:

$$S_w^{-1} S_b \phi = \lambda \phi \tag{34.18}$$

where λ is the eigenvalue which can indicate the degree of the overall separability among the classes.

34.3.2 Fault Diagnosis Based on FDA

The FDA method simultaneously uses all the data points to obtain a single lower dimensional model, used to diagnose the faults, that is to say, the FDA technique is applied to all the classes simultaneously [13]. In paper [11–13], the FDA method had been used in the process industry for the purpose of fault detection and diagnosis. In this paper, this method is studied for fault diagnosis of the rail vehicle suspension systems. The standard FDA-based fault diagnosis consists of three steps and the detail procedure is given in Fig. 34.2.

34.3.2.1 Fault Feature Extraction

In this paper, the data collected includes the training data and test data. Suppose an original training matrix $X \in R^{n \times m}$ consisting of n observations and m measurement variables. Because the scale of the data is huge, we should do some transformations to the matrix. The statistical information of time-base data and frequency-base data are used for obtaining the feature information from the training data. The moment coefficients of time-waveform data are defined as

$$m_n = \frac{1}{N} \sum_{i=1}^{N} x_i^n \tag{34.19}$$

where x_i is the *ith* time historical data and N represents the number of data.

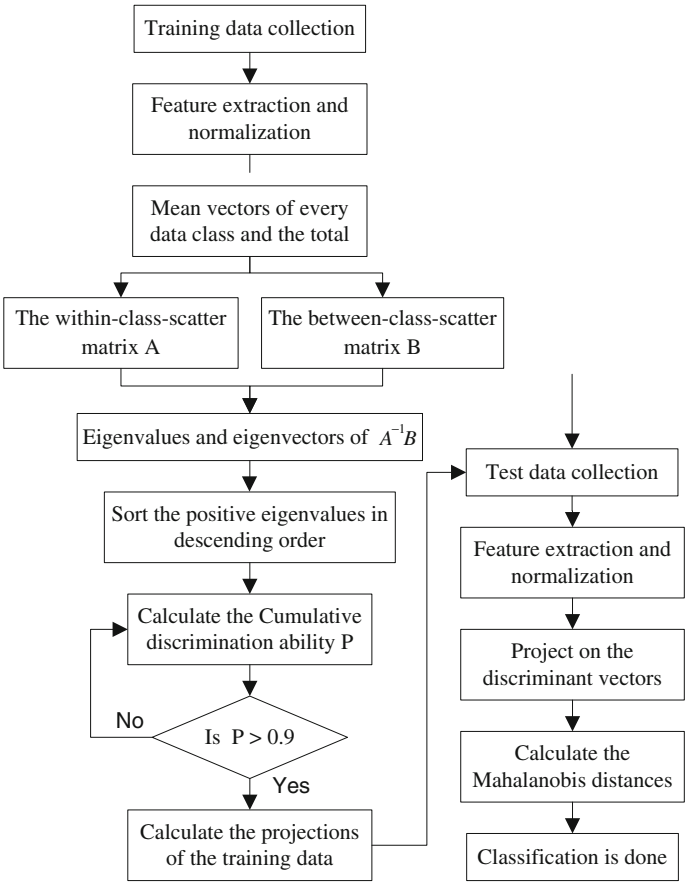


Fig. 34.2 Fault diagnosis flowchart based on FDA

The four features in the time domain, mean C_1, standard deviation C_2, skewness C_3, and kurtosis C_4, can be computed from Eq. (34.19) using the following relationships:

$$\begin{aligned} &C_1 = m_1, C_2 = m_2 - m_1^2, C_3 = m_3 - 3m_2m_1 + 2m_1^3 \\ &C_4 = m_4 - 3m_2^2 - 4m_3m_1 + 12m_2m_1^2 - 6m_1^4 \end{aligned} . \tag{34.20}$$

The frequency domain is another description of a signal and it can reveal some information that cannot be found in time domain [14]. In this paper, frequency center FC, root mean square frequency $RMSF$, and root variance frequency RVF are

$$FC = \frac{\int_0^{+\infty} fs(f)df}{\int_0^{+\infty} s(f)df}, RMSF = \left[\frac{\int_0^{+\infty} f^2 s(f)df}{\int_0^{+\infty} s(f)df}\right]^{1/2}$$
$$RVF = \left[\frac{\int_0^{+\infty} (f - FC)^2 s(f)df}{\int_0^{+\infty} s(f)df}\right]^{1/2} \tag{34.21}$$

where $s(f)$ represents the signal power spectrum. FC and $RMSF$ show the position change of main frequencies, RVF describes the convergence of the spectrum power [14]. The more detail introduction of data fusion and feature extraction can be found in paper [14].

After obtaining the feature information of the variable, we can get a new training matrix Y, which is then scaled to zero mean and to unit variance. The next step is using FDA to generate the discriminant vectors and applying the discriminant criteria to diagnose the faults.

34.3.2.2 Classification Criteria

In general, there are two approaches to separate the faults. One method is to project matrix Y on the discriminant vectors and plot the data in three-dimensional coordinate. The other is Mahalanobis distance criteria method. If the number of the classes is big and the data is a small sample, applying the first way is difficult to observe the distribution of the data. In this paper, the latter one is used.

Define $\mu = \{\mu_1, \mu_2, \cdots, \mu_m\}$ as the mean vectors of matrix Y, where m is the number of the measurement variables. Represent the *ith*class sample's covariance matrix with Σ_i. Define the Mahalanobis distance from the test data x to each class as follows:

$$D_i = d^2(x, Y) = (x - \mu)^T \sum_i^{-1} (x - \mu) \tag{34.22}$$

$$D_l = d_l^2(x) = \min_{i=1,\cdots,k} \{|d_i^2(x)|\} \tag{34.23}$$

when $i = l$, if the Eq. (34.23) holds, the test data x is classified to the class l.

34.4 Simulation

34.4.1 Data Requirement

In this paper, we build a rail vehicle model by SIMPACK/MATLAB with real parameters to obtain simulation data for this application. Different faults data in both primary and secondary suspension are simulated by this model. Here we

Table 34.1 The simulation faults

Scenario	Fault	Location	Fault pattern (%)
F_0	No fault		Lost 0
F_1	c_1	*Secondary*	Lost 50
F_2	k_1	*Secondary*	Lost 50
F_3	c_2	*Secondary*	Lost 50
F_4	k_2	Secondary	Lost 50
F_5	k_3	*Primary*	Lost 50
F_6	k_4	*Primary*	Lost 50
F_7	k_5	*Primary*	Lost 50
F_8	k_6	*Primary*	Lost 50

choose some typical faults of the suspension systems for the application, which are given in Table 34.1. The positions of the fault components are depicted in Fig. 34.1, where c_1, c_2 are the secondary dampers, k_1, k_2 are the secondary springs, and k_3, k_4, k_5, k_6 are the primary springs. The faults of these components arise from the 50 % loss of their damping performance.

34.4.2 FDA Fault Diagnosis Results

According to the steps in Fig. 34.2, the fault diagnosis results are shown in Table 34.2. In the table, the overstriking number is the smallest one of the column it belongs to, so the test data is classified to the fault class on the left of it. Nine test data are given to inspect the eight discriminant vectors and before that all the training data had been analyzed by FDA and the discriminant rate is 100 %. In Table 34.2, the numbers of every column are the Mahalanobis distances from the corresponding fault data to the training classes. In column 5, the smallest distance is 0.0026, so the k_3_50 fault is misruled to be the class of F_1. In column 7, the smallest distance is 0.1208 and the k_5_50 fault is misruled to be the class of F_2. The other test data are classified correctly. The discriminant rate of the test data is 77.8 %.

34.5 Conclusions

This paper proposes a new method to diagnose the faults of rail vehicle suspension systems by using FDA. FDA provides an optimal lower dimensional representation in the light of maximizing the separation between classes. The discriminant vectors are derived from the formula which can maximize the scatter the classes and minimize the scatter within each class.

Table 34.2 Mahalanobis distances to the training classes

Class name	Test faults name								
	c_1_50	k_1_50	c_2_50	k_2_50	k_3_50	k_4_50	k_5_50	k_6_50	No
F_0	1.4812	1.8068	1.6251	1.1400	0.0284	0.9959	0.3438	1.9581	**0.0159**
F_1	**0.0203**	0.4499	5.5561	0.2928	***0.0026***	0.3387	6.6641	0.4467	0.1768
F_2	0.2652	**0.0485**	0.2160	1.4357	1.0279	0.7640	***0.1208***	1.3101	0.1896
F_3	2.7087	0.0835	**0.1581**	0.7443	2.7877	0.1539	8.0163	9.7209	1.2196
F_4	2.0711	0.1116	0.5270	**0.0050**	0.0311	1.3341	1.5382	9.1743	1.0485
F_5	3.5572	1.3195	0.5580	1.0289	**0.0051**	0.2311	1.0197	0.6217	0.1648
F_6	0.3422	0.1920	0.7919	0.3941	0.0822	**0.1513**	0.2852	0.6172	0.0341
F_7	0.4772	0.3157	1.6159	2.7833	0.1502	0.3913	**0.2564**	2.0338	0.0504
F_8	1.4812	0.6151	0.7838	0.2194	0.0306	0.3706	1.8383	**0.0283**	0.8809

In order to inspect the validity of the vectors, we have applied them to analyze the training data and the test data. The diagnostic accuracy of the training data is 100 % and the diagnostic accuracy of the test data is 77.8 %. The more the fault classes gathered for training, the higher the diagnostic accuracy of the FDA will be. There are two more reasons which affect the diagnosis results. One is that the number of the training data is small and the difference between some classes is too small to separate. The other is the influence of the nonlinear factors. We can introduce a kernel function, i.e., KFDA, to improve this method.

In this paper, the proposed method has only been applied to the cases that 50 % coefficient reduction and the other applications on an arbitrary fault case between 0 and 100 % coefficient reduction have not been investigated yet. How to extend the proposed approach to arbitrary fault magnitude cases would be our future work.

Acknowledgment This work is partly supported by Chinese 863 program (Contract No. 2011AA110503-6) and Ph.D. Programs Foundation of Ministry of Education of China (grant number: 20110009120037).

References

1. Xiukun W, Limin J, Hai L (2012) Data-driven fault detection of vertical rail vehicle suspension systems. In: UKACC international conference on control, pp 589–594
2. Bruni S, Goodall R, Mei T, Tsunashima H (2007) Control and monitoring for railway vehicle dynamics. Veh Syst Dyn 45(7–8):765–771
3. Goodall R, Mei T (2006) Advanced control and monitoring for railway vehicle suspensions. In: International symposium on speed-up and service technology for railway and maglev systems (STECH'06), 1, Chengdu, China, pp 10–16
4. Goodall R, Roberts C (2006) Concepts and techniques for railway condition monitoring. In: IET international conference on railway condition monitoring, pp 90–95
5. Li P, Goodall R (2004) Model-based condition monitoring for railway vehicle systems. In: Control 2004, University of Bath, ID-058

6. Li P, Goodall R, Weston P, Ling CS, Goodman C, Roberts C (2006) Estimation of railway vehicle suspension parameters for condition monitoring. In: Control engineering practice, pp 43–55
7. Wei X, Liu H (2011) Fault diagnosis of rail vehicle suspension systems by using GLRT. In: Chinese control and decision conference
8. Xiukun W, Hai L, Yong Q (2011) Fault isolation of rail vehicle suspension systems by using similarity measure. In: International conference on intelligent railway transportation, pp 391–396
9. Yuting W, Yong Q, Xiukun W (2012) Track irregularities estimation based on acceleration measurements. In: 2012 international conference on measurement, information and control, pp 86–91
10. Duda RO, Hart PE, Stock DG (2001) Pattern classification, 2nd edn. Wiley, New York
11. Zhao X, Huihe S (2006) On-line batch process monitoring and diagnosing based on fisher discriminant analysis. J Shanghai Jiaotong Univ (science) E-11(3):307–312
12. Chiang LH, Kotanchek ME, Kordon AK (2004) Fault diagnosis based on fisher discriminant analysis and support vector machines. Comput Chem Eng 28(2004):1389–1401
13. Fuente MJ, Garcia G, Sainz GI (2008) Fault diagnosis in a plant using fisher discriminant analysis. In: 16th mediterranean conference on control and automation congress centre, Ajaccio, France, June 25–27, 2008
14. Yang B-S, Kim KJ (2006) Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals. Mech Syst Signal Process 20(2006):403–420

Chapter 35
Analysis of Related Factors Influencing Reliability of Railway Signaling Systems Based on Fuzzy Analytical Hierarchy Process

Hongxia Chen and Min An

Abstract A signaling system is a safety-critical part of railways. Most of the railway signaling systems currently used today are intelligent and automatic high-performance systems, for which the high level of reliability is required. However, there are many factors including those uncertainties that affect system reliability performance. In this paper, fuzzy analytical hierarchy process is further developed by using hierarchical structure to determine weights of contributions of each factor and subfactor to the reliability performance of the system. By using the proposed methodology, the results of reliability analysis indicate that the reliability of a signaling system can be assessed effectively and efficiently.

Keywords Railway signaling system · Fuzzy analytic hierarchy process · Influencing factor · Reliability

35.1 Introduction

Railway signaling systems are used to control railway traffic safely in order to prevent train accidents such as derailments and train crashes etc., particularly, in high-speed rail. Unlike other countries whose train speed is restricted by the existing conditions of the rail track [1]. Lessons learned from '7.23' high-speed train crash accident show that the failure of railway signaling system can cause serious accidents which lead to injury to personal (even death), damage to railway

H. Chen (✉)
Nanjing Institute of Railway Technology, Nanjing, Jiangsu, China
e-mail: chen.hongxia9@gmail.com

M. An
School of Civil Engineering, University of Birmingham, Birmingham, UK
e-mail: m.an@bham.ac.uk

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_35,
© Springer-Verlag Berlin Heidelberg 2014

infrastructures, and environment. Therefore, it is necessary to develop a methodology to assess and improve the reliability of railway signaling systems.

Reliability is the ability of a system or a component to perform its required functions under stated conditions for a specified period of time [2, 3]. However, there are many factors that affect safety and reliability of the system. Therefore, in practice, reliability evaluation of a railway signaling system is analyzed from a holistic point of view and considered as a multiple criteria decision-making (MCDM) problem. There are some methods existing in the literature to solve MCDM problems [4]. But applications of these methods in reliability analysis heavily rely on the availability and accuracy of the reliability data. However, in many circumstances, these methods often do not cope well with uncertainty of information. In railway system reliability analysis, there are numerous variables interacting in a complex manner which cannot be explicitly described by an algorithm, a set of equations, or a set of rules [2, 5]. Fuzzy-AHP is a useful tool to deal with imprecise, uncertain, and ambiguous data. It is more appropriate for reliability evaluation among datasets with high nonlinearity and complexity of relationships in a signaling system [6]. This paper aims to develop a model which can be used to identify and assess the causes that influence the safety and reliability of the railway signaling systems.

35.2 Methodology

Fuzzy set theory was originally introduced by Zadeh in 1965 [7]. Fuzzy reasoning approaches possess the ability to mimic the human mind to effectively employ modes of reasoning that are approximate rather than exact [5]. This method resembles human reasoning in its use of approximate information and uncertainty to generate decisions. A fuzzy set is defined as a class of objects with a continuum of grades of membership. It is characterized by a membership (characteristic) function (MF) which assigns to each object a grade of membership ranging between zero and one. A fuzzy number can describe the relationship between an uncertain quantity x and a membership function μ_A which is a real number in interval [0, 1]. A MF μ_A of a triangular fuzzy number (TFN) is given in Eq. (35.4), where l, m, and u denote real numbers [8, 9]. Common algebraic operations of TFNs can be found in Ref. [9].

$$\mu_A(x) = \begin{cases} (x - l)/(m - l), & l \le x \le m \\ (u - x)/(u - m), & m \le x \le u \\ 0, & otherwise \end{cases}. \tag{35.1}$$

The conventional analytical hierarchy process (AHP) is a method to resolve multiple criteria decision problems. The judgments are accomplished by 1 to 9 values of pair comparisons. To merge the experts' judgments, the pair comparison

matrices need to be constructed in which the relative weights of the criteria and alternatives are calculated [2, 10].

However, the conventional AHP method is incapable of handling the uncertainty and vagueness involved in the mapping of one's preference to an exact number or ratio. An et al. [2, 5], Chang [11] and Veerabathiran [12] extended AHP to Fuzzy-AHP by introducing the concept of fuzzy linguistic preference with a scale ranging from 0 to 1. The following outlines Fuzzy-AHP in reliability decision-making analysis.

Step1: Establish the Hierarchical Structure

The establishment of a hierarchical structure is a basis of reliability and safety analysis by using Fuzzy-AHP. The hierarchical structure can be developed by the decomposition of complex decision-making problems into smaller manageable elements at different hierarchical levels [11]. A four-level hierarchical structure is shown in Fig. 35.1. As can be seen that the first level of the hierarchy is an objective or a goal, the bottom level in a hierarchical structure is evaluation alternatives (options), whereas the intermediate levels are criteria and subcriteria.

Step 2: Establish the Pairwise Comparison Matrix and Calculate the Weights

The extent analysis method is used to consider the extent to which an object can satisfy the goal [12, 13]. In this method, the "extent" is quantified by using a TFN. Application of the comparison principle of Fuzzy-AHP, the weight vectors with respect to each element under a certain criterion is calculated. The details of the methodology are presented as follows:

Let $X = \{x_1, x_2, \ldots, x_n\}$ be an object set and $U = \{u_1, u_2, \ldots, u_3\}$ be a goal set. According to Ref. [11], each object is taken and an extent analysis for each goal, i.e., g_i $(i = 1, 2, \ldots, n)$, is performed. Hence, m extent analysis values for each object can be obtained as $M_{g_i}^1, M_{g_i}^2, \ldots, M_{g_i}^m$ $(i = 1, 2, \ldots, n)$ where all of the $M_{g_i}^m$ are TFNs whose parameters are l, m, and u. These parameters are of the minimum value, most likely value, and the maximum value, respectively, in which a TFN is represented as (l, m, u) and the value of fuzzy synthetic extent with respect to the ith object is defined as follows:

$$S_i = \sum_{j=1}^{m} M_{g_i}^j \otimes \left[\sum_{i=1}^{n} \sum_{j=1}^{m} M_{g_i}^j \right]^{-1}. \tag{35.2}$$

Pairwise comparison matrix is to decide relative importance of each pair of factors at the same level of a hierarchical structure. The fuzzy judgment matrix $A = \left(a_{i,j}\right)_{n \times m}$ is established by using triangular numbers and pairwise comparison [2, 9, 11, 12]. If strong important of element j is over element i, then the pairwise comparison scale can be defined by TFNs, i.e., $a_{j,i} = \left(a_{i,j}\right)^{-1} = \left(\frac{1}{u_{i,j}}, \frac{1}{m_{i,j}}, \frac{1}{l_{i,j}}\right)$.

The degree of possibility of $M_i = (l_i, m_i, u_i) \geq M_j = (l_j, m_j, u_j)$ $i \geq j$ is defined as

$$V(M_i \geq M_j) = \sup_{x \geq y} [\min(\mu_{M_i}(x), \mu_{M_j}(y))]. \tag{35.3}$$

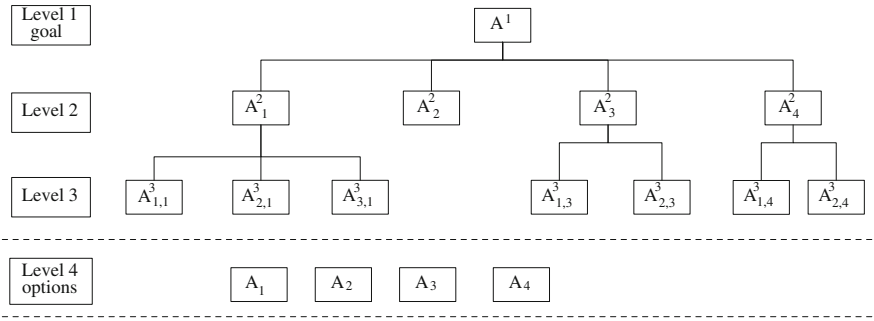


Fig. 35.1 Hierarchical structure

and can be equivalently expressed as given below.

$$V(M_i \geq M_j) = \begin{cases} 1, & m_i \geq m_j \\ \frac{(l_j - u_i)}{[(m_i - u_i) - (m_i - l_j)]}, & m_i \leq m_j, u_i \geq l_j \\ 0, & otherwise \end{cases}. \tag{35.4}$$

The degree of possibility for a convex fuzzy number to be greater than k convex fuzzy numbers $M_i (i = 1, 2, \ldots, k)$ can be defined as

$$\begin{aligned} &V(M \geq M_1, M_2, \ldots, M_k) \\ &= V[(M \geq M_1) \text{ and } (M \geq M_2) \text{ and } \cdots \text{ and } (M \geq M_K)]. \\ &= \min V(M \geq M_i), i = 1, 2, \ldots, k \end{aligned} \tag{35.5}$$

Assume the above Eq. (35.4) is

$$d'(A_i) = \min V(S_i \geq S_k), \quad k = 1, 2, \ldots, n;\ k \neq i. \tag{35.6}$$

Then the weight vector is given by

$$w' = (d'(A_1), d'(A_2), \ldots, d'(A_n))^T. \tag{35.7}$$

where $A_i (i = 1, 2, \ldots, n)$ are n elements at a level of a hierarchical structure.

Via normalization, the normalized weight vectors are given as follows:

Where W is a nonfuzzy number that gives priority weights of an alternative to others.

$$\begin{aligned} W &= (d(A_1), d(A_2), \ldots, d(A_n))^T \\ &= \left(\frac{d'(A_1)}{\sum_{i=1}^{n} d'(A_i)}, \frac{d'(A_2)}{\sum_{i=1}^{n} d'(A_i)}, \ldots, \frac{d'(A_n)}{\sum_{i=1}^{n} d'(A_i)} \right)^T. \end{aligned} \tag{35.8}$$

35.3 Application of Fuzzy-AHP to a Railway Signaling System

The proposed reliability analysis decision model as described in the previous sections is applied to analyze the influence factors. The modeling assumption was determined by taking into account the prestigious experts.

Step 1: Identify Factors and Establish Hierarchical Structure.

According to the accident-causing theory and failure model, the main factors are identified and classified into four categories as equipment (device), environment, human, and management factors as shown in Fig. 35.2. Those factors are accepted as the criteria. The subcriteria that are grouped as human, environment, equipment, and management are given in the left column in Table 35.1.

Step 2: Determine the Weight of Each Factor.

Nine-point scale is frequently used to represent the pairwise comparison and triangular fuzzy numbers (TFNs) are used to represent subjective pairwise comparison. Table 35.2 shows the pairwise comparison matrix. The global weights of the factors can then be calculated as shown in Table 35.2, whose linguistic variables proposed by Cheng and Wang [10] are used in this step.

35.4 Result Analysis

The result can be seen in Table 35.2, "Failure" is estimated to have the highest weight among all of the factors in the evaluation. The factor that has the second highest weight is "misoperation." "Man-made sabotage" and "unskilled worker" are ranked as the third and fourth, respectively. The results indicate that it is important to improve reliability of system hardware and reduce human errors in operation and maintenance activities. Because each failure of the system has a consequence at system level [12], this consequence can be evaluated both in terms of impact on the system and injury on human [13]. Each identified failure of the system should be detected, therefore, it is necessary that any equipment in a railway signaling system must have automatic high performance, self-inspection, and self-adaptation functions which can be achieved by applications of high-performance components and redundant design. Reoccurring problems do arise in railway signaling systems/equipment [14], such as a signal aspect reverts, a track circuit hangs, lightning-damaged microelectronics in a point machine needs the repair and relay switching fails which generally are given rise to be environment factors. Monitors and alarm systems are frequently required in railways. Maintenance should be carried out regularly according to the system states which can be checked in terms of the monitors. It is better to reduce failure rate and eliminate faults and guarantee safety. On the other hand, empirical workers are lack of new knowledge and cannot operate correctly the new equipment which is the main

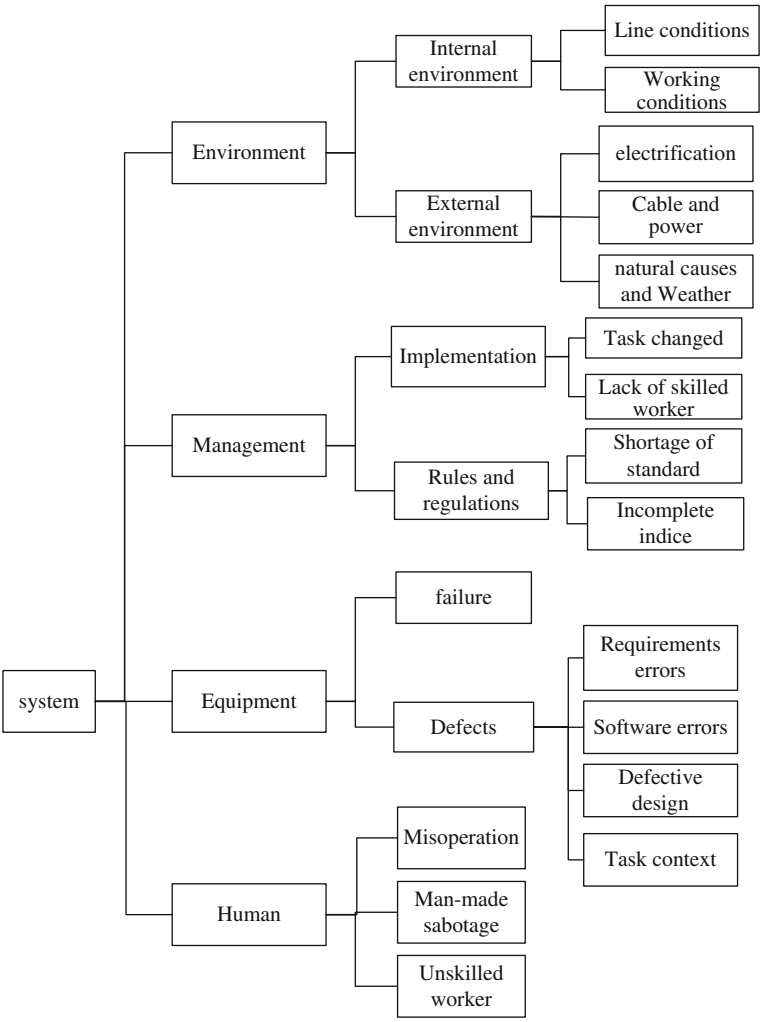


Fig. 35.2 A hierarchical structure of influence factors in the signaling system

Table 35.1 Pairwise comparison matrix

Factors	Human	Equipment	Environment	Management	Weights
Human	(1,1,1)	(1/2,2/3,1)	(2/3,1,2)	(1,3/2,2)	0.277
Equipment	(1,3/2,2)	(1,1,1)	(1,3/2,2)	(3/2,2,5/2)	0.365
Environment	(1/2,1,3/2)	(1/2,2/3,1)	(1,1,1)	(1,3/2,2)	0.233
Management	(1/2,2/3,1)	(2/5,1/2,2/3)	(1/2,2/3,1)	(1,1,1)	0.125

Table 35.2 Weights of factors

Factors	Equipment 0.365	Environment 0.233	Human 0.277	Management 0.125	Weight $\sum = 1.0$
Failure	0.333	0.000	0.000	0.000	0.1216
Misoperation	0.000	0.000	0.400	0.000	0.1108
Man-made sabotage	0.000	0.000	0.200	0.000	0.0554
Unskilled worker	0.000	0.000	0.400	0.000	0.1108
Task context	0.075	0.000	0.000	0.000	0.0274
Requirement errors	0.184	0.000	0.000	0.000	0.0671
Software errors	0.204	0.000	0.000	0.000	0.0743
Defective design	0.204	0.000	0.000	0.000	0.0743
Incomplete indices	0.000	0.000	0.000	0.250	0.0313
Shortage of standard	0.000	0.000	0.000	0.500	0.0626
Task change	0.000	0.000	0.000	0.060	0.0078
Shortage of skilled workers	0.000	0.000	0.000	0.190	0.0235
Electrification	0.000	0.240	0.000	0.000	0.0575
Cable and power	0.000	0.160	0.000	0.000	0.0362
Natural factors and weather	0.000	0.100	0.000	0.000	0.0228
Working condition	0.000	0.250	0.000	0.000	0.0583
Line condition	0.000	0.250	0.000	0.000	0.0583

reason that leads to system failure. Vocational education and technology training to the empirical workers and new staff which will reduce the human errors should be carried out.

35.5 Conclusions

This paper presents a Fuzzy-AHP methodology that is able to deal with adequately incomplete and uncertain data in determining contribution weights of influence factors to reliability of the system. A case example is used to demonstrate the proposed method. The results indicate that components intrinsic failure dominates the main reason which influences the reliability of a signaling system. On the basis of results some recommendations are given.

References

1. Zhou L, Shen Z (2011) Progress in high-speed train technology around the world. J Mod Transp 19(1):1–6
2. An M, Chen Y, Baker CJ (2011) A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: a railway risk management system. Inf Sci 181:3946–3966

3. Miee CE. EMC issues in advanced signaling transmission based train control systems. The IET, Michael Faraday House, Six Hills Way, Stevenage, Herts SG1 2AY, UK
4. Zhang G, Zhu N, Tian Z, Chen Y, Sun B (2012). Application of trapezoidal fuzzy AHP method for work safety evaluation and early warning rating of hot and humid environments. Saf Sci 50:228–239
5. An M, Lin W, Stirling A (2006) Fuzzy-based-approach to qualitative railway risk assessment. Proc IMechE Part F: J Rail Rapid Transit 220:153–167
6. Ardente F, Beccali M, Cellura M (2004) A fuzzy software for the energy and environmental balances of products. Ecol Model 176:359–379
7. Zadeh LA (1965) Fuzzy Sets. Inf Control 8:338–353
8. Timothy RJ (2010) Fuzzy logic with engineering applications (3rd ed). Wiley
9. Tesfamariam S, Sadiq R (2006) Risk-based environmental decision-making using fuzzy analytic hierarchy process (F-AHP). Stoch Environ Res Risk Assess 21:35–50
10. Chang T, Wang TC (2009) Using the fuzzy multi-criteria decision making approach for measuring the possibility of successful knowledge management. Inf Sci 179(4):355–370
11. Chang DY (1996) Applications of the extent analysis method on fuzzy AHP. Eur J Oper Res 95:649–655
12. Veerabathiran R (2012) Applications of the extent analysis method on fuzzy AHP. Int J Eng Sci Technol 4(7):3472–3480
13. Vernez D, Vuille F (2009) A method to assess and optimise dependability of complex macro-systems: application to a railway signalling system. Saf Sci 47:382–394
14. EN 50129 (2002) Railway applications: communication, signaling and processing systems—safety related electronic systems for signaling

Chapter 36
Application of GA-LSSVM in Fault Diagnosis of Subway Auxiliary Inverter

Junwei Gao, Ziwen Leng, Yong Qin, Limin Jia and Dechen Yao

Abstract Focusing on the fault diagnosis precision of subway auxiliary inverter, the diagnosis method based on genetic algorithm (GA) and least squares support vector machine (LSSVM) is proposed in this paper. First, the optimal parameters of LSSVM are obtained by GA with global search capability and the diagnosis model of the optimized LSSVM is established, then the empirical mode decomposition (EMD) is introduced to decompose the fault signal into several intrinsic mode functions (IMF), and finally we will extract the approximate entropy of each IMF as the fault feature which will be applied to test the performance of the diagnosis model. Simulation results have proved that the proposed diagnosis method is feasible to recognize each fault and has achieved higher precision.

Keywords Fault diagnosis · Subway auxiliary inverter · LSSVM · Genetic algorithm · EMD · Approximate entropy

36.1 Introduction

With the fast development of advanced technologies, the whole performance of subway vehicle has been greatly improved and it has brought people obvious convenience. As an important component of a vehicle, the structure of subway auxiliary inverter is becoming more complicated and highly integrated, and more faults have occurred. Since the working stability of auxiliary inverter will have an

J. Gao (✉) · Z. Leng
College of Automation Engineering, Qingdao University, No 308 Ningxia Road, Qingdao 266071, China
e-mail: qdgao163@163.com

J. Gao · Y. Qin · L. Jia · D. Yao
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_36,
© Springer-Verlag Berlin Heidelberg 2014

impact on the running safety of the vehicle, it is of great significance to accurately recognize kinds of faults.

Least squares support vector machine (LSSVM) is the improvement of standard SVM, which is based on statistical theory and structural risk minimization principle [1]. LSSVM [2, 3] takes the equality constraints to replace the inequality constraints, selects the least square linear system as the loss function, and transforms the solution of quadratic programming into solving linear equations. Since LSSVM has the excellent capability in learning small sample [4], it is widely applied in the field of fault classification. However, the parameter selection of LSSVM will have an influence on the diagnosis precision. With better parallelism and global search capability, genetic algorithm (GA) can search the global optimal solution in a complex space [5–7]. To improve the diagnosis efficiency and precision, GA is introduced to search the optimal parameters of LSSVM.

As an adaptive decomposition method, empirical mode decomposition (EMD) [8] can analyze nonlinear and nonstationary signal, and decompose the signal into a small number of intrinsic mode functions (IMF). Approximate entropy [9] can reflect the complexity of time series and measure the probability of newly generated pattern in fault signal. In this paper, we use EMD to process the fault signals, extract the approximate entropy of the decomposed IMF, and compose the feature vector which will be taken as the input of GA-LSSVM for fault diagnosis.

36.2 Basic Theory of LSSVM

For the given training sample $\{(x_i, y_i)|i = 1, 2, \ldots, l\}$, the objective function and constraint condition are [10]

$$\begin{cases} \min J(w, \xi) = \dfrac{1}{2} w^T w + \dfrac{1}{2} \gamma \sum_{i=1}^{l} \xi_i^2 \\ s.t. \quad y_i = w^T \varphi(x_i) + b + \xi_i \ , i = 1, 2, \ldots, l \end{cases} \tag{36.1}$$

where w is the weight vector, ξ is the error variable, γ is the penalty factor, $\varphi(x)$ is the nonlinear transformation function, and b is the bias coefficient.

The Lagrange function is introduced to solve the optimization problem

$$L(w, b, \xi, \alpha) = J(w, \xi) - \sum_{i=1}^{l} \alpha_i \left[w^T \varphi(x_i) + b + \xi_i - y_i \right] \tag{36.2}$$

where α_i is the Lagrange multiplier.

Calculate the partial derivative of w, b, ξ and a in Lagrange function, and assign the value zero to the partial derivative.

$$\frac{\partial L}{\partial w} = 0, \ \frac{\partial L}{\partial b} = 0, \ \frac{\partial L}{\partial \xi} = 0, \ \frac{\partial L}{\partial \alpha} = 0 \tag{36.3}$$

The optimization problem is transformed into solving linear equations

$$\begin{bmatrix} 0 & I^T \\ I & K + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{36.4}$$

where $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ is the kernel function which satisfies the Mercer theorem, $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_l]^T$, $y = [y_1, y_2, \ldots, y_l]^T$ and $I = [1, 1, \ldots, 1]^T$.

Calculate α and b according to Eq. (36.4) and the optimal classification function of LSSVM is obtained.

$$\bar{y}(x) = sgn\left[\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b\right] \tag{36.5}$$

Considering the good generalization performance of radial basis function, the paper selects it as the kernel function of LSSVM.

$$K(x_i, x_j) = \exp\left(-\left\|x_i - x_j\right\|^2 \Big/ 2\sigma^2\right) \tag{36.6}$$

where σ is the kernel parameter.

36.3 Parameter Selection of LSSVM by GA

After the structure of LSSVM is determined, we need to search the optimal penalty factor γ and kernel parameter σ while the appropriate selection of two parameters will affect the learning performance of LSSVM. With the capability of imitating natural selection and genetic mechanism, genetic algorithm will achieve the population evolution and search the global optimal solution by genetic operations of selection, crossover, and mutation. The optimization process of LSSVM by GA is shown as follows:

- Coding mode. Initialize the population size and maximum iteration. For the easy operation of real coding method, we select it to code the penalty factor γ and kernel parameter σ to generate initial population.
- Fitness function. The introduction of GA is to search the optimal parameters so as to improve the classification precision of LSSVM. Therefore, the correct classification accuracy is chosen as the fitness which will evaluate the performance of individuals. Higher the correct classification accuracy is, better is the adaptability of corresponding individual.
- Selection. Based on the calculated fitness, we adopt the roulette method and the optimal individual preservation strategy to guarantee GA to converge to the global solution.
- Adaptive crossover. We select arithmetic crossover to generate new individuals. Since the execution of crossover depends on the crossover probability p_c, we

choose the adaptive adjustment of p_c [11] to improve the global search capability.

$$p_c = \begin{cases} \frac{P_{c1}(f_{\max} - f)}{f_{\max} - f_{\text{avg}}}, & f \geq f_{\text{avg}} \\ p_{c2}, & f < f_{\text{avg}} \end{cases} \tag{36.7}$$

where $f_{\max}$ is the maximum fitness of population, f_{avg} is the average fitness of population, f is the larger fitness of the selected two individuals, and $p_{c1} = p_{c2} = 0.8$.

- Adaptive mutation. To keep the population diversity, nonuniform mutation is chosen. The mutation probability p_m is also adaptively selected according to population fitness.

$$p_m = \begin{cases} \frac{p_{\text{m1}}(f_{\max} - f')}{f_{\max} - f_{\text{avg}}}, & f' \geq f_{\text{avg}} \\ p_{\text{m2}}, & f' < f_{\text{avg}} \end{cases} \tag{36.8}$$

where f' is the fitness of individual to be mutated and $p_{\text{m1}} = p_{\text{m2}} = 0.2$.

- When the optimization process is finished, the global optimal solutions of penalty factor and kernel parameter are assigned to LSSVM for the training and testing of fault sample.

36.4 Extraction of Fault Feature by EMD and Approximate Entropy

Since the fault signal of auxiliary inverter is nonlinear, the accurate extraction of fault feature is important. EMD is based on the local characteristics time scale, so it is applicable to decompose the fault signal into a collection of IMF adaptively which will achieve the separation of fault feature. With good resistance to noise and excellent capability of anti-interference, approximate entropy has been applied in fault diagnosis to describe the complexity of signals. The extraction process of fault feature combining EMD and approximate entropy is shown as follows:

- Suppose that fault signal is $x(t)$, and $x(t)$ is decomposed into finite IMF with different instantaneous frequencies and a residue by EMD.

$$x(t) = \sum_{i=1}^{n} c_i + r_n \tag{36.9}$$

where c_i is the *ith* IMF, r_n is the residue.

- Select the former m IMF and calculate the corresponding approximate entropy of each IMF.
- Constitute feature vector E based on the calculated approximate entropy.

$$E = [E_1, E_2, \ldots, E_m] \tag{36.10}$$

36.5 Simulation Experiment

In this paper, we will recognize three types of faults in subway auxiliary inverter, including frequency variation, impulsive transient, and voltage fluctuation. We select the experimental data from Ref. [1]. With sampling frequency 4,096 Hz, sampling points are 4,096 and each fault has 30 groups of data. The fault diagnosis of auxiliary inverter is achieved by MATLAB programming.

36.5.1 Extraction of Fault Feature

We process fault signals by EMD and the decomposition results of frequency variation signal are presented in Fig. 36.1. For different fault signals, the number of decomposed IMF varies. Therefore, we calculate the approximate entropies of the former four IMF to compose the fault feature vector. Table 36.1 shows parts of the fault feature sample.

36.5.2 Diagnosis Results

For each type of fault feature sample, we randomly take 16 groups of data to make up training sample and the remaining 14 groups are taken as testing sample. The population size of GA is 20, the maximum iteration is 100, and the search ranges of penalty factor γ and kernel parameter σ are [10, 1000] and [0.1, 10] respectively.

In this paper, we adopt the 'One VS All' method to establish multiple classifiers of LSSVM. When the optimization process of GA is over, the optimal parameters of LSSVM are determined $\gamma = 518.32$ and $\sigma = 1.47$. To compare the diagnosis effect of the proposed GA-LSSVM, we also adopt the common LSSVM whose parameters are selected by crossvalidation (C-LSSVM). The diagnosis results of two models are given in Table 36.2. As is seen in Table 36.2, the diagnosis accuracy of GA-LSSVM is 95.24 % which is higher than that of C-LSSVM, so GA-LSSVM is feasible and effective for fault diagnosis.

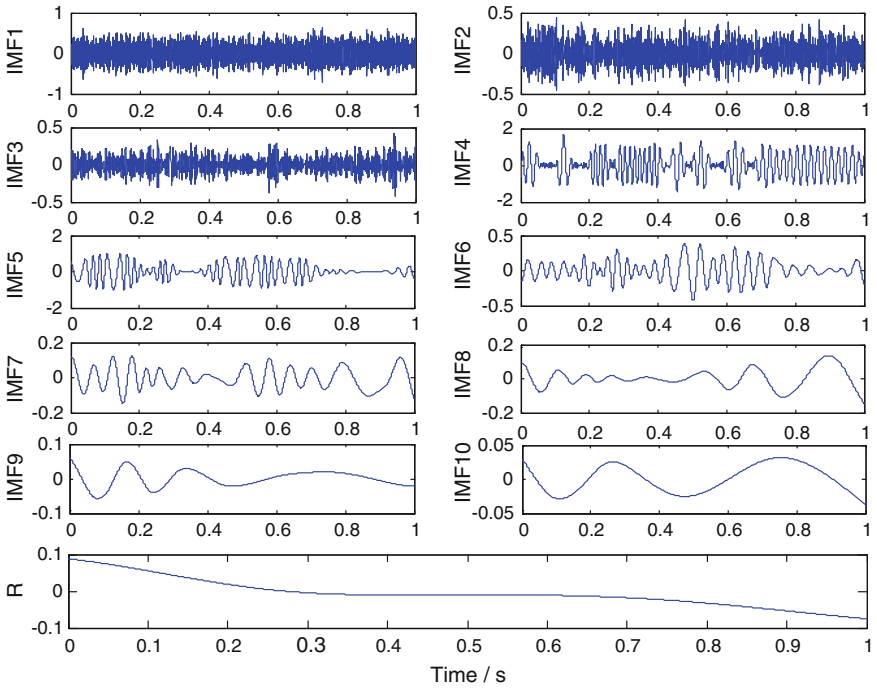


Fig. 36.1 Decomposition of frequency variation signal

Table 36.1 Parts of the fault feature samples

Fault	Series	E_1	E_2	E_3	E_4
Frequency variation	1	0.9759	0.3423	0.2891	0.6062
	2	1.0081	0.4143	0.3207	0.5654
	3	1.1489	0.5233	0.4261	0.6160
Pulse transient	1	0.7295	0.5181	0.7295	0.4331
	2	0.7303	0.4471	0.6129	0.5796
	3	0.6989	0.4270	0.6368	0.5354
Voltage fluctuation	1	0.7517	0.2632	0.2336	0.5138
	2	0.7607	0.2553	0.2427	0.4940
	3	0.8044	0.2730	0.2625	0.5013

Table 36.2 Diagnosis results of GA-LSSVM and C-LSSVM

Diagnosis model	Penalty factor	Kernel parameter	Frequency variation	Pulse transient	Voltage fluctuation	Diagnosis accuracy (%)
GA-LSSVM	518.32	1.47	14/14	13/14	13/14	95.24
C-LSSVM	140.68	2.93	12/14	13/14	13/14	90.48

36.6 Conclusion

The paper adopts genetic algorithm to search the optimal parameters of LSSVM, builds the multiple classifiers of LSSVM, and applies the optimized LSSVM in the fault diagnosis of subway auxiliary inverter. Experiment results show that the proposed GA-LSSVM has fully taken advantage of the small sample, overcome the random selection of parameters, and achieved higher diagnosis accuracy.

Acknowledgments This work is partially supported by the National Science and Technology Pillar Program (2011BAG01B05), the Foundation of Shandong Province (BS2011DX008, ZR2011FM008), and the State Key Laboratory of Rail Traffic Control and Safety Foundation (RCS2011K005), Beijing Jiaotong University.

References

1. Gao JW, Yu JP, Leng ZW et al (2013) The application of PSO-LSSVM in fault diagnosis of subway auxiliary inverter. ICIC Express Lett B Appl 4(3):777–784
2. Suykens JAK, Vandewalle J (2000) Recurrent least squares support vector machine. IEEE Trans Circuits Syst-I 47(7):1109–1114
3. Liao RJ, Bian JP, Yang LJ et al (2012) Forecasting dissolved gases content in power transformer oil based on weakening buffer operator and least square support vector machine-Markov. IET Gener Transm Distrib 6(2):142–151
4. Li W, Shi LS, Liang CL (2009) Forecasting model of research octane number based on PSO-VB-LSSVM. Chin J Sci Instrum 30(2):335–933 (in Chinese)
5. Liao WJ, Balzen Z (2013) LSSVM network flow prediction based on the self-adaptive genetic algorithm optimization. J Netw 8(2):507–512
6. Gao JW, Leng ZW, Zhang B et al (2013) Traffic flow forecasting based on wavelet neural network optimized by GA. In: Proceedings of the 32nd Chinese control conference, CCC2013, pp 8708–8712
7. Gao JW, Leng ZW, Zhang B et al (2013) Short-term forecasting of traffic flow based on genetic algorithm and BP neural network. Lect Notes Electr Eng 254:745–752
8. Huang NE, Shen Z, Long SR et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc London A 454:903–995
9. Hu HY, Ma XJ (2006) Application of local wave approximate entropy in mechanical fault diagnosis. J Vibr Shock 25(4):38–46 (in Chinese)
10. Zhang DW, Duan ZM, Li P et al (2010) Fault diagnosis method based on genetic method and LSSVM. Comput Measur Control 18(8):1745–1747 (in Chinese)
11. Srinivas M, Patnaik LM (1994) Adaptive probabilities of crossover and mutation in genetic algorithm. IEEE Trans Syst Man Cybern 24:656–667

Chapter 37
The City Rail Safety Detection Sensor Network Platform Using Online Track Detection

Shaohuang Pang, Weiping Zhao, Honghui Dong, Wenlong Peng and Jianxiao Chen

Abstract Corrugation is a common rail damage phenomenon worldwide, delays in processing will directly lead to the deterioration of the subway train wheel rails' relationship, and affect traffic safety [1]. We can get the real-time kinetic data by use of the similar urban rail sensor network platform which was researched in Part five of the "863" project, and through discrimination operation, exclude the wheel, bearing, and other factors and get the track's real-time status information. In this way, we can find the track's bad state and locate the accurate position, promptly repair the track in a polished method, and truly achieve the repair according to the condition. If we configure 2–3 trains with this function in each line, we can finish monitoring of all tracks effectively, which has a practical significance for promotion.

Keywords Train safety detection network · Track detection · Corrugation

37.1 Introduction

Today, the city rail transit's development and subway lines' extension has greatly eased the city's ground transportation. But because the subway adopts the same wheel rail system with our country's railway, it is inevitable to have rail corrugation phenomena, especially the subway line with features of small radius curves, multiples, high slope, and large vibration damping measures [2]. Through the collected statistics, we can see that the corrugation's wavelength range is generally

S. Pang (✉) · W. Zhao
Guangzhou Metro Corporation, Guangzhou, China
e-mail: 525707170@qq.com

H. Dong · W. Peng
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

J. Chen
Zhuzhou CSR Times Electric Co., Ltd., Zhuzhou, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_37,
© Springer-Verlag Berlin Heidelberg 2014

30–600 mm; the corrugation develops rapidly and often appears in the small radius curve and the track that adopts elastic damping element for reducing the rigid support. The smaller the radius, the faster the speed of development; the bigget the slope, the more quickly the rail corrugation formation. Therefore, the subway line's characteristics makes rail corrugation to form quickly and develop seriously, but when the normal line is in good condition, the development of the rail corrugation is not balanced.

The development of corrugation exacerbates the rail vehicle's wheel relationship. A lot of the city's rail vehicles in our country have appeared with different degrees of abnormal damage in walking system units.

37.2 Reason Analysis

Generally, the subway train is an EMU and the running part is a two-axle bogie. The bogie includes bogie frame, axle wheel set, axle box, two-line suspension device, traction device, traction motor, and gear box [3]. Each component has certain strength and rigidity in design structure and can withstand the wheel rail's alternating stress which is produced during operation and in accordance with the requirement of EN61373 "railway locomotive vehicle equipment—vibration and shock testing."

The rail corrugation due to rail damage will produce corresponding vibration acceleration during the wheels pass through, the corrugation's wavelength, wave depth, and speed all of which are related to the vertical vibration acceleration. A city subway train dynamics test showed that the wheel rail's vertical vibration acceleration in the damaged track had been more than 150 g, far more than the relevant standards; if failed to be timely treated, in long-term operation it will lead to premature fatigue failure in vehicle walking units, including axle box bearing, motor bearing, and so on [4].

Secondly, the wheel's ellipticity will also influence the wheel track's vertical vibration acceleration, but the wheel ellipticity has a character of periodic vibration.

37.3 Measures

37.3.1 The Direction of Measure

In our country, the city Metro Inc pays more attention to the repair work of damaged tracks. During the process, the corrugation section's location and status cannot be measured directly; the traditional method of polishing the track regularly has the feature of heavy workload, low pertinence, and efficiency, and the whole line polish also will make the track early to the limit for change and improves the

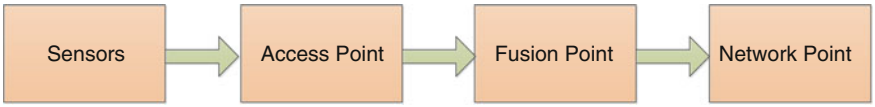


Fig. 37.1 Sensor network's data flow and connection structure

cost of maintenance. Now in the world lots of countries employ track inspection car, but because of the bogie structure, axle load, and real speed cannot equate with the operation train, the test result is not believable.

In recent years, in order to solve the problem of vehicle vibration spring, Guangzhou Metro Corporation and Southwest Jiao Tong University carried out many dynamics tests, the result of three-line north extension section test shows that the vertical vibration acceleration caused by the wave grinding track which with 30–50 mm wavelength and amplitude of 0.06 mm is 3–19 times than rail grinded, beyond the vehicle parts' design input conditions, and it is the main reason for vehicle and track parts' vibration [5]. But the test to wavelet grinding amplitude needs higher requirements for the test equipment and personnel; at present the detection of wavelet mill track status can only rely on pedestrian detection with precision equipment, the method with low efficiency and difficulty to meet the needs of line maintenance.

So, we will solve this problem from the direction of how to determine which section of rail should be polished, determine the effective positioning and specific polish, ensure that the vehicle is running in good condition, and reduce the workload. This paper puts forward for the first time to fix the acceleration sensors in the trains, a series of spring and the use of generalized resonance and resonance demodulation principle; each line is equipped with two columns and each column equipped in two cars. It can complete the analysis to wheel's ellipticity, bearing, and the service life trend at the same time, realizing the monitoring to the rail corrugation status and track state information.

The project of "City Rail train in road detection and early warning technology" undertaken by Guangzhou Metro Corporation has achieved initial results, forming a complete safety sensor network. In the train's journal box body of running part equipped with acceleration sensor device which can detect temperature can monitor the temperature of the bearing, and make full use of the resonate demodulation technology to complete the running parts' gears, gear, and other key components for comprehensive monitoring and diagnosis [6]. Track traffic safety monitoring sensor network's data flow and connection structure is shown in Fig. 37.1.

37.3.2 Technical Proposal

We should make full use of the existing "863" project "City Rail train in road detection and early warning technology" platform and the original acceleration

sensor devices equipped journal box body of running part, in addition redevelopment of the existing safety sensor network's analysis software, pay more attention to the transmission data's deep analysis, exclude the impact of wheel's ellipticity and crankcase bearing, so as to obtain the real status of rail damage [7]. We also should combine with the "863" network platform to obtain the train control system TCN network's location information and with its help realize the train running location and vibration data's consistency. At the same time, in the process of software development, we can by the method of increasing the vibration allowable value's 85 % of metro vehicle unsprung as the early warning target value, and realize the real-time warning purposes, it can be used as a further guide for the turning wheel and rail's polish work. In this way, the real-time detection to the track condition is realized, and it has great importance to ensure that the vehicle always runs in the track in accordance with the vehicle design standards.

37.4 Conclusion

Fixing the acceleration sensors in trains running part of a spring, and in safety sensor network platform similar to the "863" project adopt the generalized resonance and resonance demodulation principle to redevelopment walking system fault diagnosis system's analysis software not only achieved the life trend judgement of key device (bearing, wheel, tread), but can also easily know the full rail's corrugation, development speed, and accurate location. This has brought a new breakthrough to the maintenance of subway trains [8]. With its help, the main components of metro trains have changed from periodical repair to preventive repair, largely reducing the train and track's repair cost and effort. At the same time it also brings the track into the category of road detection and early warning technology.

This paper did not analyze the wheel stress relationship to find out the effective means for reducing corrugation formation, but creatively put forward by using subway vehicles to detect the actual track acceleration, and in this way determining whether the track needs to be repaired. This method is simple, practical, easy to reform, and convenient to spread.

Acknowledgments This work is supported by the National High-Tech Research and Development Program of China "863 Project" (Grant No.2011AA110505), the State Key Laboratory of Rail Traffic Control and Safety (Contract No. RCS2010ZT004), and the Star of Science and Technology Program of Beijing (Grant Z1211106002512027).

References

1. Liu H (2004) Vehicle maintenance and equipment of railway construction. China Railway Publishing House (in Chinese)
2. Shi H (1994) Wheel surface scratch and its harm. Railw Veh 32:56–57 (in Chinese)

3. Jia Meng (2005) The present situation and development of rail abrasion detection. Railw Tech Superv 1:34–36 (in Chinese)
4. Knothe K, Grassie SL (1993) Modelling of railway track and vehicle/track interaction at high frequencies. Veh Syst Dyn 22:209–262
5. Inhida M, Moto T, Takikawa M (2002) The effect of lateral creepage force on rail corrugation on low rail at sharp curves. Wear 253:172–177
6. Wang W (2007) High speed railway's rail grinding technology and its application. J Southwest Jiao Tong Univ 05 (in Chinese)
7. He Z (2000) The thinking of foreign rail grinding technique and its application. China Railw 10 (in Chinese)
8. Liu Z (2003) High speed railway's rail grinding technology (Conference papers)

Part III
Rail Transportation Computer Technology

Chapter 38
Research for Algorithm of the Super Low Delay Image Coding in Airborne Photo-Electricity Survey Equipment

Qihu Li, Yunfeng Wen and Ruicai Jia

Abstract In order to solve the problem that the wireless channel bandwidths undertake the huge pressure, which caused by load product the large amounts of data in the airborne photoelectric survey equipment. A low complex image coding algorithm has been proposed based on the theory of the wavelet analysis. This algorithm has an ability that is very adaptive ultra-low delay realized by the VLSI, and draw a conclusion that is ultra-low delay image coding is implemented by VLSI requirements what feature to coding algorithm. The related redundancy between the pixel of the image has been removed by the integer lifting wavelet of the Le Gall(53) in the algorithm, and then the visual redundancy has been removed by the technique of the optimal quantity for the image, which has been transformed. Lastly, the probability redundancy has been removed by the zero run-length coding joint index coding of the Columbus. Simulation results show that the algorithm proposed has better quality of the image coding than traditional coding algorithm based on the wavelet analysis, and the structure of the algorithm is simple, and very adaptive ultra-low delay realized by the hardware. Another, the conclusion of the analysis can provide the principal evidence for the algorithm of image coding high-speed and ultra-low delay realized based on the hardware.

Keywords Wavelet analysis · Ultra-low delay compression · Optimal quantity · Zero run-length Exp-Columbus

38.1 Introduction

With the rapid development of the science and technology, airborne photoelectric measuring system load capacity, resolution, sampling rate are increasing, this limited transmission bandwidth of the channel has been unable to meet the actual

Q. Li (✉) · Y. Wen · R. Jia
The 54th Research Institute of CETC, Shijiazhuang, China
e-mail: gucas07@126.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_38,
© Springer-Verlag Berlin Heidelberg 2014

needs of the people. In order to ensure that the airborne photoelectric measuring systems for data precision airborne optical measurement system for precise control of the loop, which makes the image information that must be collected photoelectric measurement system load lowest latency compression, transmission, decompression, displayed on the terminal system [1]. However, the current international market image compression dedicated chip often only the pursuit of high-quality image compression, while ignoring the code delay. The airborne optical measurement system load data collection, low-latency compression is particularly important, traditional data compression methods; however, there are different degrees of limitations. Such as differential pulse code compression ratio is not high, the sharp increase of the computational complexity of the vector quantity method with the increase in dimension, JPEG again the presence of the relatively high compression box effect [2]. It is good when the frequency of the local area and multiresolution analysis capabilities are widely used in aerospace monitoring airborne optical image measurement system developed in the traditional Fourier analysis based on a variety of image compression based on wavelet analysis theory compression field.

At present, many inland and overseas scholars presented many wavelet image compression algorithm combined different coding measures, and in which the EZW, SPIHT, EBCOT, and some other algorithm that based on the wavelet analysis has more improvement and application, however, the traditional EZW and SPIHT encoding both used the linked list structure, which has a larger storage space and amount of calculation, even if the Listless structure improved subsequently, because its image compressing quality is not very good, and its application is greatly limited [3–5]. This article designed a low complexity and high-quality image compression algorithm based on integer wavelet transform.

38.2 Principle of Algorithm

The purpose of compression is remove the redundancy in the image, so the fewer bites rate can be used to represented the image information. The framework of the transform encoding is mainly composing of three parts. Transform removes the correlation redundancy. Quantization removes the visual redundancy, simultaneity it is the most way to achieve data compression. Entropy coding is removes the redundancy of probability. Author design the flow diagram as show in Fig. 38.1.

2D_IWT express the two dimension inter wavelet transform, OPT-Quantity express the optimal quantity. LL-Predict express the predict coding for frequency band of LL. ARLC express the adaptive run-length coding, and Exp_Golomb express the coding of Exp_Golomb.

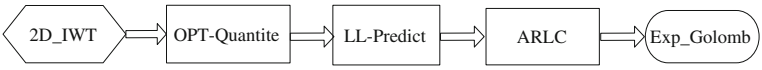


Fig. 38.1 Processes of algorithm

38.2.1 Theory of Wavelet Analysis

For given a basic function $\psi(t)$, for

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{38.1}$$

Both a and b are constants ($a > 0$), the function of (38.1) can be obtained by the basic function $\psi(t)$. When the a and b are constantly changing, a set of functions $\psi_{a,b}(t)$ can be obtained. So for the signal that is integrate square, for example $x(t)$, $x(t) \in L^2(R)$, the continuous wavelet transform can be defined as:

$$WT_x(a,b) = \frac{1}{\sqrt{a}}\int x(t)\psi^*(\frac{t-b}{a})dt = \int x(t)\psi^*_{a,b}(t)dt = <x(t), \psi_{a,b}(t)> \tag{38.2}$$

$\psi^*_{a,b}(t)$ is the conjugated expression of the $\psi_{a,b}(t)$ in the functions of (38.2). When the values of a and b continuous changed, it will result in a lot of redundancy to the base set of functions. So it is necessary to disperse the parameter a and b. To the function $a = a_0^m$ and $b = nba_0^m$, $m, n \in Z$, the function of wavelet basis can be expression as:

$$\psi_{m,n}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0) \tag{38.3}$$

The form of forward transform as:

$$X_{WT} = (m,n) = <x(t), \psi(t)> \tag{38.4}$$

The convolution algorithm has been adopted for the traditional wavelet transform. The MALLAT algorithm which based on filter structure is the typical representative, process of realization is complex, large amount of computation, poor real-time, and hardware implementation is difficult. The second generation wavelet transform which based on lifting structure has become the practical method for the wavelet transform. In this paper, the integer wavelet which is Le Gall(5/3) has been used in the algorithm of the low complexity image compression. The filter form of this integer wavelet corresponding to the form (38.5) and (38.6).

$$\tilde{g} = \left\{-\frac{1}{2}, 1, \frac{1}{2}\right\}; \tag{38.5}$$

$$\tilde{h} = \left\{ -\frac{1}{8}, \frac{1}{4}, \frac{3}{4}, \frac{1}{4}, -\frac{1}{8} \right\} \tag{38.6}$$

$\tilde{g}$ and $\tilde{h}$ express, respectively, the tap coefficients for the high-pass filter and low-pass filter to the analysis-side. When to wavelet transform for the one-dimensional sequence which length is N, requires 2N multiplication's operation if adopt the algorithm of MALLAT, but it only requires N multiplication's operation if adopt the algorithm of lifting, that can reduce the half amount of multiplication. In addition, by analyzing the characteristics of the data transformation of the filter, which can be drawn that the Le Gall(5/3) transformation process is typical of the same address same position operation [6]. The Le Gall (5/3) wavelet forward transform mathematical expression is:

$$Y(2n+1) = X_{\text{ext}}(2n+1) - \left| \frac{X_{\text{ext}}(2n) + X_{\text{ext}}(2n+2)}{2} \right| \quad i_0 - 1 \le 2n + 1 < i_1 + 1 \tag{38.7}$$

$$Y(2n) = X_{\text{ext}}(2n) - \left| \frac{Y(2n-1) + Y(2n+1) + 2}{4} \right| \quad i_0 \le 2n < i_1 \tag{38.8}$$

Y is the transformation result of the one-dimensional signals X in the formula, and $X_{\text{ext}}(n)$ is X symmetric periodic extension of the one-dimensional signals X. i_0 and i_1 is X sequence number, which is the first sample and last sample, respectively. The two-dimensional wavelet transform of the image is based on one-dimensional wavelet transform that is transform, respectively, for the line and columns of the image. The transformed image will be divided into four frequency bands, which respectively are LL band, HL band, LH band, and HH band. These bands reflect the different characteristics of the image in the different directions. When carried out the next level wavelet transforms, only operated to LL band accordance with the lifting algorithm [7].

38.2.2 Optimal Quantity and Prediction Operator

After wavelet transform, the information of the different image sub-band has the different weightings to the entire image reconstruction quality. Different bit rate allocation of the sub-band should be equivalent to the quantity step size selection of different sub-band. Taking into account the conversion gain and the human visual system (HVS) infection to the image compression system, where introduce two variables that are factors of human visual weighting (W_b) and factor of different band energy gain (G_b) the optimal quantity step size of sub-band is given by function (38.9).

$$\Delta_b = \sqrt{\frac{G_0 W_0}{G_b W_b}} \Delta_0 \tag{38.9}$$

Δ_0 is the basic quantity step in the function. By changing the Δ_0, the compression rate and distortion can be adjust. For the orthogonal wavelet transform, this function $G_b \approx 2^{2b}$ can be approximated. According with the function (38.9), three level wavelet transform optimum quantity step has been shown in Table 38.1.

For the image pixels in the low-frequency sub-band still has a very strong correlation, a prediction operator has been proposed, which has a low complexity and high accuracy. Forecast template shown as in Fig. 38.2. According with the prediction formula (38.10), the predicted value corresponding to prediction template can be obtained.

$$N_{i,j} = \begin{cases} \min(N_{i-1,j}, N_{i,j-1}) & N_{i,j} \geq \max(N_{i-1,j}, N_{i,j-1}) \\ \min(N_{i-1,j}, N_{i,j-1}) & N_{i,j} \leq \min(N_{i-1,j}, N_{i,j-1}) \\ N_{i+1,j-1} + N_{i,j-1} - N_{i-1,j-1} & \text{others} \end{cases} \tag{38.10}$$

In the lowest sub-band, with the predicted value and the true value to o perform the arithmetic operation of subtraction. The value of residuals has been used for adaptive zero run-length coding, and for the result of the coding of the zero run-length using Exp-Columbus encoding of the K = 0 [8]. It is worthiness that when encoding for the first row or first column of the lowest sub-band, due to lack of the corresponding pixel value in the LL sub-band, according to prediction formula, so it is necessity that offset a rows of data and a column of data, and values of these data are zero. When encoding for the HH sub-band, HL sub-band, or LH sub-band, as the wavelet transform has the good feature of decorrelation, and these bands will be has a large number of zero after the operator of quantity, so these bands can be directly coding by the zero run-length with the Exp-Columbus. The goal of this algorithm is pursuit of low delay coding, so entropy coding adopt the zero run-length with the Exp-Columbus, because it is without statistical probability of the source data [9]. These two kinds of entropy coding have good coding efficiency, and this coding mode is simpler relative to arithmetic coding, and convenient for engineering implementation.

38.3 Performance Analysis for the Algorithm Simulation

In order to test the effect of the algorithm which based on the wavelet analysis, author developed the simulation procedure of the codec at the environment of VC6.0, and compare with the algorithm of SPIHT and JPEG2000 all of which based on the wavelet analysis. The data of test result adopt the least squares acquire a curve, which shown in Fig. 38.3.

Table 38.1 Quantity step of 3 levels DWT sub-bands

band/level	1	2	3
HH	8Δ	4Δ	2Δ
HL	4Δ	2Δ	Δ
LH	4Δ	2Δ	Δ
LL	–	–	Δ

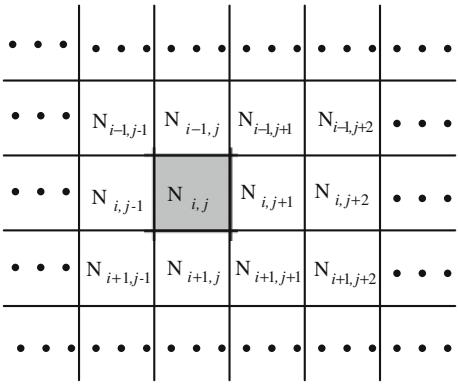


Fig. 38.2 Prediction template of the lowest frequency sub-band

It can be drawn from the results of the experimental data:

(1) The compression performance of proposed algorithm has better than the classical algorithm of SPIHT, slightly lower than the algorithm of JPEG2000, and simpler than the algorithm of SPIHT and JPEG2000 at the structure. The proposed algorithm neither has three linked list that is LSP, LIP, and LIS, nor assigned a storage unit which encode the each code block in the JPEG2000 standard algorithm. So the proposed algorithm saves a lot of storage capacity for the hardware.

(2) The compression performance of proposed algorithm has lower relative to the JPEG2000 standard. This is because the JPEG2000 standard adopted the algorithm of Embedded Block Coding with Optimized Truncation (EBCOT) which has high complexity structure. The algorithm of EBCOT adopted the optimal PCRD-opt to search the optimized truncation, which is optimized at the theoretically. When using the JPEG2000 algorithm to code, the each bit plane of each code block are used three channel scanning coding mode. The efficiency of entropy coding is superior to the proposed. The JPEG2000 algorithm has superior performance of coding and application is broader, but it is difficult to do high-speed implementation by the VLSI. At present, the sole fastest chip of the JPEG2000 coding speed is cannot satisfied the speed of the 30 frames images per second, which the resolution achieve the 1024 × 1024. The reason can be summed up as follows. First of all, the optimal interception of the JPEG2000 standard is based on the optimal rate-distortion optimal

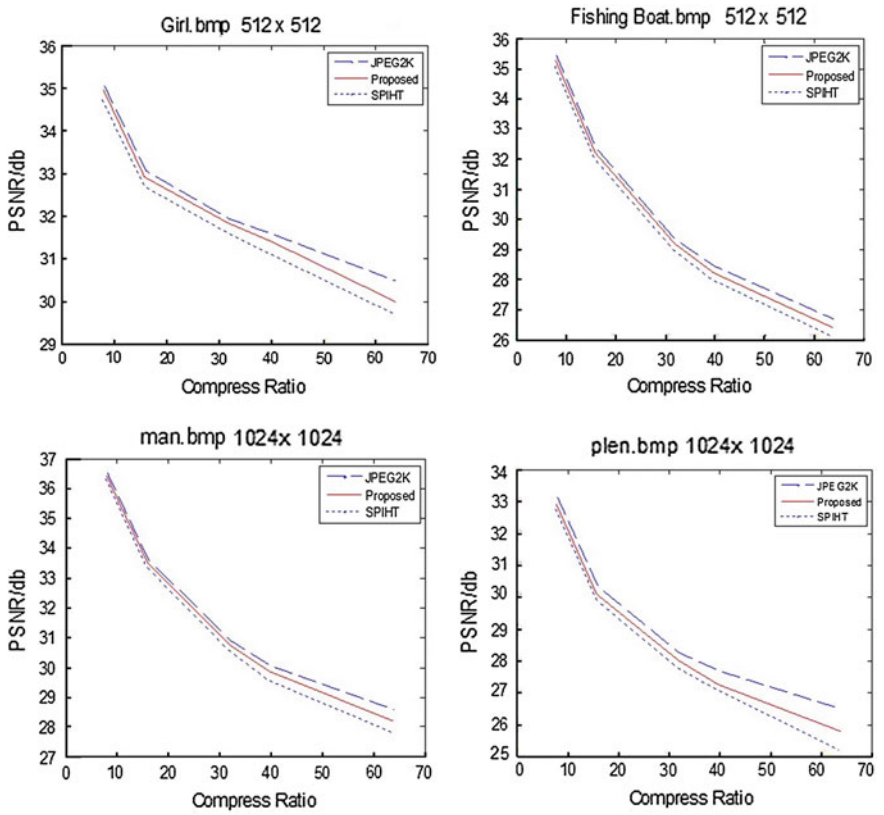


Fig. 38.3 Performance comparison of three coding algorithms

search which is optimal at the theoretically, but this operation will result in some feedback at the algorithm internal. This feedback operation is extremely conducive to the parallel implementation of the hardware. Secede, the coding core of the JPEG2000 standard is very complex, it is difficult implementation by the hardware. Those difficult of the algorithm implemented by the hardware result in it cannot be widely used in the super low delay image coding in airborne photo-electricity survey equipment.

(3) The proposed low complexity image algorithm can support the any coding rate from the loss to lossless. If the data which has been transform by the wavelet cannot be quantized, and directly encoded accordance with the coding format, the original image can be completely decoded, and without losing any information. The experiment result show that the performance of the lossless compression better than the algorithm of JPEG2000, and considerable with the algorithm of JPEG_LS.

38.4 Conclusion

In this paper, in order to solve the problem that the wireless channel bandwidths undertake the huge pressure, which caused by load product the large amounts of data in the airborne photoelectric survey equipment. A low complex image coding algorithm has been proposed which based on wavelet analyzes and coding unit based on the image pixels. First, multilevel integer wavelet transform has been done for an image, and then the optimal quantity step has been used for different sub-band. Second, a simple and effective predictive operator has been introduced for the data correlation of the lowest frequency band. Third, adaptive coding of zero run-length coding with the coding of Exp-Columbus has been used for the entropy coding. The result of the experiment show that proposed algorithm has better coding effect for the different type of image. Another, the feasibility of the algorithm has been validated in the practice project, and provides the principal evidence for the algorithm of image coding high-speed and ultra-low delay realized based on the hardware.

Acknowledgments The work was supported by the China Postdoctoral Science Foundation (2013M541202).

References

1. Taubman D (2000) High performance scalable image compression with EBCOT. IEEE Trans Image Process 9(7):1158–1170
2. Li QH, Ren GQ, Wu QZ et al (2011) An adaptive rate control algorithm for JPEG2000 based on optimal rate-distortion. Acta Geodaet Cartograph Sin 40:204–209
3. Chuang JC, Hu YC (2011) An adaptive image authentication scheme for vector quantization compressed image. J Vis Commun Image Represent 22(3):440–449
4. Li L, Guo M, Wang ZS (2009) Adaptive rate control for motion JPEG2000 . IET Image Proc 3(2):75–87
5. Auli-Llinas F, Serra-Sagrista J (2007) Low complesity JPEG2000 rate control through reverse subband scanning oder and coding pass concatenation. IEEE Signal Process Lett 14(4):251–254
6. Lina CJ, Chen KF (2001) Lifting based discrete wavelet transform architecture for JPEG2000. In: proceedings of the 2001 IEEE international symposium on circuits and system (ISIC-2001), Singapore, 2001, pp 497–500
7. Li H, Li G, Wen C (2006) Fast mode decision algorithm for inter-framecoding in fully scalable video coding. IEEE Trans Circuits Syst Video Technol 7(16):889–895
8. Glantz A, Krutz A, Haller M, Sikora T (2009) Video coding using global motion temporal filtering. In: proceedings of the 16th IEEE international conference on image processing (ICIP 2009), Cairo, Egypt, 11 November 2009
9. Su Y, Sun MT, Hsu V (2005) Global motion estimation from coarsely sampled motion vector field and the applications. IEEE Trans Circuits Syst Video Technol 15(2):232–242

Chapter 39
A Novel Recursive Algorithm for Training RBF Networks

Peng Zhou and Zhu Yang

Abstract A recursive learning algorithm is presented for basis selection of radial basis function (RBF) neural network. It is based on an adaptive kernel width algorithm, which can select basis functions recursively in the nonorthogonal space and assign an appropriate number of hidden units of RBF network. This also makes the model structure independent of the selected term sequence and assures an optimal RBF network even if the RBF original basis is nonorthogonal. Its effectiveness is demonstrated by the simulated results.

Keywords Radial basis function · Bayesian information criterion · Recursive

39.1 Introduction

Radial basis function neural networks have been widely used in signal processing, chaotic signals prediction, simulation, and design of circuits almost all engineering disciplines. Several learning algorithms have been proposed in the literature for training RBF neural networks, such as k-means, resource allocation network (RAN), RAN via extended Kalman filter (RANEKF), Rival Penalized Competitive Learning (RPCL), and growing and pruning RBF(GAP-RBF) [1]. The aforementioned algorithms may break down or have improper number of centers when its corresponding parameters or thresholds cannot be specified correctly. It is due to their intensive dependency on the distribution of the input–output samples.

P. Zhou (✉) · Z. Yang
State Key Laboratory of Optical Communication Technologies and Networks, Wuhan Research Institute of Posts and Telecommunications, Wuhan, China
e-mail: zhoupengsirok@163.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_39,
© Springer-Verlag Berlin Heidelberg 2014

A significant contribution to training RBF networks is made by Chen and Xiong through development of orthogonal least square (OLS) algorithm [2] and backward selection methods. However, the Achilles heel is that the final model obtained by these selection methods is not optimal [3]. A further improvement to the forward and backward methods is made by Kang Li et al. [4] and Jing Deng et al. [5] by introducing the two-stage subset selection algorithm (TSA), but the second stage only make a local improvement to the first stage and still cannot assure an optimal network.

To generate more compact network, in this letter, a novel recursive algorithm is presented. The new algorithm performs forward selection and backward refinement recursively in the "new information" space. The rest of the paper is organized as follow. Section 39.2 presents the RBF network architecture and understands how it works. The recursive algorithm is discussed in Sect. 39.3. Section 39.4 shows simulation experiments, our conclusions are exposed in Sect. 39.5.

39.2 Radial Basis Function Neural Network

A scalar output of RBF neural network with M inputs is depicted in

$$f(\mathbf{x}) = \theta_0 + \sum_{i=1}^{M_s} \theta_i \varphi(\|\mathbf{x} - \mathbf{c}_i\|^2 / 2\sigma^2) \tag{39.1}$$

To understand how this works, it is essential to view an RBF neural network (RBFNN) as the special case of the linear regression models. The corresponding geometric interpretation is best revealed by the following form. One has

$$\mathbf{d} = \mathbf{q}\theta + \mathbf{e} \tag{39.2}$$

where the regressor vectors $\mathbf{q}_i$ are regressor terms forming a set of basis vectors (when the fixed width σ is given, $\mathbf{q}_i = \exp(-\|x - \boldsymbol{c}_i\|^2/(2\sigma^2))$), $\mathbf{d}$ is output sample data, $\mathbf{e}$ is the model residual sequence. In the way, the problem of how to select a suitable set of RBFNN centers from the dataset can be regarded as an example of how to select a significant subset from a given candidate set of regressor.

39.3 Recursive Algorithm

A recursive method in the nonorthogonal space is employed to realize an optimal RBF network, in which the adaptive kernel width method is incorporated to consider the real distribution of input–output samples fully.

39.3.1 Automatic Model Selection

Among various subset selection algorithms, the model size is determined using an information measure, such as the final prediction error (FPE), minimum description length (MDL), Akaike information criterion (AIC), and Stein's unbiased risk estimator (SURE). These criteria provide a trade-off between the fitting error and model complexity, but unfortunately the tuning parameter for penalized term is hard to determine and is application dependent. As a result, the information criteria only provide a way of stopping the selection procedure, but stop point itself may well vary. However, as can be seen from the expression of Bayesian information criterion (BIC) [6], it is an important mechanism in balancing the two above mentioned. The convergence value of BIC is a desired trade-off. And we can also find proper values of kernel width by applying a meta-heuristic (Differential Evolution) using the convergence BIC for fitness calculation. Then numbers of hidden units and kernel width can be specified simultaneously.

Under the Gaussian model, Schwarz's procedure can be written as

$$BIC(N) = N\log(S_k^2) + L\log(N) \tag{39.3}$$

where S_k is a standard deviation of training set.

39.3.2 Forward Selection

In the forward model selection, the key step is to effectively compute the contribution of each regressor term to network output. For a given sample data d, the first center selection procedure of RBFNN is defined as follows. For $1 \le i \le M$, compute

$$[Q]_i = (\mathbf{q}_i^T \mathbf{d})/(\mathbf{q}_i^T \mathbf{q}_i) \tag{39.4}$$

where $\mathbf{q}_i$ denotes the ith column of regressor matrix. Note that since $[Q]_i$ denotes the contribution of any regressor term$\mathbf{q}_i$ to network output, the corresponding vector $\mathbf{q}_i$ is selected when $[Q]_i$ takes the maximum value. We have

$$[Q]_{i\max} = \max\{[Q]_i, 1 \le i \le M\} \tag{39.5}$$

In the way, ximax and $\mathbf{q}_{i\max}$ become the first selected center and regressor vector.

To generate an optimal RBF neural network, we must avoid redundant information as well as grasp necessary information of the hidden units. As a consequence, the concept of new information is proposed to achieve this target. Assume that the k regressor terms $\mathbf{R}_k\{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_k\}$ are added to the model. The corresponding selected centers are $c_k(x_1,x_2,\ldots, x_k)$. The remaining M–k centers and

regressor terms are c_k^* and $\mathbf{R}_k^*$. To formulate forward selection iteration process, the projection of $\mathbf{q}_i^*$ onto $\mathbf{R}_k$ is defined as

$$\mathbf{u}_i = \mathbf{R}_k(\mathbf{R}_k^T\mathbf{R}_k)^{-1}\mathbf{R}_k^T\mathbf{q}_i^* \tag{39.6}$$

And then the new information expressed by the residual regressor $\mathbf{q}_i^*$ can be written as

$$\boldsymbol{\delta}_i = \mathbf{q}_i^* - \mathbf{u}_i \tag{39.7}$$

Now, $\boldsymbol{\delta}_i$ span a new information subspace and the square modulus of projection of d onto$\boldsymbol{\delta}_i$ is obtained

$$[Q]_k^{(i)} = (\boldsymbol{\delta}_i^T\mathbf{d})/(\boldsymbol{\delta}_i^T\boldsymbol{\delta}_i),\ i = 1,\ldots,M-k \tag{39.8}$$

where $[Q]_k^{(i)}$ indicates the contribution of each residual regressor term $\mathbf{q}_i^*$ ($i = 1,\ldots,M–k$) which is used by the network to approximate d. The larger $[Q]_k^{(i)}$, the more information the residual regressor terms $\mathbf{q}_i^*$ contain. If one selected term achieves the maximum contribution among all remaining candidates and then the corresponding input vector is added to the model. The maximum contribution $[Q]_k^{(i\max)}$ is defined as

$$[Q]_k^{(i\max)} = \max\{[Q]_k^{(i)}, i = 1,\ldots,M-k\} \tag{39.9}$$

In such a way, the $(k + 1)$th center, which is the ikth column of $x_k^{(i\max)}$, can be obtained, $\mathbf{q}_k^{(i\max)}$ becomes the $(k + 1)$th selected regressor vector. The selected centers and regression matrix is given by

$$\mathbf{R}_{k+1} = [\mathbf{R}_k\ \mathbf{q}_k^{(i\max)}] \tag{39.10}$$

After a latest selected center is added to the model, whether or not the current model is to be expanded must be determined by the cost function. Suppose a subset of previously selected model contains k hidden units, the sum squared errors (SSE) is calculated by

$$\varphi_k = 1/\mathbf{R}_k^T\mathbf{R}_k,\ \mathbf{e}_k = \mathbf{d} - \mathbf{R}_k(\varphi_k(\mathbf{R}_k^T\mathbf{d})),\ S_k^T = \mathbf{e}_k^T\mathbf{e}_k/M \tag{39.11}$$

And then applying BIC criterion, the cost function is obtained by

$$BIC_k = M\ln(S_k^2) + k\ln(M) \tag{39.12}$$

Similarly, the cost function BIC_{k+1} of the (k + 1)th intermediate model is acquired by applying $\mathbf{R}_{k+1}$ to (39.11) and (39.12). If $BIC_{k+1}<BIC_k$, $c_{k+1} = [c_k, x_k^{(i\max)}]$. The cost function is gradually reduced with increasing the model size. The forward selection procedure continues until the cost function satisfies the following condition.

$$BIC_{k+1} > BIC_k \tag{39.13}$$

In order to reduce the computational burden, a fast recursive algorithm is used to update the key step φ_{k+1} in the process of computing BIC_{k+1}. When one more regressor term $\mathbf{q}_k^{(i\max)}$is added to model according to (39.10), the augmentation $\varphi_{k+1}$of φ_k is not calculated according to (39.11) but is implemented as follow [7].

$$\boldsymbol{\alpha} = \varphi_k \mathbf{R}_k^T \mathbf{q}_{i_k}^*,\ \bar{\mathbf{e}} = \mathbf{q}_{i_k}^* - \varphi_k \boldsymbol{\alpha}\beta = \bar{\mathbf{e}}^T \bar{\mathbf{e}} \tag{39.14}$$

$$\varphi_{k+1} = \begin{bmatrix} \varphi_k + \boldsymbol{\alpha}\boldsymbol{\alpha}^T/\beta & -\boldsymbol{\alpha}/\beta \\ -\boldsymbol{\alpha}^T/\beta & 1/\beta \end{bmatrix} \tag{39.15}$$

39.3.3 Backward Refinement

If the RBF original basis is nonorthogonal, the energy contributions of different basis vectors are mixed. As a consequence, after one latest selected regressor term is added to model, the contribution of some previously added units may change, and some may consequently become very little to the network output. So the stepwise forward selection procedure performs a series of constrained optimization. This process does not minimize the cost function BIC and is unable to find out the most compact permutation of regressor terms with maximal contribution to network output in the global sense. Even though each step yields a regressor with the largest possible contribution, the final selected regressor terms are not most compact, compactness for hidden layer. To solve the problem resulted by the stepwise forward method, a backward refinement procedure is proposed.

To evaluate the contribution of each selected term $\{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_k, \mathbf{q}_{k+1}\}$ and locate the regressor vector of the minimum contribution to net output, a new matrix $\mathbf{R}_{k,i}$, should be constructed, which is equal to $\mathbf{R}_{k+1}$ without the ith term of it.

$$R_{k,i} = [\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_{i-1}, \mathbf{q}_{i+1}, \ldots, \mathbf{q}_{k+1}] \tag{39.16}$$

And then the cost function ($BIC_{k,i}$) of the hypothetical models without the ith term of the selected regressor terms is acquired by applying $\mathbf{R}_{k,i}$ to (39.11) and (39.12). If the minimum value $BIC_{k,i\min}$ satisfies the following criterion:

$$BIC_{k,i\min} < BIC_k \tag{39.17}$$

So let k = k − 1, $BIC_k = BIC_{k,i\min}$, $\varphi_{k+1} = \varphi_{k,i\min}$, $\mathbf{R}_{k+1} = \mathbf{R}_{k,i\min}$, $c_{k+1} = c_{k,i\min}$. The refinement iteration procedure is terminated and then the forward selection process restarts if $BIC_{k,i\min} > BIC_k$.

Similar to φ_{k+1}, we can also use matrix transformation skills to speed up calculation of $\varphi_{k,i\min}$ which should be equal to $1/(\mathbf{R}_{k,i\min}^T \mathbf{R}_{k,i\min})$ according to

(39.11). The ith row and ith column of matrix φ_{k+1} move to the end, then new matrix φ_{k+1} and $\varphi_{k,i}$ can be written as

$$\varphi_{k+1} = \begin{bmatrix} \varphi & \gamma \\ \gamma^T & r \end{bmatrix}, \ \varphi_{k,i} = \varphi - \gamma\gamma^T/r \tag{39.18}$$

39.4 Computer Simulation

In this subsection, the results for approximating the bandlimited function are presented. Bandlimited function is given by the following equation:

$$F(x) = \sin(2\pi 5x) + \sin(2\pi 10x) + \sin(2\pi 15x) \tag{39.19}$$

A random sampling of the interval (0, 0.3) is used in obtaining 100 input–output data for the training set. The output set [8] can be modeled adequately as $\mathbf{y}_k = F(x_k) + fe_k$, where f describes the noise amplifying factor, which is a constant, $\mathbf{e}_k$ denotes the system noise, and it satisfies the norm distribution with mean value 0 and covariance 1, $\mathbf{y}_k$ (output set) is a fitting set and $F(x_k)$ are 100 uniformly sampled data for model test. The model test output is $\mathbf{y}_k^{(test)}$ and the test error rate (MSE) can be defined by

$$MSE = \sum (y_k^{(test)} - F(x_k))^2/N_t \tag{39.20}$$

where N_t denotes the number of test samples and here is 100. In order to verify the effectiveness of the proposed algorithm for different noise bandlimited function systems, both the proposed algorithm and two-stage algorithm (TSA) with the Bayesian information criteria (BIC) are applied to the noisy training dataset, whose amplifying factor f is uniformly sampled in the interval (0, 0.2). We fix the amplifying factor acquired just and then make the procedure run ten times in different random datasets. The fixed width σ is given by one or more iterations of the Differential Evolution (DE) algorithm. The experimental average results of 10 times from the two methods are shown in Figs. 39.1 and 39.2.

Figure 39.1 shows the behavior of the proposed method with noise standard variance 0.2 on the bandlimited function. As it is seen in Fig. 39.2a, the model constructed by the proposed method presents a slightly reduced number of nodes in the hidden layer under different noise levels. Then this means TSA will pick more basis vectors than needed. Figure 39.2b shows the model generalization performance using the MSE for the two methods. It can be seen that the proposed method resulted in a slightly lower approximation error under different noisy training conditions. So it can be seen that the proposed method can achieve a lower approximation error and also constructs a smaller RBF network than TSA (i.e., less number of hidden units) and also BIC is always an appropriate cost function under different noise levels.

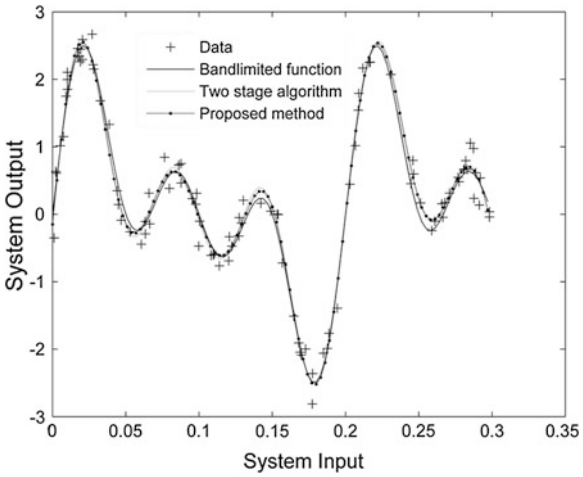


Fig. 39.1 Behavior of algorithms with the standard noise variance 0.2

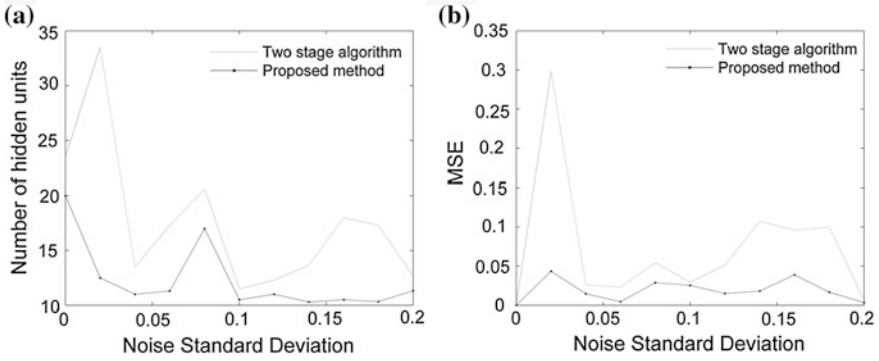


Fig. 39.2 Number of hidden units (**a**) and MSE (**b**) under different noise levels

39.5 Conclusion

In this paper, we have presented a method to assure an optimal RBF neural network even if the RBF original basis is nonorthogonal, in which BIC is used as fitness calculation. The experimental results show that the proposed algorithm can achieve much better performance over the classical method and BIC is an appropriate cost function.

References

1. Huang GB, Saratchandran P, Sundararajan N (2005) A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation. IEEE Trans Neural Network 16(1):57–67
2. Chen S, Cowan CFN, Grant PM (1991) Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans Neural Network 2(2):302–309
3. Sherstinsky A, Picard RW (1996) On the efficiency of the orthogonal least squares training method for radial basis function networks. IEEE Trans Neural Network 7(1):195–200
4. Li K, Peng JX, Bai EW (2009) Two-stage mixed discrete-continuous identification of radial basis function (RBF) neural model for nonlinear systems. IEEE Trans Circuits Syst 56(3):630–643
5. Deng J, Li K, Irwin GW (2010) A two-stage algorithm for automatic construction of RBF neural model. In: 15th IEEE mediterranean electrotechnical conference, vol 5795-9. pp 166–171
6. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464
7. Li K, Peng J, Irwin GW (2005) A fast nonlinear model identification method. IEEE Trans Autom Control 50(8):1211–1216
8. Peng Z, Dehua L, Hong W, Feng C (2011) The automatic model selection and variable width for RBF neural networks. Neurocomputing 74:3628–3637

Chapter 40
Argumentation Pattern: An Approach to Issuing Software Reliability Case

Boxuan Wang and Minyan Lu

Abstract Software is a major component of most important system applications. Because the software component typically provides critical functions, faults in the software may cause the system to fail in a significant way. Such system failures due to direct software faults are what we classify as "software failures." Thus, it is important to use methods and techniques that provide evidence that the software component has been designed, implemented, tested, installed, and, as necessary, updated without faults that might result in undesirable system failures. Originate from safety case, which has been used in the aerospace and nuclear industries, reliability case is one method to make sure that the software's reliability has been reached. This paper briefly reviews some of the previous research in the area of safety case. And then introduce the concept of reliability case. This is followed by a discussion of the approach to issuing reliability case—argumentation pattern. Finally, we draw the conclusion and discuss related future works.

Keywords Software reliability · Reliability case · Argument · Argumentation pattern

40.1 Introduction

Recent thinking with software reliability practice has led to a cooperative approach which should present evidence throughout the project that software reliability requirements are consistent with system level requirements, are achievable, are understood by the development organization, and that ambiguities have been resolved. This approach is known as the "reliability case" and it has some parallels

B. Wang (✉) · M. Lu
School of Reliability and Systems Engineering, Beihang University, No. 37 Xueyuan Road, Haidian, Beijing, China
e-mail: winnerwbx@163.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_40,
© Springer-Verlag Berlin Heidelberg 2014

with, but is not the same as, the better known "safety case" as used particularly in the aerospace and nuclear industries.

The concept of safety cases has been evolving since the mid-1980s. A safety case is a generally accepted practice for reporting data needed by contractual, regulatory, and/or independent third-party certification authorities. A generic safety case is structured to include: system safety requirements and their allocation, assumptions, and/or claims based on preexisting systems, evidence, conclusions, and recommendations. The safety case addresses all components of system safety, including hardware and software. It provides a systematic process for collecting, analyzing, and interpreting data throughout the lifecycle, which can be used as evidence and for monitoring whether or not a project is on track for meeting stated system safety goals.

40.2 The Concept of a Reliability Case

As an equivalent concept with safety case, reliability case is needed to bridge the gap among hardware reliability, software reliability, system safety, and reliability. Accordingly, this paper will present an approach to issuing software reliability cases, based on the international standards SAE JA 1002, Software Reliability Program Standard [1] and SAE JA 1003, Software Reliability Implementation Guide [2], published by the Society of Automotive Engineers (SAE).

The software reliability case captures the assumptions, claims, arguments, and evidence necessary to provide a convincing argument that a software product has a specified reliability. This collection of information, referred to as the Software Reliability Case, must be accurate, current, and complete and presented in a convincing manner in order to obtain sign-off by the customer and/or certifying authority. It consists of six components: system context description; software reliability goals, objectives, and requirements; assumptions and claims; evidence and reasoned arguments to support the claims; conclusions/recommendations; and certification records. Figure 40.1 is an overview of the software reliability case.

There are three types of software reliability case: (1) pre-development software reliability case, (2) development software reliability case, and (3) in-service software reliability case. The pre-development software reliability case provides an analysis of the supplier's approach to reliability achievement. It provides evidence that:

- the individuals and organizations involved are capable of supplying software that is corresponding with the reliability requirements of the proposed system;
- the software reliability plan is appropriate for the proposed system reliability requirements;
- technical proposals and decisions upon which the tender is based are appropriate for the proposed system reliability requirements; and
- the supplier will be able to demonstrate reliability achievement during the project.

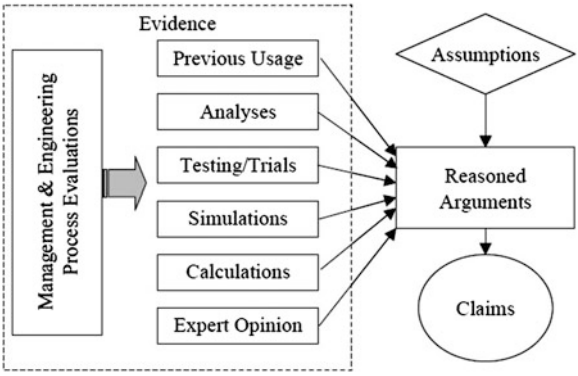


Fig. 40.1 Reliability case: claim based on evidence

During development, the supplier should update the software reliability case with a summary and assessment of the results of the activities that contribute to the reliability evaluation. By the time of acceptance into service, the case should contain the complete set of evidence that the reliability requirements of the software will be met [3]. The software reliability case should describe measurements taken of the software products and the software engineering processes that provide evidence development is proceeding satisfactorily. Versions of the case should be planned for appropriate milestones in the software development lifecycle, for instance after reliability tests and trials. The current version of the case should be presented at design reviews and the outcome included in the case. Versions of the case should be managed under a configuration control mechanism to ensure compatibility with development status.

In-service software reliability management includes the collection of operational data and the maintenance of an in-service software reliability case. The case should contain a description of field experience with the software and an analysis of the impact of software failures on system reliability. The analysis should address the potential consequences of a failure, root causes, mitigation strategies, and the lessons learned for the software engineering process [4].

40.3 Using Argumentation Pattern to Issue Reliability

In software field, the concept of "pattern" is applied in the form of "design pattern," which used by designers as a means of abstracting fundamental design strategies from the details of particular designs [5]. Based on the principles of Design Patterns, particularly the concept of structured documentation together with diagrams, Kelly [6] from University of York has developed the concept of Safety Case Patterns as a means of documenting and reusing successful safety argument structures. He argues that patterns should use diagrams to describe the overall structure of the solution succinctly, and structured supporting text to

Table 40.1 Case pattern elements

Pattern elements	Explanation
Pattern number	The labeled number of pattern for easy identification and management
Pattern name	The pattern name should communicate the key principle or central argument being presented by the safety argument pattern. This will be the label by which people will identify this pattern
Intent	The statement answering what the pattern is trying to achieve and what is being attempted
Motivation	Describes why the pattern was constructed. The motivation can be expressed in terms of previous experiences, problems, etc.
Structure	Using the Goal Structuring Notation to present the structure of the argument pattern. Using the GSN pattern extensions. It can be indicated clearly where the argument is complete, where information must be provided (e.g., where instantiation must occur) and where the argument requires further development
Participants	Augment the structural description by providing a description of each of the elements of the goal structure pattern (i.e., the goals, the contexts, the strategies, the solutions, etc.)
Collaborations	Describe how the different elements of the pattern (sources of contextual information, argument strategies, and goals) work together to achieve the desired effect of the pattern—to present an effective argument
Applicability	This section should record under what circumstance this argument can and should be applied. Of particular concern for safety arguments, it should make clear the assumptions and principles underlying the argument pattern
Consequences	This section should make clear what work remains after having applied or carried out an argument pattern
Implementation	Communicate how the application of this pattern should be carried out; communicate hints or techniques that would ease successful application of the pattern; and record common misinterpretations of the terms or concepts used in the pattern
Examples	Provide examples that illustrate the instantiation of the pattern
Known uses	This section should refer to known uses of the form of argument presented in the pattern
Related patterns	This section should refer to other reliability case patterns that are related to this pattern, e.g., patterns that share the same intent but are admissible under different applicability conditions

document important details of how that pattern may be instantiated, together with underlying rationale [7]. He emphasizes the necessity to consider the following 2 issues:

- Represent (in diagrammatic form) the structure of a generalized safety argument;
- The format and role of the text that should support such a diagrammatic description.

Summing up the above research, this paper defines that:

Table 40.2 GSN multiplicity extensions (for structural abstraction)

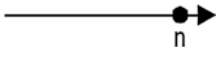
Notation	Explanation
	A solid ball is the symbol for many (meaning zero or more). The label next to the ball indicates the cardinality of the relationship
	A hollow ball indicates "optional" (meaning zero or one)
	A line without multiplicity symbols indicates a one-to-one relationship
These symbols are defined for use as annotations on existing GSN relation types. Multiplicity symbols can be used to describe how many instances of one entity relate to another entity	

Table 40.3 GSN optionality extensions (for structural abstraction)

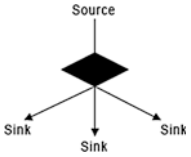
Notation	Explanation
	one source has three possible sinks. Multiplicity relations can be combined with optionality relations. Placing multiplicity symbols prior to the "choice" vertex (squashed diamond) describes a multiplicity over all the optional relations. Placing a multiplicity symbol on individual optional relations describes a multiplicity over that relation only
This symbol is defined for use over the existing GSN relation types. Choice can be used to denote possible alternatives in satisfying a relationship. It can represent 1-of-n and m-of-n selection	

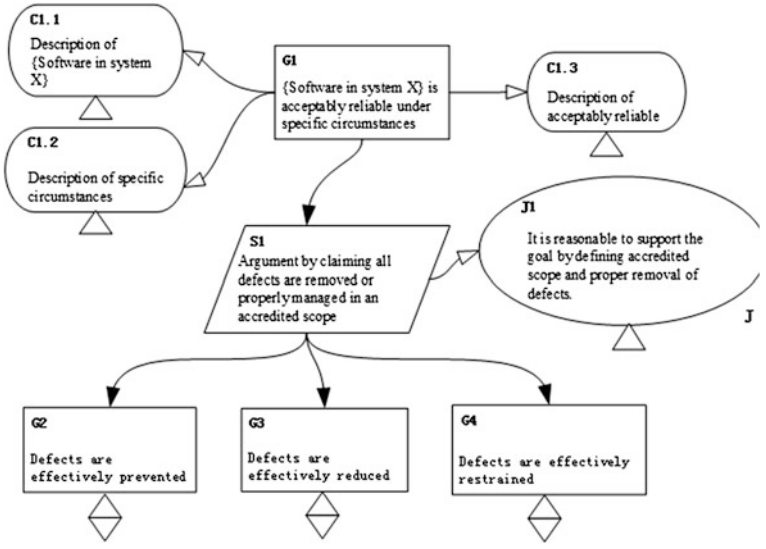
Table 40.4 GSN extensions for entity abstraction

Notation	Explanation
	The Is_A relation provides a basis for the expression of supertype and subtype relations between GSN entities (e.g., "Failure rate is less than 1×10^{-6} per year" Is_A Failure Rate Claim) and can therefore be used to establish type hierarchies
	This notation indicates that the attached entity remains to be instantiated, i.e., at some later stage the "abstract" entity needs to be replaced (instantiated) with a more concrete instance
	This notation indicates that the attached entity requires further development, i.e., at some later stage the entity needs to be (hierarchically) decomposed and further supported by subentities

Argumentation Pattern abstractly describes some basic principles and core argument structures during system or software case construction processes. It contains a number of pattern elements to provide auxiliary information, and through instantiating abstract structure elements, it can be used to build concrete case module structures.

Argumentation Pattern has characteristics of abstraction, reusability, autonomy, collaboration, and so forth. It includes some pattern elements like GSN (Goal Structuring Notation) structure, and provides necessary information to construct argument structures. The reliability case pattern documentation format is as Table 40.1.

Table 40.5 Reliability case argumentation pattern

Pattern number	P1
Pattern Name	Defect-based reliability case top-level argumentation pattern
Intent	The intent of this pattern is to argue the reliability of a system by demonstrating that all defects are removed or properly managed in an accredited scope [11]
Motivation	The motivation for this pattern is the need to decompose a high-level goal (that is difficult to substantiate "as-is") into subgoals that are hopefully easier to address
Structure	

(continued)

Table 40.5 (continued)

Pattern number	P1
Participants	G1 Defines the overall objective of the pattern
	C1.1 Description of {Software in system X}, mainly requirements and design specifications
	C1.2 Description of specific circumstances under which the system runs
	C1.3 Definition and description of acceptably reliable can be from certain standards or regulatory authorities. It is the basis of argumentation structure
	S1 Presents the strategy adopted to support G1
	J1 Provide the justification to support strategy S1
	G2 Subgoal from the aspect that defects are effectively prevented
	G3 Subgoal from the aspect that defects are effectively reduced
	G4 Subgoal from the aspect that defects are effectively restrained
Collaborations	1. C1.1, C1.2, and C1.3 provide essential contextual information and constraints for top-level goal G1 and the argumentation structure below
	2. J1 provides the justification to support strategy S1. It plays an important role in decomposing G1
	3. G2, G3, and G4 declare the argument goal from three different aspects. The G1 is true only if G2, G3, and G4 are all true
Applicability	This is a very general pattern and, as such, has a wide applicability In order to apply the pattern it is necessary to instantiate C1, C2, and C3
Consequences	After instantiating this pattern, three unresolved goals will remain for further expansion: 1. G1: Argument that defects are effectively prevented 2. G2: Argument that defects are effectively reduced 3. G3: Argument that defects are effectively restrained
Implementation	In implementing this pattern it should follow a top-down methodology. Instantiate goals, contexts, and others successively *Possible pitfalls*: Lack of sufficient analysis and negotiation, inappropriate definition of "acceptably reliable," which would affect the confidence of G1 and the argumentation structure follows
Examples	None for now
Known uses	The ideology of gradual decomposition and expand respectively
Related pattern	The future work of patterns of G1, G2, and G3

According to pattern elements' auxiliary information, argumentation pattern instantiate abstract GSN structure elements. And through collaborations between different patterns, it builds a complete structure of reliability case. Compared to other pattern elements, GSN structure is the most critical one [8]. It pictorializes the core argument structure of argumentation pattern and so as to form a special class of argumentation structure model, as known as "argumentation pattern."

The Goal Structuring Notation (GSN) has been developed for the description of safety arguments: relating the breakdown of safety requirements to argument based upon available evidence [9]. GSN can be used to articulate a specific argument [10]. However, to be able to generalize the specific details of reliability argument and represent patterns of argument rather than simply instances the GSN must also support abstraction. Therefore, we must extend the GSN notation elements to support structural abstraction and entity abstraction. The extensions are detailed in Tables 40.2, 40.3 and 40.4.

Using GSN and the extensions provided above, we can now describe Reliability Case Patterns. Table 40.5 is an example of a documented Reliability Case Argumentation Pattern.

40.4 Conclusion

The main contribution of this paper is to present an approach to issuing software reliability case—argumentation pattern. It is a means of supporting and promoting the reuse of successful reliability arguments between reliability cases. By expanding the argumentation pattern provided above, researchers can generate their own reliability case patterns on specific areas. Future work could involve the development of practical examples and the applicability upon software engineering.

References

1. SAE (1998) Software reliability program standard. Warrendale, PA
2. SAE (2004) Software reliability program implementation guide. Warrendale, PA
3. Jones JA, Marshall J, Newman B (2004) The reliability case in the REMM methodology
4. Lipson H (2008: 2011) Assurance cases overview. US Department of Homeland Security
5. Hawkins R, Kelly T (2010) A systematic approach for developing software safety arguments. J System Safety 46(4):25–33
6. Gamma E, Helm R, Johnson R, et al. (1995) Design patterns: elements of reusable object-orientated software. Addison-Wesley, Boston
7. Kelly T (1998) Arguing safety: a systematic approach to managing safety cases. University of York, York
8. Ayoub A, Kim B, Lee I et al (2012) A safety case pattern for model-based development approach: NASA formal methods symposium (NFM). Norfolk, VA

9. Matsuno Y, Taguchi K (2011) Parameterised argument structure for GSN patterns
10. Conner SA, Watts PE (2012) Transformation of thermal ink-jet product reliability strategy
11. Minyan L (2011) Software reliability engineering. National Defense Industry Press, Beijing (In Chinese)

Chapter 41
Control Modeling and Signal Processing of a Library Self-Delivery Robot and Its Applications

Shuping Dang, Meizun Meng, Daniel Mathews and Rinat Kakimzhanov

Abstract Nowadays, the cost of human labor is increasing dramatically in modern society in turn the time in which demands for certain tasks are to be completed is also decreasing. Meanwhile, most people do not have the patience to find out a book from a great number of bookshelves within a vast collection; even if the title of the book has been given. Fortunately, all of books in a library must be collated and organized according to a structure that distinguishes certain topics or genres. As this is a standard procedure for most libraries, this provides the fundamental structure for designing a robot to assist humans in acquiring the books they need. This paper outlines the theories, structure, and model of the robot called "library self-delivery robot (LSDR)," and especially focuses on the signal processing module. Additionally, applicable implementation methods and application aspects will also be discussed briefly.

Keywords Self-delivery · Control modeling · Signal processing

41.1 Introduction

In recent years, most of the libraries around the world have uploaded their hardware facilities to meet the requirements of convenience and economical demands. According to a related report, 85 % of the libraries around the world have more or less self-service facilities, including self-check machines, drive-through windows, vending machines with books, and DVDs [1]. This global tendency provides a positive opportunity for introducing more high-level automatic devices to assist people to administrate and make use of libraries. Owing to the plentiful number of

S. Dang (✉) · M. Meng · D. Mathews · R. Kakimzhanov
School of Electrical and Electronic Engineering, The University of Manchester, Sackville Street Building, Manchester M13 9PL, UK
e-mail: shuping.dang@gmail.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_41,
© Springer-Verlag Berlin Heidelberg 2014

books and other various materials in a library, normally, it is relatively tough to administrate and utilize a library via automatic devices and self-service systems. However, in accordance with the state-of-the-art coding techniques, signal processing algorithms and control techniques, it is entirely possible to construct the fundamental model of a library self-delivery robot (LSDR) to assist librarians and users to administrate and use libraries effectively and conveniently.

In this paper, the detailed structure and model of LSDR will be introduced and elaborated in several aspects, including signal processing, actuator driving along with other devices and components. In addition, the simulation results based on Matlab will be further analyzed in terms of system stability, robustness, and practical applicability.

41.2 Basic Theories and Structure of LSDR

41.2.1 Outline of LSDR

The basic theories of LSDR are relatively simple and accessible. First of all, having a quick glance at the signal processing theory, it is obvious that the input signal of the system is the demand of the user, i.e., which book they would like to borrow from the library? Obviously, the output should be the operation of delivering the specified book for the user. Therefore, input devices must be required, i.e., keyboard or voice receiver to detect the demand from the user. Then, if the input signal now becomes the title of a book, the embedded microcomputer of the system should be able to transfer this name to a stored book code. This will in turn command the LSDR to move to the related bookshelf according to the internal coordinate system. Following this the LSDR is capable of ascertaining and acquiring the book corresponding to the unique code of the book precisely, via a scanning sensor that will read the stored book code. Ultimately it will deliver the book to the user. Therefore, the flow path can be depicted from the procedure described above; a microcomputer, a series of motors, sensors, mechanical arms, as well as some audio-visual devices are required. The associated techniques are signal detection, signal processing, motor driving, image identification, and mechanical arms control. Thus, the structure of signal flow can be concluded and shown in Fig. 41.1.

Figure 41.1 shows a user inputs a requirement of a certain book to LSDR, i.e., asks for the service by giving the name of the book. Then LSDR will transfer the characters inputted by the user into a digital stored book code, which can be read by the electronic devices. Then the location of the book will be returned to the embedded microcomputer in terms of wireless signal from a cloud database, the purpose of this signal is to tell the robot where the book is situated and what is the optimal path to take in order to reach this book. Concurrently, the control signal will drive the motors of LSDR to the specified bookshelf, and make use of the image identification sensors to scan and detect the code labels attached in the

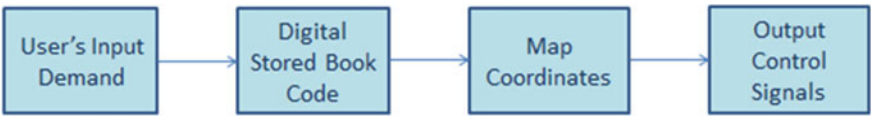


Fig. 41.1 Structure of signal flow of LSDR

books according to a certain algorithm, and then switch on the mechanical arms to obtain the book. Eventually, LSDR will come back to the start point and delivers the book to the user.

41.2.2 Location Technique

Once the LSDR has received the input demand, it will firstly check the current position in which the LSDR is situated and send these coordinates back to the cloud database, which will check the position of destination, and optimize the path to it. For example, assuming that demand for book 005 is inputted to the robot, then it will obtain the code of the book 001 is 005(01), which means the book is the fifth book on the first bookshelf, and then the location information of first bookshelf, e.g., (a, b) will be returned from the cloud database. Consequently, LSDR can decide where to go and find this book in terms of the designed a path sent from the cloud database as shown in Fig. 41.2.

41.2.3 Scamming and Acquiring

Then, after arriving at the bookshelf, LSDR will commence to find the label of the bookshelf at the interval of $(a \pm \Delta x, b \pm \Delta y)$, where Δx and Δy represent the coordinate fault-tolerant values in two-dimensional coordinate system, respectively [2]. The purpose of setting these fault-tolerant values is to make sure the label of the bookshelf can be detected in complex cases, e.g., artificial disturbance, so that the robustness of this system will be enhanced along with the efficiency of the book selection.

Furthermore, if the bookshelf has been confirmed correct, then it will commence to scan the digital stored book codes row by row in an established sequence for the purpose to find out the code of 005(01).

If the code scanned matches the digital code stored in the RAM of LSDR, then mechanical arms will move to clamp this book. There are three mechanical arms, and the method to clamp a book can be concluded as the three steps below and shown in Fig. 41.3.

- Step 1: press the book and make it tilt outward by the first mechanical arm.
- Step 2: clamp the book tightly by the second and third mechanical arms.
- Step 3: pull out the book by three mechanical arms.

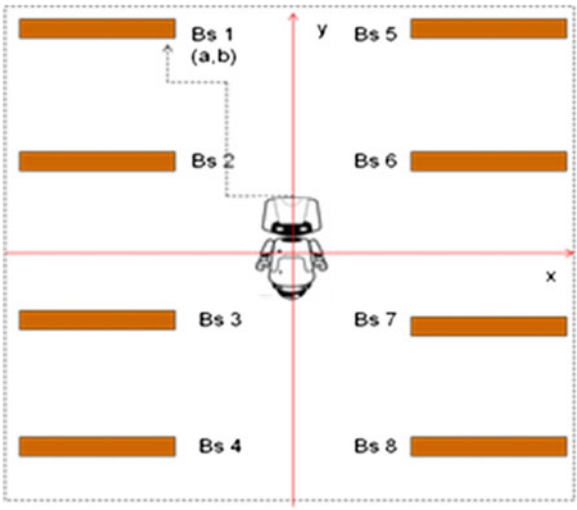


Fig. 41.2 Demonstration of the algorithm to find the bookshelf position

41.2.4 Control Modeling Constructions and Signal Processing

Depending on the structure and functions of the LSDR, the core modules of it are obviously the location identification and actuator driving. For the reason that the location identification is normally supported by the cloud database with respect to feedback transmitted via wireless network in the library, thus it could be more precise and effective to transmit and feedback the data to the LSDR. Hence, in this part, we mainly analyze the model and signal processing inside the control system of the LSDR apart from external input processing. Under the assumption that the input analog signal has been transmitted precisely applying a classical feedback control loop [3] can be used to achieve the goal of control of the mechanic arms shown in the Fig. 41.4.

- **Input X(s)**: The amplitude value of the step function represents the different input position information.
- **Controller C(s)**: To keep the LSDR universal [4], classical PID control strategy has been utilized in order to obtain stable, efficient response, and eliminate overshot. Where k_d, k_p, k_i, and γ are the control parameters needed to be further discussed later.
- **Transfer function G(s)**: The transfer function represents the behavior and operation of mechanical arms in one-dimension. Where a, b, and c are the control parameters needed to be further discussed later.

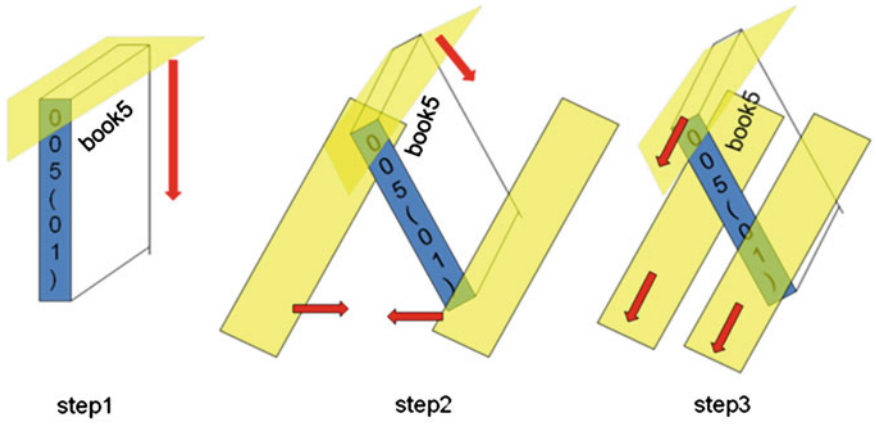


Fig. 41.3 Steps to clamp a book

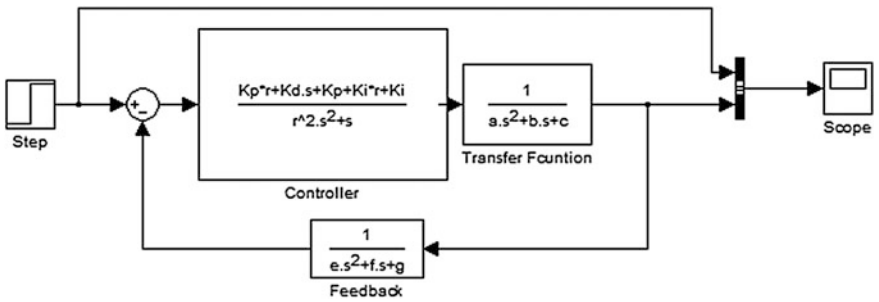


Fig. 41.4 Block diagram of the control model of LSDR

- **Feedback function F(s)**: Feedback with second-order function. Where e, f, and g are the control parameters needed to be further discussed later.
- **Output Y(s)**: It represents the actual position of the mechanical arms in one-dimension [5].

Depending on this block diagram, their mathematical input-to-output relationship can be obtained by

$$[X(s) - F(s)Y(s)]C(s)G(s) = Y(s) \tag{41.1}$$

Then the closed-loop transfer function can be obtained as

$$T(s) = \frac{C(s)G(s)}{1 + C(s)G(s)F(s)} \tag{41.2}$$

Then selecting these parameters below, and thus the simplified closed-loop transfer function is

$$\begin{cases} k_d = \gamma = 0 \\ k_p = k_d = K \\ a = 1 \\ b = 2 \\ c = 1 \end{cases} \qquad T(s) = \frac{Ks + K}{s^2 + (K+2)s + (K+1)} \tag{41.3}$$

Obviously, the parameter remained in the transfer function is the only variable needed to adjust in order to meet the specifications [6].

To create a stable and oscillation-free system, the requirement below must be satisfied:

$$\begin{cases} \mathrm{Re}\left\{ \dfrac{-(K+2) \pm \sqrt{(K+2)^2 - 4(K+1)}}{2} \right\} < 0 \\ \xi = \dfrac{K+2}{2\sqrt{K+1}} > 1 \end{cases} \Rightarrow K > 1 \tag{41.4}$$

Thus, for the purpose to obtain a proper set point, choosing K = 200 gives

$$T(s) = \frac{200s + 200}{s^2 + 202s + 201} \tag{41.5}$$

41.3 Simulations and Analysis

Mathematic model of the electrically controlled system can be constructed by the Matlab statements and after executing the Matlab statements above, the according step response simulation is shown in Fig. 41.5.

From the plot of step response, the settling time is approximately 0.018 s, which is relatively fast enough for processing and responding to an input demand [7, 8].

Moreover, Overshot and oscillation have been eliminated in this simulation assuming K is 200, which in fact implies that the system is a stable and oscillation-free system.

In addition, the steady-state error is only 0.5 %, and it is acceptable for a qualified steady-state response value.

41.4 Implementation

The algorithm of scanning the digital codes of books is illustrated in Fig. 41.6.

To confirm the practical applicability of LSDR, the case study of the Joule Library (Manchester, UK) has been investigated, and the necessary conditions for finding out a specified book are satisfied and shown in Fig. 41.7.

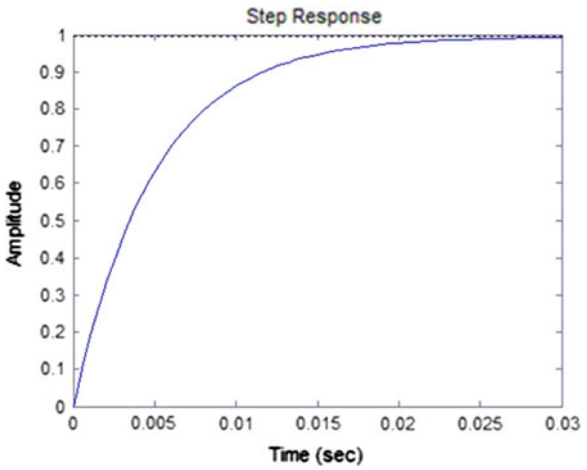


Fig. 41.5 Plot of step response of the system

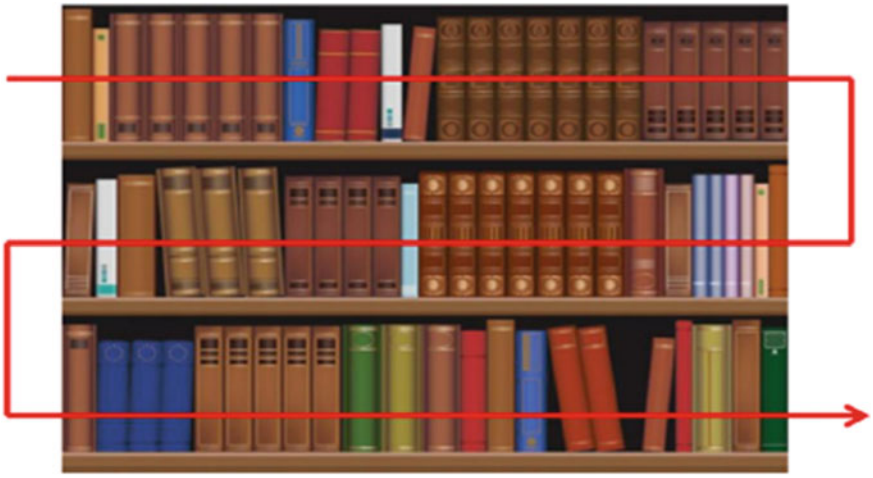


Fig. 41.6 Sequence to scan the stored book codes in the bookshelf

As the Fig. 41.7 shown, it can be clearly seen that each bookshelf has a unique code and each book also has a specified code, which can provide the distinct identification signal for LSDR. Furthermore, if a more accurate identification signal for practical application is desired to be provided, the two-dimensional code will be a cheap and easy way to achieve the goal.



Fig. 41.7 Bookshelves and books in the Joule library, Manchester, UK

41.5 Conclusion

To sum up, the LSDR would help people save both time and expenditure to find out a book needed as well as administrate a modern library. As the analysis above shows the techniques corresponding to this application is relatively mature now and with the aid of some existing modules combined with a series of computer codes a simple prototype can be created. In addition, owing to its low cost and easy construction approach, it is useful and attractive for most libraries. Predictably, if it is implemented the time for users to find a book will be reduced sharply, and the degree of modernization and automation of our libraries will be increased significantly.

References

1. Dempsey B (2010) Do it yourself libraries. Libr J: 24–28
2. Sun T, Zhou D (2010) Automatic identification technology and application of two-dimensional code. In: 2011 IEEE international conference on automation and logistics, pp 164–168
3. Dorf RC, Bishop RH (2000) Modern control system, 9th edn. UK
4. Heath W (2013) Lecture notes for 2nd year EEEN20030 in control system I. School of Electrical and Electronic Engineering, The University of Manchester
5. Junjie Z, Tianfei S et al (2010) Research of mechanical arm real-time motion control. In: E-business and e-government, pp 4321–4323
6. Dudek P (2012) Lecture notes for 2nd year EEEN20023 (2012/3) of digital system design II. School of Electrical and Electronic Engineering, The University of Manchester
7. Hu Z (2012) Lecture notes for 2nd year EEEN20027 (2012/3) of signals and systems. School of Electrical and Electronic Engineering, The University of Manchester
8. Alsusa E (2012) Lecture notes for 2nd year EEEN20031 (2012/3) of analogue and digital communication. School of Electrical and Electronic Engineering, The University of Manchester

Chapter 42
Analysis of Moore's Law on Intel Processors

Jia Chen

Abstract Moore's law is a famous law in the field of computer hardware, stating that the number of transistors on integrated circuits doubles approximately every 2 years. In this report, we do an analysis of Moore's law on consumer-class Intel processors to see how well Moore's law has been doing on predicting the number of transistors of Intel processors, and whether this trend will continue in the future. We take a statistical approach by using linear regression and extrapolation. By matching actual data against the model of Moore's law and calculating r-square, we are able to conclude that although Moore's law has made close predictions for single-core processors, the time interval for a double in multi-core processors is actually longer than Moore's law, thus Intel will fall behind Moore's law in the future.

Keywords Moore's law · Transistor · Processor

42.1 Introduction

As a famous law in the field of computer hardware, Moore's law has never failed to become one of the topics whenever Intel or AMD announces its new generation of processors. Moore's law states that the number of transistors on integrated circuits doubles approximately every 2 years. On one hand, Moore's law keeps inspiring Intel innovation and the forecast of Moore's law drives groundbreaking new technologies and more power-efficient processors. Intel claims that it has maintained the pace of Moore's law for decades [1]. On the other hand, in a recent conference, Intel said that keeping up with Moore's law is

J. Chen (✉)
Department of Computer Science, College of Engineering, University of Missouri, Columbia, MO, USA
e-mail: jckg7@mail.missouri.edu

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_42,
© Springer-Verlag Berlin Heidelberg 2014

becoming a challenge [2]. Such opinions as "Moore's law is dead" have been around for years. Thus, two questions arise: How well has Moore's law been doing on predicting Intel processors, and whether Intel is likely to maintain this trend in the years to come.

In this report, we do an analysis of Moore's law on Intel consumer processors and attempt to answer these two questions. Previous works on Moore's law have taken different approaches. Mack [3] focuses on identification and quantification of various drivers of Moore's law. Rupp and Selberherr [4] focus on economy and states a barrier of Moore's law when fabrication cost reaches total revenue. Powell [5] focuses on physics and states that Heisenberg uncertainty would be the fundamental limit. In this report, we decide to go back to the origin of Moore's law by taking a statistical approach similar to the one in [6], in which Moore's law first appeared. We use linear regression and extrapolation. By matching actual data against the model of Moore's law and calculating r-square, we find that Moore's law has indeed made close predictions on the number of transistors till 2005, till the introduction of multi-core processors. We also find out that after 2005 the time interval for a double in the number of transistors has been longer than that of Moore's law. Finally, we believe that Intel will fall behind the prediction of Moore's law in the future.

42.2 Background

Moore's law is named after Gordon E. Moore, a cofounder of Intel. In sum, Moore's law is a prediction and it went through one proposition in 1965, one revision in 1975 and gradually got its commonly accepted form. Moore first described the trend in his 1965 paper [6].

> The complexity for minimum component costs has increased at a rate of roughly a factor of two per year. Certainly over the short term this rate can be expected to continue, if not to increase.

As an observation that still holds for today, at a given time, the number of components on an integrated circuit with one unit size may vary, but as the number of components increase, the unit price of a component always follows a U-shape curve. At a certain number or complexity, unit component cost shall reach its minimum, and in his 1965 paper, Moore stated that this complexity for minimum component cost had increased by a factor of two per year, and he expected this rate to continue. However, in his 1975 paper [7], Moore altered his prediction to a double every 2 years instead of one.

> With this factor disappearing as an important contributor, the rate of increase of complexity can be expected to change slope in the next few years. The new slope might approximate a doubling every 2 years, rather than every year, by the end of the decade.

In this 1975 paper, Moore divided the contribution to the increase into three parts: larger chip area, higher component density, and more circuit cleverness.

He suggested a limit to the contribution of circuit cleverness, thus setting the rate down. Circuit cleverness describes how clever we can arrange the components on an integrated circuit.

Moore's papers at 1965 and 1975 set the foundation, and over time they evolve into the commonly accepted Moore's law, which states that the number of transistors on integrated circuits doubles approximately every 2 years.

42.3 Methods

In order to see how well Moore's law has been doing in the past, we extract a linear model from Moore's law and set both the time interval between doubling and the initial number of transistors as variables. We give two sets of values to the variables to create two specific models. In the first model we set the variables to represent Moore's law, and in the second one we calculate their optimal values by fitting actual data. Our primary concern is that the introduction of multi-core processors may have an impact on Moore's law, so we use data of processors from 1971 to 2002. We then calculate r-squares to see how close these two models are to the data and make a comparison.

For the second question: whether Intel is likely to maintain this trend in the future, we collect data of multi-core processors from 2005 to 2012, concatenate it with the data from 1971 to 2002 and fit it onto the linear model of Moore's law to get the global trend. Then we do another fitting using multi-core processors only. We calculate r-squares and compare among the results.

42.4 Results

42.4.1 Modeling Moore's Law

In order to build a linear model, we remove the uncertainty of Moore's law by stating that the number of transistors on integrated circuits doubles every 2 years. Then we can transform this modified Moore's law into a functional form

$$y = a \times 2^{\frac{x}{2}} \tag{42.1}$$

in which y is the number of transistors, x is the number of years starting at 0, and a is the number of transistors at year 0.

We generalize (42.1) into an exponential form

$$y = a \times 2^{bx} \tag{42.2}$$

and apply $\log_2$ to both sides of (42.2) to get a linear form

$$\log_2 y = \log_2 a + bx \tag{42.3}$$

which is ready for data fitting. Similar models for Moore's law can be found in [5, 8].

42.4.2 Fitting Data and R-Square from 1971 to 2002

In order to verify how well Moore's law fits into Intel processors, we build two specific models using (42.3) based on a dataset. The dataset we use is an Intel press release, illustrating the history of Intel processors from the 4004 processor in 1971 to the quad-core Core 2 Extreme processor in 2007. We choose two fields for each processor, the year of introduction and the number of transistors. On the selection of processors, we take two issues into consideration:

- The dataset includes many nonconsumer processors, such as the ones under Itanium and Xeon. Since our focus is on consumer processors, we decide not to use data from nonconsumer processors.
- The introduction of Pentium D processor acts as a gap between single-core processors and multi-core processors. The first Smithfield Pentium D processors have nearly twice the transistors as contemporary Pentium 4 processors, because a Smithfield Pentium D essentially consists of two Pentium 4 cores. Also, the introduction of Core microarchitecture in 2006 may have an impact on the number of transistors. Based on these uncertain factors, we decide to use the data before Pentium D, from 1971 to 2002, to verify the closeness of Moore's law to the data.

Details of the Intel press release dataset and the data we select is in Appendix A. In order to linearize our data, we take $\log_2$ to the number of transistors. Also to match our model, we subtract the year by 1971 to set the first year to 0.

We build the first model to represent Moore's law. We set b to 0.5, representing a double every 2 years. Also according to the dataset, we set a to 2300, which is the number of transistors at year 0. Thus, we get

$$y = 2300 \times 2^{\frac{x}{2}} \tag{42.4}$$

which is

$$\log_2 y = 11.167 + 0.5x \tag{42.5}$$

in linear form.

Then we build a second model to be the optimal model for the dataset. We apply linear regression on (42.3) as in Fig. 42.1 using MATLAB function *polyfit* and get the optimal fit

$$\log_2 y = 11.366 + 0.457x \tag{42.6}$$

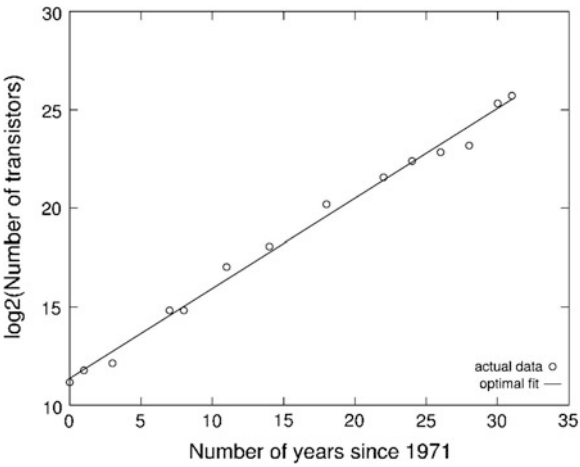


Fig. 42.1 $\log_2$ of number of transistors versus the number of years, both actual data and optimal fit, from 1971 to 2002

which is

$$y = 2638.6 \times 2^{\frac{x}{2.1882}} \tag{42.7}$$

Then we calculate r-square by

$$1 - \text{SSresid}/\text{SStotal}$$

where SSresid is the sum of the squared residuals from the regression, and SStotal is the sum of the squared differences from the mean of the dependent variable [9]. The r-square for (42.6) linear optimal model is 0.992, and the r-square for (42.5) linear Moore's law is 0.973.

42.4.3 Fitting Data and R-Square from 2003 to 2012

While looking into the question of whether Intel is likely to maintain Moore's law in the future, we notice a possible impact multi-core processors may have on Moore's law. Here, we propose two hypotheses:

1. With multiple cores, the total number of transistors in processors shall follow Moore's law.
2. Despite having multiple cores, the number of transistors per core shall follow Moore's law.

We apply (42.7) and make predictions for processors starting from 2003 to 2012. Then we retrieve data on the number of transistors for contemporary consumer Intel processors. The data we use are mainly from the ARK of Intel, and details of the data are in Appendix B.

Table 42.1 Prediction of Moore's law, actual total transistors, and transistors per core of Intel processors from 2003 to 2012

Year	Prediction	Actual total	Actual/Core	Processor codename	Core count
2003	66626000	77000000	77000000	Banias	1
2004	91457000	125000000	125000000	Prescott	1
2005	125540000	230000000	115000000	Smithfield	2
2006	172330000	291000000	145500000	Conroe	2
2007	236550000	582000000	145500000	Kentsfield	4
2008	324710000	731000000	182750000	Bloomfield	4
2009	445730000	774000000	193500000	Lynnfield	4
2010	611850000	382000000	191000000	Clarkdale	2
2011	839880000	1160000000	290000000	Sandy Bridge	4
2012	1152900000	1400000000	350000000	Ivy Bridge	4

Table 42.2 Dataset used in Sect. 42.4.2, the year of introduction and the number of transistors of Intel processors, from 1971 to 2002

Year	Transistor
1971	2300
1972	3500
1974	4500
1978	29000
1979	29000
1982	134000
1985	275000
1989	1200000
1993	3100000
1995	5500000
1997	7500000
1999	9500000
2000	42000000
2002	55000000

Introduction of the first dual-core processor Smithfield Pentium D happened in 2005, and the switch to Core microarchitecture happened in 2006, with the introduction of Conroe Core 2 processors. Table 42.1 roughly shows that between 2003 and 2004 the total number of transistors followed the prediction well, but in 2005 the dual-core Smithfield doubles the number of transistors from last year's single-core Prescott. Between 2005 and 2009, the actual number of transistors is close to the prediction times two. And in 2011 and 2012, the actual number of transistors is close to the prediction of Moore's law again. As to the number of transistors per core, this number has been increasing since the introduction of Core microarchitecture in 2006, but it is always much smaller than the prediction. Thus, we see that none of the two hypotheses have a close description of the trend.

In order to get a global trend, we do a fitting on data from 1971 to 2012. We concatenate data from Tables 42.2 and 42.3 and apply linear regression on (42.3) again. We get

Table 42.3 Dataset used in Sect. 42.4.3, year of introduction, core count, and transistor count of Intel processors, from 2003 to 2012

Year	Codename	Core	Transistor
2003	Banias	1	77 million
2004	Prescott	1	125 million
2005	Smithfield	2	230 million
2006	Conroe	2	291 million
2007	Kentsfield	4	582 million
2008	Bloomfield	4	731 million
2009	Lynnfield	4	774 million
2010	Clarkdale	2	382 million
2011	Sandy Bridge	4	1.16 billion
2012	Ivy Bridge	4	1.4 billion

$$\log_2 y = 11.222 + 0.472x \tag{42.8}$$

which is

$$y = 2388.2 \times 2^{\frac{x}{2.1167}} \tag{42.9}$$

and the r-square for (42.8) is 0.993. Figure 42.2 shows both (42.6), the optimal fit from 1971 to 2002, and (42.8).

Next, we do a fitting on multi-core processors. The data for 2010 are not increasing, so we remove it from fitting. We select the rest of the data from 2005 to 2012, using the year and the total number of transistors and fit them on (42.3). We get

$$\log_2 y = 28.020 + 0.362x \tag{42.10}$$

which is

$$y = 272210000 \times 2^{\frac{x}{2.759}} \tag{42.11}$$

and the r-square for (42.10) is 0.925. Figure 42.3 shows the three optimal fits we have so far, (42.6), (42.8), and (42.10).

42.5 Discussion

For the first question: how Moore's law has been doing in the past. At the end of Sect. 42.4.2, we calculate two r-squares, 0.992 for linear optimal model and 0.973 for linear Moore's law model. We can see that Moore's law predicts 97.3 % of the variance in the dependent variable, of the number of transistors, and this is quite close to 99.2 % by the optimal model. Thus, we conclude that Moore's law fits the actual data of Intel single-core processors quite well, ranging from the 4004 processor in 1971 to Pentium M processor in 2002.

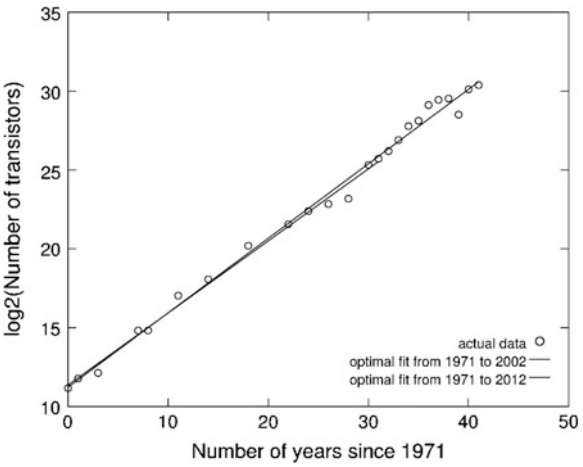


Fig. 42.2 $\log_2$ of number of transistors versus the number of years, actual data, and two optimal fits, one from 1971 to 2002 and the other from 1971 to 2012

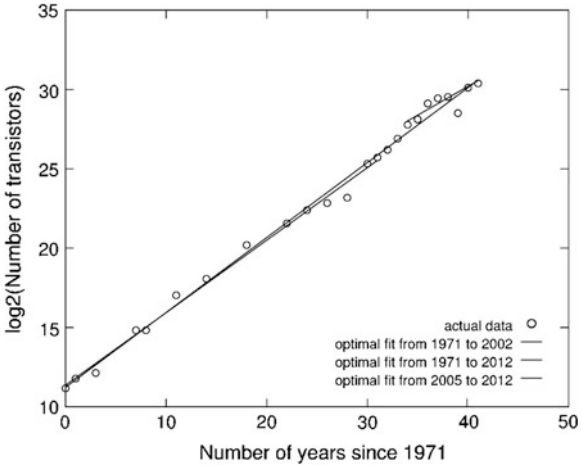


Fig. 42.3 $\log_2$ of number of transistors versus the number of years, actual data, and three optimal fits, one from 1971 to 2002, one from 1971 to 2012, and one from 2005 to 2012

For the second question: whether Intel is likely to maintain the trend of Moore's law in the future. R-square for the global fit is 0.993, but in Fig. 42.2 we can see that most of the actual data points beyond 2005 are above the optimal fit. Considering the introduction of multi-core processors in 2005, we believe a separate fit beyond this point is necessary. R-square for the fit between 2005 and 2012 is 0.925, which is quite close to 1. However, the optimal time interval for a double is 2.759, resulting in the smallest slope in Fig. 42.3, and this is quite far from Moore's law's double every 2 years. So even though by using multiple cores we started out higher than the prediction of Moore's law, as the time interval for a double has increased, the prediction will catch up with the total number of transistors, and this explains the data for 2011 and 2012. Thus, we conclude that with a double every 2.7 years from 2005 to 2012, Intel is unlikely to maintain Moore's law in the future.

Coming back to previous works on Moore's law. Mack [3] concludes in general that the industry has enjoyed 50 remarkable years of Moore's law, but there are unlikely to be many more years left. Rupp and Selberherr [4] say that we will soon deplete the economic driver, thus it will take more years till we hit a fundamental physical barrier for Moore's law. Powell [5] gives a prediction that we will meet the fundamental limit of Heisenberg uncertainty to Moore's law in 2036. Our general idea that Intel will not catch up with Moore's law agrees with those of previous works.

42.6 Conclusion

In this report, we do an analysis of Moore's law on Intel consumer processors, and come to two conclusions.

1. In the times of single-core processors, the number of transistors follows Moore's law quite well, and we are able to get an r-square of 0.973 which is quite close to 1.
2. In the times of multi-core processors, the time interval for a double has been 2.7 years from 2005 to 2012 compared to 2 years of Moore's law, and although Intel is keeping up for now, it will fall behind Moore's law in the future.

Moore's law is often regarded as a self-fulfilling prophecy [10], but if Intel shall fail to follow Moore's law in the future due to economic or technical reasons, Intel may decide not to focus on Moore's law in press releases, and Moore's law may eventually fade out of our sight.

Possible improvement of our work may include taking more factors into consideration. The introduction of Core microarchitecture in 2006 and the introduction of HD Graphics in 2010 may also have impacts on transistors. Also based on our conclusion, we see a possible future work of building a more complicated model to predict the number of transistors. Additional factors may include economic and physical variances. The foundation of Moore's law is linear regression. By taking a more systematic approach, we may eventually get an adaptive prediction model.

A.1 Appendix A

This appendix shows details on the data we use in Sect. 42.4.2. The dataset we use is an Intel press release entitled "The Evolution of a Revolution," it is an overview of Intel processor history. We retrieved it at:

http://download.intel.com/pressroom/kits/IntelProcessorHistory.pdf

After removing all data after Pentium D (Pentium D is also removed) and all data from nonconsumer processors, the data we use for fitting are in Table 42.2.

A.2 Appendix B

This appendix shows details on the data we use in Sect. 42.4.3. We retrieve the data in Table 42.3 from the ARK of Intel, all except the last two entries for 2011 and 2012.

http://ark.intel.com/

We retrieved the entries for 2011 and 2012 from a press report on ANANDTECH.com

http://www.anandtech.com/show/4798/ivy-bridge-148b-transistors

References

1. Moore's Law Inspires Intel Innovation, Intel. http://www.intel.com/content/www/us/en/silicon-innovations/moores-law-technology.html. Accessed 24 May 2013
2. Intel: keeping up with Moore's law is becoming a challenge | PCWorld, PCWorld. http://www.pcworld.com/article/2038207/intel-keeping-up-with-moores-law-becoming-a-challenge.html. Accessed 8 May 2013
3. Mack CA (2011) Fifty years of Moore's law. IEEE Trans Semicond Manuf 24(2):202–207. doi:10.1109/TSM.2010.2096437
4. Rupp K, Selberherr S (2011) The economic limit to Moore's law. IEEE Trans Semicond Manuf 24(1):1–4. doi:10.1109/TSM.2010.2089811
5. Powell JR (2008) The quantum limit to Moore's law. Proc IEEE 96(8):1247–1248. doi:10.1109/JPROC.2008.925411
6. Moore GE (2006) Cramming more components onto integrated circuits. Reprinted from Electronics 38(8):114 Ff (19 April 1965). IEEE Solid-State Circ Soc Newsl 11(5):33–35. doi:10.1109/N-SSC.2006.4785860
7. Moore GE (1975) Progress in Digital Integrated Electronics. In: International electron devices meeting, vol 21, pp 11–13
8. Nagy Béla et al (2013) Statistical basis for predicting technological progress. PLoS ONE 8(2):e52669. doi:10.1371/journal.pone.0052669
9. Linear Regression—MATLAB & Simulink. http://www.mathworks.com/help/matlab/data_analysis/linear-regression.html. Accessed 29 May 2013
10. Disco C, van der Meulen B (1998) Getting new technologies together: studies in making sociotechnical order (Walter de Gruyter 1998), pp 206–208

Chapter 43
A Reliable QoE-aware Framework for Cloud Service Monitoring and Ranking

Yuchao Zhang, Hongfu Liu, Bo Deng and Fuyang Peng

Abstract With the popularity of cloud services, consumers are faced with the optimal selection among a diversity of services from federated clouds. Since the consumers might have individual requirements on applications, the selection of service has become embarrassed. To solve this problem, the research on effective approaches of ranking service to ensure the selection has been put forward as a demanding work. In this paper, we present a reliable QoE-aware cloud service ranking framework based on Markov chain model, which integrates QoE metrics to optimize the rank results. Our approach not only focuses on the individual rank, but also introduces the prediction model to assure the reliability of rank.

Keywords QoE · Rank · Reliability · Cloud service

43.1 Introduction

Cloud computing has attracted a lot of vendors to migrate their traditional application services to cloud in recent years [1]. The so-called "XaaS" paradigm of cloud not only reduces the silo investment from vendors, especially for those small and medium enterprises (SME), but also provides consumers with more opportunities to choose those services which satisfy them most. Every coin has two sides. When facing such a huge mass of services, consumers would always suffer from making decisions for service selection. Moreover, with the popularity of federated cloud, such as Cloud@Home [2], service selection would attract much more attention from the consumers during the process of deploying their cloud applications.

Y. Zhang (✉) · B. Deng · F. Peng
Beijing Institute of System Engineering, Beijing, China
e-mail: dragonzyc@163.com

H. Liu
Beihang University, Beijing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_43,
© Springer-Verlag Berlin Heidelberg 2014

In terms of previous service selection, quality of services (QoS) has played an important role for making decisions, which can be described in different dimensions. However, the QoS cannot be used for reflecting feedbacks from users outside clouds. Since users engaging in different countries might have individual requirements and evaluations for the same services; we need to take the user experience into consideration when making selection. Quality of experience (QoE) [3], which is widely used to describe the overall performance of a system from the views of users, could be one of the candidates cutting to this point. The QoE is often employed for evaluation in the field of multimedia services.

Prior to service selection, we should conduct the service monitoring and ranking based on the various QoE criteria. As mentioned in the researches [4], the monitoring and ranking results tend to be unreliable because of the dynamics in cloud and the subjective evaluations from different users. Hence, we also need to introduce the probability prediction model to improve the reliability and accuracy.

In this paper, we present a reliable QoE-aware cloud service ranking framework based on Markov chain model, which integrates QoE metrics with the user experience to optimize the rank results. Our approach initiates to employ the user evaluation, similar to the WOT [5], into the service ranking. In order to assure the reliability of the rank results, a prediction model based on Markov chain is introduced as well. Services in cloud can be various, including application services, web services, and utility computing services [6]. Here, we do not distinguish them.

The rest of our paper is organized as follows. Section 43.2 formulates the monitoring and ranking problem in our context. The design of our QoE-aware cloud framework and its improvement of reliability based on a Markov model are respectively described in Sects. 43.3 and 43.4. Section 43.5 discusses some related works on cloud service monitoring and ranking. Section 43.6 presents the conclusion and introduces some future works.

43.2 Problem Formalization

To fully evaluate the overall performance, we need to deploy agents in both user-end and cloud respectively. The agent in the user end would be employed for collecting feedbacks from user for each service transaction, while the agent in the cloud end would be in charge of monitoring the QoS relative parameters. On some occasions, a couple of services would be credited with the identical rank level, as they might be invoked by unified workflows, like the Web services flows.

43.2.1 Definition of QoE Criteria

Compared with the previous QoS-driven evaluation methods, the QoE-driven methods can better describe the overall performance of cloud services. As the QoE

criteria consisted of the objective as well as the subjective ones, it could be monitored in different manners. First, we can make use of the objective mapping from QoS criteria to evaluate the QoE criteria. When the QoS parameters are observed, we could easily derive the QoE from them. Secondly, we can employ the subjective user experience directly to replenish the above mapping method. On this occasion, we should design such a user ranking agent deployed in the user end that the QoE could be evaluated in a direct way.

Here, we formally define S_{ij} to denote the QoS value of the service j from cloud vendor i. Thus, we could get a matrix S_x to represent the relationship between the services of user x and its cloud vendors. We assume that those services located at the same column in the matrix have similar functions to some extent. On account of S_x from a specific user x, S_{ij} might be a vector, each element of which stands for a QoS parameter. For example,

$$S_{ij} = <\mathrm{qos}_0,\ \mathrm{qos}_1,\ \mathrm{qos}_2,\ \ldots,\ \mathrm{qos}_n>,$$

where qos_0, qos_1 and qos_n may respectively describe the response time, throughput, and packets drop rate of the service S_{ij}.

43.2.2 Formulation of User Preference

We define user preference ranking to reflect the satisfaction and even trust for each service from users' point of view. As the user preference ranking is evaluated by users from different zones, the preference of different users for the same service would be a great disparity. The more a service satisfies a user, the higher the level it would get. The user preference for each service could be categorized into five different levels, namely the highest, high, medium, low, and the lowest levels. As for the highest ranking level, users might feel most satisfactory for the overall performance of this service, and vice versa.

Similarly, for each user U_{xn}, we also define a matrix R_x to denote the user preference ranking and its element r_{ij} represents the ranking value for service j in cloud i. The r_{ij} is an integer value from set $\{-2, -1, 0, 1, 2\}$, where 2 stands for highest preference level, while -2 stands for the lowest. Thus, a possible matrix R_x might be as follows:

$$R_x = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 2 & -1 & -2 \\ 1 & -1 & 1 & 0 \\ -2 & 0 & 2 & 2 \end{pmatrix}$$

The value r_{ij} is an overall evaluation for the service. As it is difficult for users to perceive the different ranking dimensions, respectively, we should not require them to evaluate a service according to its different QoE metrics. The symbol U_{xn} above defines the user x of group n, in which the users nearly have the same preference.

In order to derive the final ranking results for selection, we would also define a rank matrix CQ, which is a combination of both user ranking and cloud ranking results. In the same way, its element c_{ij} denotes the overall rank of service j in cloud i. Therefore, our goal is to design a framework for deriving CQ from R and S.

43.3 QoE-aware Cloud Framework Overview

43.3.1 Framework

As mentioned before, QoE-aware cloud functioning as a middleware comprises three main components on both sides: monitoring module, evaluating module, and ranking module. After the above procedures, a combination of both ranking results would be used to make decisions. Our framework is presented in Fig. 43.1.

First, from the point of view of the user end, the monitoring module works as a collector to gather the feedbacks of preference ranking results for each service. This module would provide an open interface for users to easily submit their individual ranking values, similar to WOT [5]. Since different users might focus on different services and even distinguished QoE dimensions, we should take the influence of individuality into consideration. Besides, the evaluating module is responsible for refining the ranking results to divide users into different groups, so that users who have similar requirements would be clustered into the same group. After that, once a user selects a new service, we would first identify its promising group so as to improve the selection efficiency. To reduce the overheads during data updating and exchanging, our ranking module would update the matrix R_x every t_1 interval, which is set by the users autonomously.

Secondly, for cloud end, the monitoring module is exploited by each user to monitor its services in various clouds. The monitoring results would be recorded in matrix S_x. *For* every t_2 interval ($t_2 > t_1$), evaluating module would recalculate the QoE values based on different user groups. Once the new QoE values are obtained, ranking module would update the ranking results for each user group. We project our framework in open API to attract the involvement of more ordinary users. Accordingly, participants would have access to the desirable ranking as a return.

43.3.2 Clustering Algorithm in the User End

It is obvious that ranking matrixes R_x of similar users are almost the same. Furthermore, they might have analogous experiences for identical services. When acquiring the ranking matrixes R_x, we could distinguish them into different clusters automatically with clustering approaches. In this case, our user preference ranking matrixes could be regarded as the image. Each element in matrix R_x could be regarded as a pixel of image, which can be described as follows:

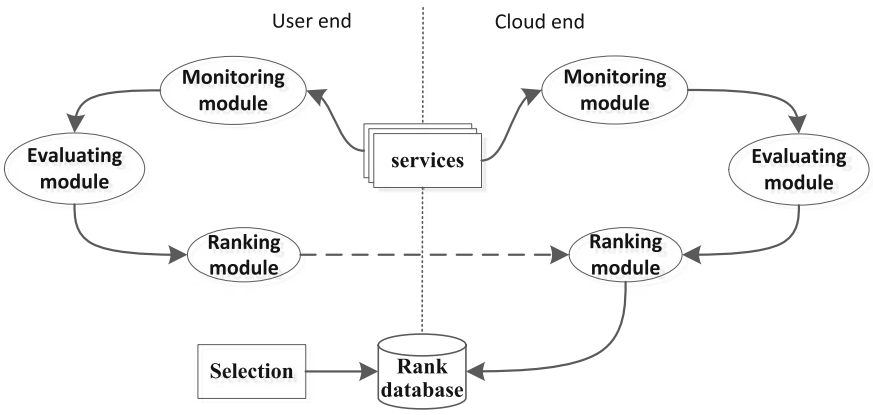


Fig. 43.1 QoE-aware cloud framework

Step 1: covert $m \times n$ matrix R_x into an $m \times n$ vector
Step 2: extract the principal component from the deduced vectors
Step 3: use clustering algorithm for clustering users
Step 4: for every t_1 interval, repeat Step 1–3.

After finishing the procedures above, we could get different user groups as $\hat{U}_1$, $\hat{U}_2$,…, $\hat{U}_m$, where $\hat{U}_n = \{U_{1n}, U_{2n}, \ldots, U_{kn}\}$ represents a set of users in the same cluster.

43.3.3 Ranking Algorithm in the Cloud End

Based on the definition in Sect. 43.2, we could easily infer that each column of S_x stands for similar services, which is called a service class. In terms of a service class, we may calculate their QoE values depending on the user clusters. Hence, the new ranking results would be various in groups. To avoid a skewed evaluation, we would not take the average of QoE values as the overall rank.

Step 1: initiate the element c_{ij} of matrix CQ to zeros.
Step 2: extract the S_x of user x in each group $\hat{U}_k$.
Step 3: for each column SL_x in S_x, calculate the new SL as follows:

$$SL = k * \frac{\prod_{x \in \hat{U}_k} SL_x}{\sum_{x \in \hat{U}_k} SL_x},$$

where k is the number of users in the $\hat{U}_k$.
Step 4: rank service of each service class in an ascend order. The better rank the service has, the higher the weight it would get.
Step 5: add the rank value to c_{ij} to update the matrix CQ for service j in cloud i.
Step 6: For every t_2 interval, repeat Step 2–5.

In general, the user preference ranking of services are first submitted to the user evaluating module. Then the user ranking module enforces clustering algorithm to group the users, while cloud service monitoring module collects logs from services in the runtime. Once a user requests a new service, cloud service selection module would invoke the cloud service ranking module to make decisions depending on its predicted group label.

43.4 Reliability and Further Discussions

The first phase of the framework has focused on the monitoring and ranking of cloud services. Next, we present how to assure the reliability of the ranking results.

In terms of the previous evaluation methods in our approach, two ideal assumptions, namely reliable and real-time message pass, had got in the way of ranking the QoE results. Since the log information from the user end would be submitted to the cloud for central processing through unpredictable networks, message loss and message delay must be taken into consideration [4]. Therefore, we intend to introduce the trust probability based on the Markov chain model to predict the objective reputation of the derived QoE values.

43.4.1 Markov Chain Model

To illustrate our design, we define the three kinds of states for services in our Markov chain model, which includes reliable state, acceptable state, and unreliable state. If none of the QoE values for a service exceed their upper bound, e.g. the threshold value negotiated in the Service Level Agreement (SLA) [7], we define it at a reliable state. If more than one of QoE values exceeds the upper bound, the service is regarded at an acceptable state. Accordingly, the unreliable state indicates that all of the QoE values exceed their upper limit. To simplify the specifications, we denote them as '*r*', '*a*' and '*u*' respectively for short.

Let p_{xy} be the transition probability from state x to state y. The transition diagram and matrix between the three states is described in Fig. 43.2.

As a result, the trust probability of each QoE value could be obtained from the transition matrix. In detail, we first calculate the p_{xy} in the diagram as follows:

$$p_{xy} = p(y|x) = \frac{p(xy)}{p(x)} = \frac{\#(\text{from state } x \text{ to state } y)}{\#(\text{state } x)},$$

$$p_{xx} = p(x|x) = \frac{p(xx)}{p(x)} = \frac{\#(\text{from state } x \text{ to state } x)}{\#(\text{state } x)}.$$

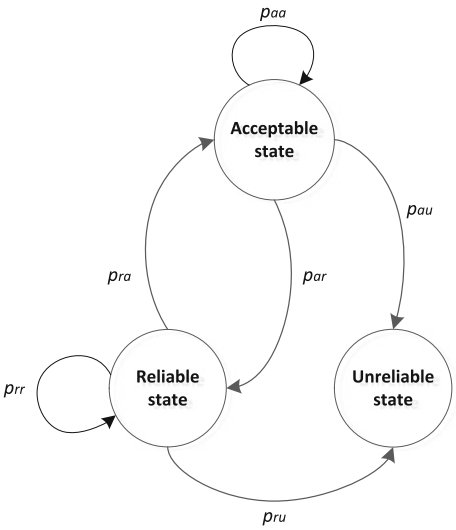


Fig. 43.2 State transition diagram

To emphasize it, we do not need to calculate all the probabilities that do not contribute to our selections, such as the p_{ur}, p_{uu} and p_u, etc. Assuming the current time to be T, we would use the transition matrix to predict the states within the time between T and 10T. After every 10T interval, we would relearn the parameters in the transition matrix. After that, we could obtain the final ranking results integrated with its predicted state for each service. On account of the successive selection on the optimal service, one of the possible solutions is that a user could choose the best ranked service belonging to its group with the maximum transition probability from a less reliable state to a reliable one.

43.5 Related Works

Recently, there have been a large number of researches on service monitoring evaluations for the optimal selection. However, these researches only focus on QoS metrics, while neglecting the user feedbacks related to QoE.

To monitor the service throughout performance, Rehman et al. [8] proposed a user feedback framework to share the QoS values from diverse users who take advantage of different benchmarks. Remano et al. [9] put forward the paradigm of "monitoring as a service" to effectively monitor the overall service performance under the uniformed standards. In addition, some business tools are also widely used in service monitoring. These tools are often used for evaluating the cloud services in an objective way, which is different from our approach.

As for the evaluations of services, Zheng et al. [10] presented a CloudRank framework to rank services, which took past component usage experience into

consideration. Compared with our approach, it neglected feedbacks from the real world, as well as the reliability during the monitoring. Ashok et al. [11] presented a trust evaluation model to improve ranking results, which is similar to the work of Jemal et al. [12] under a hybrid cloud environment. However, neither of them took the individuality of service ranking into consideration. Similarly, many researches pay attention to the ranking approaches on Web services [13, 14]. These studies also help to improve our present framework. With the increase in social networks, some researchers proposed approaches derived from recommendation system to solve service ranking problems, including the PageRank of services [15].

43.6 Conclusion

In this paper, we present a reliable QoE-aware cloud service ranking framework. Unlike the previous ranking methods, our approach focuses on the user preference related to the QoE metrics. Accordingly, we put forward a user preference ranking algorithm, which would contribute to the overall ranking of a federated cloud. Moreover, our prediction model based on the Markov chain would also drive the reliable evaluation of the cloud service.

As the immediate future work, first we plan to develop a prototype of our framework. Our future implementation would focus on minimizing the complexity of our algorithms, while optimizing the selections policies. Besides, our approach would also be evaluated based on the existing benchmarks of Web services. After that, we would also design a selection decision system to help to select the optimal services depending on our ranking results.

References

1. Armbrust M, Fox A, Griffith R et al. (2010) A view of cloud computing. Commun ACM 53(4):50–58
2. Distefano S, Puliafito A (2011) QoS management in cloud@home infrastructures. In: CyberC 2011, pp 190–197
3. Lin C, Hu J, Kong X (2012) Survey on models and evaluation of quality of experience. Chin J Comput 35(1):1–15 (in Chinese)
4. Meng S, Iyengar A, Rouvellou I, Liu L, Lee K, Palanisamy B, Tang Y (2012) Reliable state monitoring in cloud datacenters. In: IEEE CLOUD 2012, pp 951–958
5. Web of Trust (2012) Available http://www.mywot.com
6. Garg SK, Versteeg S, Buyya R (2012) A framework for ranking of cloud computing services. Future Gener Comput Syst 29(4):1012–1023
7. Patel P, Ranabahu A, Sheth A (2009) Service level agreement in cloud computing. In Cloud Workshop at OOPSLA, 2009
8. ur Rehman Z, Hussain OK, Parvin S, Hussain FK (2012) A Framework for user feedback based cloud service monitoring. In: CISIS, pp 257–262

9. Romano L, De Mari D, Jerzak Z, Fetzer C (2011) A novel approach to QoS monitoring in the cloud. In: CCP 2011, pp 45–51
10. Zheng Z, Zhang Y, Lyu MR (2010) CloudRank: a QoS-driven component ranking framework for cloud computing. In: SRDS 2010, pp 184–193
11. Chandrasekar A, Chandrasekar K, Mahadevan M, Varalakshmi P (2012) QoS monitoring and dynamic trust establishment in the cloud. In: GPC 2012, pp 289–301
12. Abawajy J (2011) Establishing trust in hybird cloud computing environments. In: Proceedings of the 2011 IEEE 10th international conference on trust, security and privacy in computing and communications. Washington, DC, USA: IEEE Computer Society. IEEE TrustCom 2011, pp 118–125
13. Meng L, Li J, Sun H (2011) WSRank: a collaborative ranking approach for web service selection. In: CIT 2011, pp 103–108
14. Wu Q, Iyengar A at el. (2009) Combining quality of service and social information for ranking services. In: ICSOC/ServiceWave 2009, pp 561–575
15. Hu R, Dou W, Liu XF, Liu J (2011) WSRank: a method for web service ranking in cloud environment. In: DASC 2011, pp 585–592

Chapter 44
Learning Distance Metrics with Feature Space Performance for Image Retrieval

Xin Luo, Guowen Wu and Kenji Kita

Abstract Learning from samples in cases where many high-dimensional vectors but only few samples are available is commonly considered a challenging problem in content-based image retrieval (CBIR). In this paper, we propose an algorithm for metric learning based on spatial distribution of image features. The optimal distance metric is then found by minimizing the divergence between the two distributions. The key idea is to construct a global metric matrix that minimizes the cluster distortions, namely, one that reduces high variances and expands low variances for the data to make a spherical form as good as possible in the high-dimensional data spaces. Experimental results show that our approach is effective in improving the performance of CBIR systems.

Keywords CBIR · Image feature space · Cluster geometry · Learning distance metric

44.1 Introduction

Content-based image retrieval (CBIR) is often done by computing the distance from a query image to images in the database, followed by the retrieval of nearest neighbors. The retrieval performance mainly depends on two related components: the image representation and the distance function used. Given a specific image

X. Luo (✉) · G. Wu
School of Computer Science and Technology, Donghua University, No. 2999 North Renmin Road, Songjiang District, Shanghai, China
e-mail: xluo@dhu.edu.cn

K. Kita
Faculty and School of Engineering, The University of Tokushima,
2-1 Minamijyousanjima-cho, Tokushima, Japan
e-mail: kita@is.tokushima-u.ac.jp

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_44,
© Springer-Verlag Berlin Heidelberg 2014

representation, the quality of the distance function used is the main key to a successful system. However, the distance function so far has been largely defined, usually by a weighting scheme and a simple cosine similarity, equivalently, a Euclidean dot product.

Distance metric learning aims to learn a distance metric from the training data that tries to maintain the cluster information of examples by their distances, i.e., examples sharing the same cluster are close to each other while examples from different clusters are separated by a large distance. During the past few years, a large number of studies are devoted to distance metric learning [1]. In many machine learning problems, the distance metric used over the input data has critical impact on the success of a learning algorithm. For instance, k-nearest neighbor classification, and cluster algorithms such as k-means rely on if an appropriate distance metric is used to faithfully model the underlying relationships between the input data points [2]. Much research effort has been spent on learning a Mahalanobis distance (MD) metric from labeled data [3, 4].

In this chapter, we propose an optimal distance function that is parameterized by a global metric matrix. This metric is optimal in the sense of global quadratic minimization, and can be learned from the given clusters in the training data. These clusters are often attributable with many forms, such as paragraphs, documents or document collections, as long as the items in the training data are not completely independent.

44.2 Related Works

Distances and metrics are now an important problem in information retrieval, machine learning, and pattern recognition. The performance of algorithms for data classification and cluster often depends heavily on the availability of a good metric. In addition, distances and metrics have found application in a number of real-world problems, including face recognition, visual object recognition, and automated speech recognition.

As an example, consider content-based retrieval in image databases, the space of features is a vector space, but it is not obvious how to introduce a norm because of the incommensurability of the components. The similarity between descriptors is usually computed with either the Euclidean distance (ED) or the MD measure.

A ED matrix is one in which the (x, y) entry specifies the squared distance between particle x and particle y [5].

$$d_E(x, y) = \sqrt{\sum_{i=1}^{p} (x_i - y_i)^2} \tag{44.1}$$

However, there is a main problem with this distance, namely, the correlation between features is ignored [6]. The major proposal is to use a distance function that allows not only for different weights of each attribute, but also for correlations.

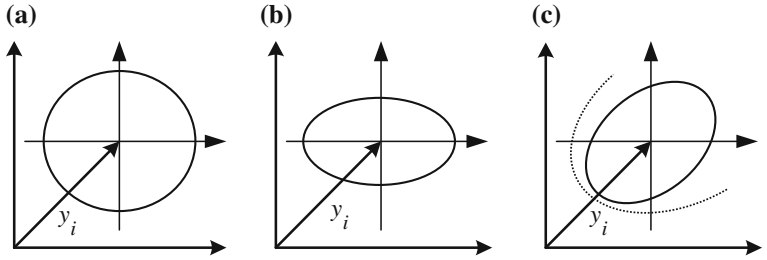


Fig. 44.1 Shape of distance functions. **a** Euclidean **b** Weighted euclidean **c** Mahalanobis

For example, see Fig. 44.1: the straight ED has circles for are surfaces (Fig. 44.1a); a weighted ED has ellipses (Fig. 44.1b), whose major axis is aligned with the coordinate axis [7]. But the weighted Euclidean is still not considered relationship may exist between features. In fact, nor by an algorithm to feature attach an appropriate weight, the weighting is inevitably arbitrary.

Usually, the distance functions result in ellipses that are not necessarily aligned with the coordinate axis [7]. In those cases, the simple ED is not an appropriate measure, while the MD will adequately account for the correlations (see Fig. 44.1c). According to definition, MD between two points $x = (x_1, \ldots, x_p)^T$ and $y = (y_1, \ldots, y_p)^T$ in the p dimensional space R is defined as

$$d_M(x, y) = \sqrt{(x - y)^T \Sigma^{-1} (x - y)} \tag{44.2}$$

In the original variable space, the MD takes into account the correlation in the data, since it is calculated using the inverse of the variance–covariance matrix of the data set of interest. However, the computation of the variance–covariance matrix can cause problems. The MD gives better matching results but has three disadvantages, viz [8]

(1) It requires a covariance matrix estimated from training data;
(2) Performance results will depend on the training set used;
(3) It is a global measure, not optimally adapted to the local structure at any feature point of interest.

Based on these problems, we propose a more generic measure similar to the metric matrix and take into account the local structure at the feature points of interest. In this case, the covariance matrix is obtained directly from the differential structure at each interest point. The matrix can be obtained in analytical form and reflects the actual behavior of the descriptor due to small perturbations. In the next section, we present the details of this approach.

44.3 The Proposed Algorithm

For the problems above-mentioned, we are on the basis of the related features of every component and pose a classification structure based on the training data. The method is by deducing the optimal distance matrix for learning and makes new classification be found by learning these matrixes.

44.3.1 The Ellipsoid Distance Function

As stated above, vectors in the same cluster must have a small distance between each other in the ideal geometry. When we measure an L2-distance between x and y by a MD parameterized by M:

$$D^2(x,y) = (x-y)^T M(x-y) \tag{44.3}$$

where symmetric metric matrix M gives both corresponding feature weights and feature correlations.

Generally, when setting M to be an identity matrix, the distance in Eq. (44.3) becomes the common ED.

For any symmetric matrix M the following equation holds:

$a_{ij} = a_{ji}(i,j = 1,2,\ldots,p)$ and $M^T = M$. M is also a positive definite matrix, then it is easy to show that $x \neq 0$, $x^T Mx > 0$.

By the learning matrix $M = [m_{ij}]$, Eq. (44.2) can be rewritten as follows:

$$D^2(x,y) = \sum_{i}^{p}\sum_{j}^{p} m_{ij}(x_i - y_i)^T(x_j - y_j) \tag{44.4}$$

where learning matrix M denotes the weight of any component and the relationship among components.

Because M is a symmetric matrix, then:

$$D^2(x,y) = (M^{1/2}(x-y)^T)(M^{1/2}(x-y)) \tag{44.5}$$

Note that this distance is global, and different from the ordinary MD in pattern recognition that is defined for each cluster one by one, using a cluster specific covariance matrix.

Therefore, we require an optimization over all the clusters in the training data. Generally, data in the clusters are distributed as in Fig. 44.2a, comprising ellipsoidal forms that have high covariances for some dimensions and low covariances for other dimensions.

Further, the cluster is not usually aligned to the axes of coordinates. When we find a global metric matrix M that minimizes the cluster distortions, namely, one that reduces high covariances and expands low covariances for the data to make a

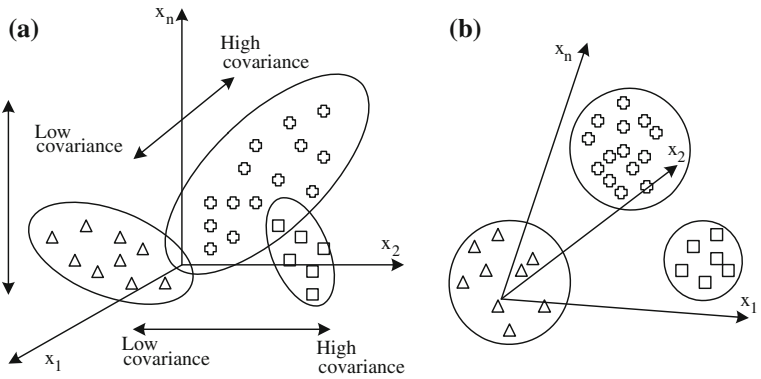


Fig. 44.2 Cluster geometry of feature space. **a** Original space **b** Mapped space

spherical form as good as possible in the $M^{1/2}$ mapped space (Fig. 44.2b), we can expect it to capture necessary and unnecessary covariations and correlations on the features, combining information from many clusters to produce a more reliable metric that is not locally optimal. We will find this optimal M below.

44.3.2 Learning the Optimal Distance Metric

As we define the distance among classes, gravity method was chosen to perform the sum of simples in each cluster.

Suppose that R^n is divided into K clusters $(X_1, \ldots, X_K)$ for each cluster X_i, cluster centric c_i is calculated as

$$c_i = \frac{1}{X_i} \sum_{x \in X_i} x \qquad (x = 1, 2, 3, \ldots, K) \tag{44.6}$$

where X_i denotes the number of data in X.

We seek the metric matrix M that minimizes the distance between each data x and the cluster centroid c_i, $d_M(x, c_i)$ for all clusters $X(x \in X_i)$. Mathematically, this is formulated as a quadratic minimization problem:

$$M = \arg\min_M \sum_{i=1}^{K} \sum_{x_j \in X_i} d_M(x_j, c_i)^2 \quad v \tag{44.7}$$

Subject to a constraint:

$$\det(M) = 1 \tag{44.8}$$

The constraints are shown in [9]. Without any constraints, the zero matrixes would give the minimum.

We expand Eq. (44.7). Then, we have

$$\begin{aligned}&\sum_{i=1}^{K}\sum_{x_j\in X_i}(x_j-c_i)^T M(x_j-c_i)\\&\quad=\sum_{i=1}^{K}\sum_{x_j\in X_i}\left[\sum_{k=1}^{n}\sum_{l=1}^{n}(x_{jk}-c_{ik})m_{kl}(x_{jl}-c_{il})\right]\end{aligned}\tag{44.9}$$

where $x_j=\left[x_{jk}\right]$,$c_i=[c_{ik}]$ and n is the dimension.

And from the constraint (44.8), for all k

$$\sum_{l=1}^{n}(-1)^{k+l}m_{kl}\det(M_{kl})=1\tag{44.10}$$

Therefore,

$$\sum_{k=1}^{n}\sum_{l=1}^{n}(-1)^{k+l}m_{kl}\det(M_{kl})=n\tag{44.11}$$

where M_{kl} denotes an adjuvant matrix of m_{kl}.

By introducing the Lagrange multiplier λ, we have

$$\begin{aligned}L=&\sum_{i=1}^{K}\sum_{x_j\in X_i}\left[\sum_{k=1}^{n}\sum_{l=1}^{n}(x_{jk}-c_{ik})m_{kl}(x_{jl}-c_{il})\right]\\&-\lambda\left[\sum_{k=1}^{n}\sum_{l=1}^{n}(-1)^{k+l}m_{kl}\det(M_{kl})-n\right]\end{aligned}$$

Differentiating by m_{kl} and setting to zero, we obtain

$$\begin{aligned}\frac{\partial L}{\partial m_{kl}}=&\sum_{i=1}^{K}\sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})\\&-\lambda(-1)^{k+l}\det(M_{kl})\end{aligned}$$

Let us define $\frac{\partial L}{\partial m_{kl}}=0$. Then,

$$\sum_{i=1}^{K}\sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})=\lambda(-1)^{k+l}\det(M_{kl})$$

Therefore,

$$\det(M_{kl})=\frac{\sum_{i=1}^{K}\sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})}{\lambda(-1)^{k+l}}$$

The inverse matrix $M^{-1} = \left[m_{kl}^{-1}\right]$ can be represented as:

$$\begin{aligned} m_{kl}^{-1} &= \frac{(-1)^{k+l}\det(M_{kl})}{\det(M)} = (-1)^{k+l}\det(M_{kl}) \\ &= (-1)^{k+l}\frac{\sum_{i=1}^{N}\sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})}{\lambda(-1)^{k+l}} \\ &= \frac{\sum_{i=1}^{N}\sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})}{\lambda} \end{aligned} \tag{44.12}$$

Let $A = [a_{kl}]$ be the matrix $a_{kl} = \sum_{x_j\in X_i}(x_{jk}-c_{ik})(x_{jl}-c_{il})$.

Form Eq. (44.12), we have $A = \lambda M^{-1}$, then, $\det(A) = \lambda^n \det(M^{-1}) = \lambda^n\lambda = [\det(A)]^{\frac{1}{n}}$.

Therefore

$$M = \lambda A^{-1} = [\det(A)]^{\frac{1}{n}}A^{-1} \tag{44.13}$$

44.4 Experiments and Evaluation

44.4.1 Experimental Data

Based on our proposal, a working system for CBIR has been established. To date, we have tested our CBIR system on a general purpose image database with 51,138 images from COREL CDs. These images have 60 categories with 100 images in each category. Every category represents a different semantic topic, such as building, mountain, beach, dog, horse, etc. Two image features, color and texture, are used for image retrieval. The color feature is a 48-dimensional vector generated by QBIC [10]. The texture feature is a 48-dimensional vector generated using the algorithms proposed in [11].With these two types of features; 96-dimensional feature database is used for retrieval in our experiment.

After then, we chose 369 images of the most representative from the CDs, and these images with complex background and composition. For example, image of snow mountain and a lake, image of grasslands, bronze statue and trees, etc. In general, training cluster doesn't take into account the cluster structure in the features in the extraction of images, where neither extra weighted, nor does it consider the diagonal relationship of features, the positive and negative weighted are usually even distributions. Actually, diagonal element is able to reflect the main body from features. The learning clusters of "Fireworks" shown in Fig. 44.3, which indicates the distribution of the four image features. In addition, from the distribution diagram, concentration of feature can faithfully reflect the description and subject of image of the cluster. Figure 44.4 shows that one of the learning

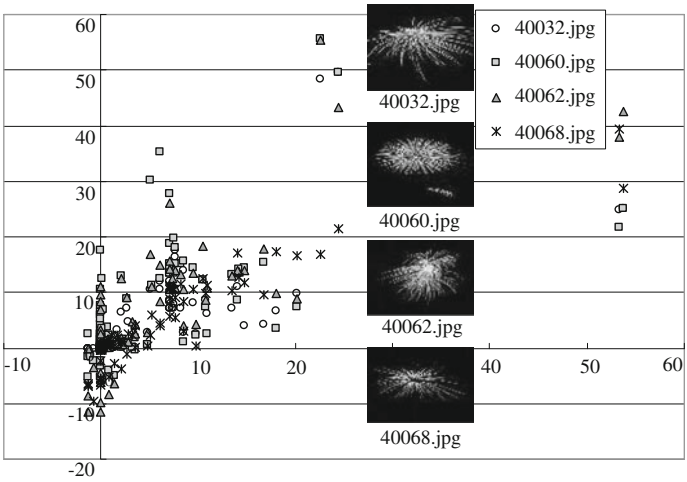


Fig. 44.3 Data distribution of "fireworks" cluster

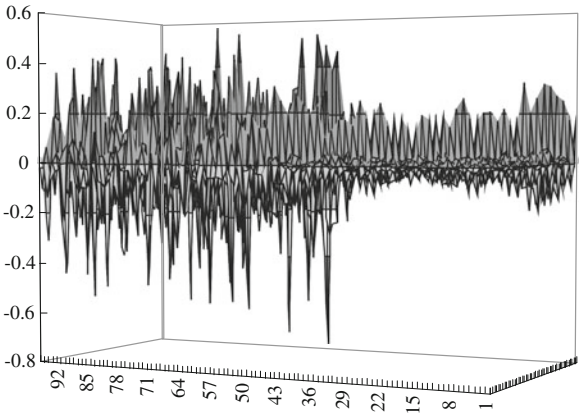


Fig. 44.4 The diagonal distribution of feature of one cluster

clusters of metric matrix with the algorithm, and by diagonal element, the main body of feature and the relationship of cluster can be enhanced.

Our retrieval system runs according to the following processes:

(1) To manually select the images for learning data;
(2) Calculate the distance matrix M of the learning data;
(3) For a given query, according to a similarity metric retrieves in the system;
(4) Output the top 100 images of ranked as a retrieval result.

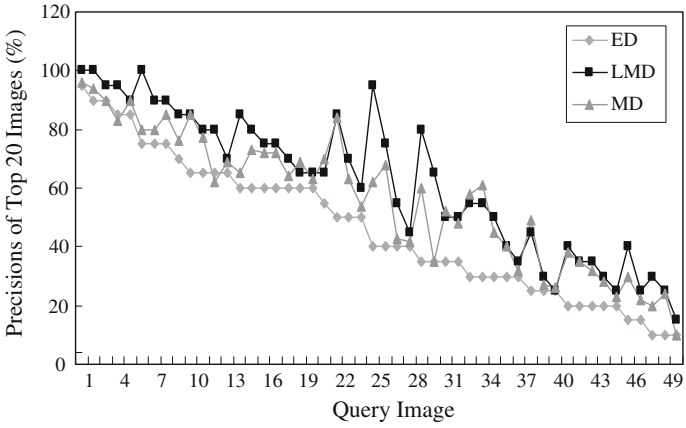


Fig. 44.5 Precisions of in-cluster at various distances metric

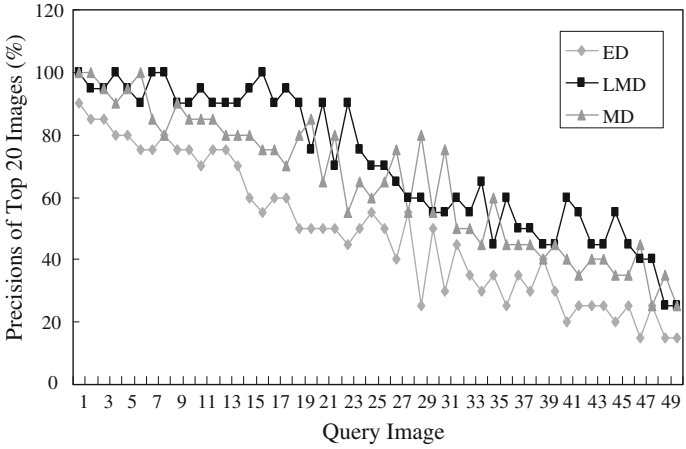


Fig. 44.6 Precisions of out-of-cluster at various distances metric

44.4.2 Experimental Results

For evaluating the performance of the algorithm, 100 images of randomly selected are chosen from the database as the test query images due to their clear semantic meaning, including 50 images in learning cluster and 50 images in out-of-cluster. The query image was selected from the database or the test query images and it would be the first image in the result list. Other images in the result list were retrieved and ranked based on the similarity to the query image. Based on precision and recall, we adopt ED, MD and presented algorithm (learning metric distance, LMD) to a query, respectively. The performance of combining methods

are similar, though "Rank Euclidean" is a little behind. Figure 44.5 shows that comparative precisions in aspect of learning cluster. Figure 44.6 shows that the comparison of precisions to query of out-of-cluster. From comparison of answer set, it is indicated that retrieval accuracy of the classified image is higher than out-of-cluster. More important, the retrieval results using learning metric matrix are better than unused.

Thus, we find in our learning set those images that are similar to the focal image according to at least one elementary distance measure. For each of the M metric matrix measures, we find the top K closest images. For example, "apple," "red rose," "dinosaur," "playing cards," "surfing," etc., in the learning cluster images and "red flower," "sun," "man," "aircraft," "bud," etc., in the out-of-cluster images. If all K images are in-cluster, then we find the closest out-of-cluster image according to that distance measure and make K triplets with one out-of-cluster image and the K similar images. Especially these query images with complicated background, general methods tend not to be adequate in retrieval, but our algorithm's performance in this test to emphasize more.

44.5 Conclusion

In this chapter, we proposed a global metric distance that is useful for image retrieval where ED and MD have been used. The key idea is to construct the kernel matrices for given training data based on the idea of minimum cluster distortion in training data. This distance is optimal in the sense of quadratic minimization over all the clusters in the training data. Through our theoretical analysis and experimental results, we conclude that improvements over ED and Mahananobis distance, with a significant refinement with tight training clusters in image retrieval.

Acknowledgment This research was partially supported by "the Fundamental Research Funds for the Central Universities (No. 13D11205)."

References

1. Hoi SCH, Liu W, Lyu MR, Ma W-Y (2006) Learning distance metrics with contextual constraints for image retrieval. In: Proceedings of the computer vision and pattern recognition
2. Shental N, Hertz T, Weinshall D, Pavel M (2002) Adjustment learning and relevant component analysis. In: ECCV 2002, pp 776–792
3. Davis JV, Kulis B, Jain P, Sra S, Dhillon IS (2007) Information theoretic metric learning. In: Proceedings of the international conference on machine learning, Corvalis, Oregon, pp 209–216
4. Yang L, Sukthankar R, Hoi SCH (2010) A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. IEEE Trans Pattern Anal Mach Intell 32(1):30–44

5. Chen J, Wang R, Shan S, Chen X, Gao X (2006) Isomap based on the image euclidean distance. In: The IEEE 7th international conference on pattern recognition (ICPR2006), pp 1110–1113
6. Wang Liwei, Zhang Yan, Feng Jufu (2005) On the Euclidean distance of images. IEEE Trans Pattern Anal Mach Intell 27(8):1334–1339
7. Luo X, Shishibori M, Ren F, Kita K (2007) Incorporate feature space transformation to content-based image retrieval with relevance feedback. Int J Innovative Comput Inf Control (IJICIC) 3(5):1237–1250
8. Balmachnova E, Florack L, ter Haar Romeny B (2007) Feature vector similarity based on local structure. In: SSVM 2007, pp 386–393
9. Ishikawa Y, Subramanya R, Faloutsos C (1998) MindReader: querying database through multiple examples. In: Proceedings of the 24th international conference on very large database, pp 218–227
10. Niblack W, Barber R, Equitz W, Flickner M, Glasman E, Pektovic D, Yanker P, Faloutsos C, Taubin G (1993) The QBIC project: querying images by content using color, texture, and shape. In: Proceedings of SPIE storage and retrieval for image and video databases, pp 173–181
11. Manjunath BS, Ma WY (1996) Texture features for browsing and retrieval of large image data. IEEE Trans PAMI 18(8):837–842

Chapter 45
A Fine-Grained Authentication Model Based on Perceptual Hashing and Grid Descriptor for Remote Sensing Image

Kaimeng Ding and Yuhai Wang

Abstract In this paper, a fine-grained authentication model for remote sensing image based on perceptual hashing and grid descriptor is proposed. Most perceptual hashing algorithms generate the hash value from an image's global features, while remote sensing images are generally of huge amount and large size, so they are not suitable for remote sensing images authentication applications with high security demand. In this work, we apply grid descriptor to divide a remote sensing image, then generate the perceptual hash value of each region, and organize these hash values by embedding them into the corresponding region with watermarking technique. The grid descriptor is applied to detect and represent the tamper of the image. Compared with other authentication algorithms, the model can authentically remote sense image with different granularity.

Keywords Authentication · Remote sensing image · Perceptual hashing · Grid descriptor

45.1 Introduction

Remote sensing image plays an important role in a variety of fields of national living because of its rich information magnitude and gradually improved resolution. For the leak in the security of storing and network transmitting, the protection of the content of remote sensing image has become more and more important. A number of approaches have already been proposed to protect remote sensing images, including cryptography techniques, watermarking techniques, digital signatures, and perceptual hashing (PH).

K. Ding (✉) · Y. Wang
Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing, People's Republic of China
e-mail: dingkaimeng@foxmail.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_45,
© Springer-Verlag Berlin Heidelberg 2014

Cryptography techniques are not suitable for image authentication, as they are so sensitive that even one bit change of the input data will lead to a significant change in the output hash. In order to make up for the disadvantage of the traditional cryptographic hash functions in image authentication, robust image hashing was first introduced by Schneider et al. [1]. Perceptual hashing (perception hash) technology, which converts image data into a string of binary sequences, has emerged as an attractive way to verify the authenticity of digital images [2]. Perceptual hashing has the special property that the hash value is dependent on the content, so that the same content always maps to the same hash value and it remains approximately the same if the content is not significantly modified.

However, remote sensing images are generally of huge amount and large size (some may be several GB), while most perceptual hash algorithms compute the hash value from an image's global features that are not sensitive to local modification. Thus those algorithms are generally vulnerable to minor modifications in the remote sensing images. This means that the existing perceptual hash algorithms are not suitable for remote sensing images authentication with high security demand.

We propose a fine-grained remote sensing image content authentication model based on PH and grid descriptor (GD). In this work, we apply GD to segment the image and represent the tamper. The resolution of the grid is not fixed and is set according to the actual demand.

45.2 Related Work

45.2.1 Perceptual Hash Algorithm

A perceptual hash algorithm should possess the properties of one-way and collision-resistant. What is more, the perceptual hash value should be insensitive to legitimate content distortion, but sensitive to malicious modification and incidental distortion. As shown in Fig. 45.1, a perceptual hash algorithm generally consists of four pipeline stages: the transformation stage, the feature extraction stage, the quantization stage, and the compression stage.

The feature extraction is the key process of hash sequence generation. The main existing feature extraction methods include: the statistical method based on image, image representation, extraction method based on the relation, and roughly said image feature point. Above all the feature extraction methods, in order to better satisfy the perception of image hash robustness, safety and uniqueness, using DCT transformation to extract feature is considered as the ideal method [3].

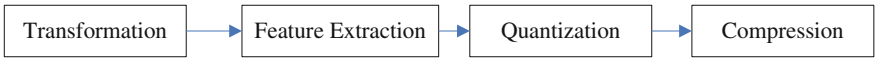


Fig. 45.1 Perceptual hash generation

45.2.2 Grid Descriptor

The GD was first proposed by Lu and Sajjanhar for shape description and similar shape retrieval [4]. In this paper, we adopt GD to segment remote sensing image, to compute the perceptual hash value of each block, and to represent the temper description of the authentication. Let PH_{ij} denote the perceptual hash value (PH) of the grid cell of coordinates (i, j), it denotes perceptual hash value of region of remote sensing image covered by the grid cell of coordinates (i, j).

In the approach proposed by Lu and Sajjanhar [5], a grid of cells is overlaid on a shape and each cell has the same size. The cells covered by the shape are assigned 1 and those not covered by the shape are assigned 0.

In order to suit the remote sensing image authentication, we propose to design an authentication model by a GD-based perceptual hash algorithm. This means that the remote sensing image is first shadowed on a grid with fixed resolution (the resolution is set according to actual authentication demand), then compute the perceptual hash value of each region covered by the grid cell, and denote PH_{ij}.

In addition to the approach proposed in [5], the remote sensing image is shadowed on a grid with fixed resolution. What is different is that the cells are assigned PH_{ij}, the perceptual hash value of region, rather than simply assigned 1 or 0.

For example, as shown in Fig. 45.2, we assume that a remote sensing image is overlaid with 10×10 grid, and the grid is 3000×3000 pixels, each cell being 300×300 pixels in size. Then we compute each perceptual hash (PH_{ij}) value of region covered by the grid cell. The GD denotes the perceptual hash values matrix.

45.2.3 Watermarking Technique

Watermarking technique, which has always been used to hide information, is the art of protecting the multimedia data by hiding a special mark that can be detected again [6]. It should meet two requirements. First, a watermarked image should be perceptually equal to the original one. Secondly, the watermark can be successfully detected, even if some attacks had been applied on the watermarked image.

Remote sensing image has more urgent demand for watermarking techniques and quite a number of scholars have devoted to study of the watermarking technology for digital remote sensing images which can be classified into spatial domain algorithms and frequency domain algorithms [7].

Least significant bit (LSB) algorithm, which is the simplest spatial domain algorithm of watermarking algorithm, embeds watermark by adjusting the most

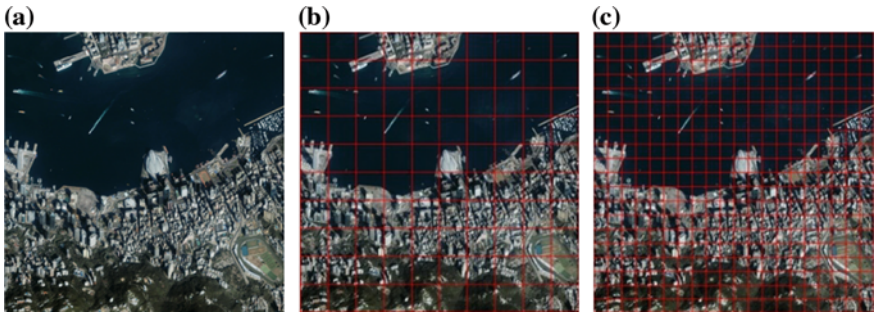


Fig. 45.2 Different resolution grid. **a** Original image, **b** 10 × 10 grid, **c** 20 × 20 Grid

unperceivable pixels and modifying the bit plane of image color [3]. Frequency domain algorithms are better than spatial domain algorithms in terms of robustness, which embed watermark by modifying the coefficients in frequency domains after transforming image via discrete cosine transform (DCT), discrete wavelet transform (DWT), etc. Such a domain features the concentrative energy distribution and the relevant pixel points, making it compatible with the human visual characteristics and contributing to the invisibility and robustness.

45.3 The Proposed Model

In this section, the fine-grained authentication model based on PH and GD for remote sensing image is proposed, which organizes the perceptual hash value by embedding them into the corresponding region.

The single perceptual hash value of one whole remote sensing image cannot meet the security requirement for remote sensing image, because of the huge amount of data and the large size of the remote sensing image. As a response to this, the remote sensing image is first shadowed on a grid with fixed resolution (the resolution is set according to actual authentication demand), then the perceptual hash value of each region is computed covered by the grid cell denoting PH_{ij}.

The GD is used to detect and represent tamper localization. Therefore, the realization of tamper location ability is built on the resolution of GD. The quality of the authentication improves as we go to higher resolutions. The higher the resolution of the GD, the more fine-grained the authentication granularity can be. But high GD resolution also raises the computational cost: at higher resolutions, we need to segment the remote sensing image into more blocks, which increases both the time for computing perceptual hash values as well as comparing the values. The choice of the GD resolution thus presents a tradeoff between the cost and tamper location ability. Our work is aimed at designing such an authentication model with good balance between cost and performance.

45.3.1 Hash Generation and Embedding Module

The flowchart of the proposed algorithm of hash generating and embedding for block is shown in Fig. 45.3, and the details of scheme is as follows:

Step 1 Overlay a GD with fixed resolution, say 10×10 or 20×20, on a remote sensing image. The selection of the resolution is according to the actual authentication demand.

Step 2 Dividing the image into blocks which is region covered by the cell of the grid. The size of the blocks (cell) is corresponding to the resolution of the GD.

Step 3 For each block, compute its perceptual hash value denoted as PH_{ij}. First change the block into 255 gray image and distinguish ability into 32×32.

Step 4 Change the 32×32 block into small pieces of 8×8, then transform each piece with DCT, and keep one DC coefficient and three AC coefficients for each piece.

Step 5 Combine all of the coefficients as a vector, and quantize the vector with lengths of 64 bits.

Step 6 Encrypt the vector by RC4 algorithm, and then get the final perceptual hash. RC4 algorithm has been one of the most widely used software stream ciphers.

Step 7 After getting the perceptual hash value of the block, embed it into the corresponding region of the remote sensing image using watermarking techniques, and then we get the final remote sensing image. Here, we employ LSB algorithm as an example to embed the perceptual hash value.

45.3.2 Authentication Module

During authentication process, the embedded image has been taken as the input data and the embedded perceptual hash value is extracted from the corresponding region. The process of regenerating the perceptual hash is the same as before. The authentication done by the following procedures are shown in Fig. 45.4.

Step 1 Overlay the same descriptor grid (GD) on the received remote sensing image to be authenticated, and divide it into blocks according to the same GD.

Step 2 For each block, regenerate the perceptual hash, denoted PH_{ij}, and extract the embedded unit perceptual hash, denoted PH'_{ij}.

Step 3 If PH_{ij} is equal to PH'_{ij}, it means the image block is not significantly modified, and it remains approximately the same.

Step 4 If not, it means the image block has been tramped and the content has been modified perceptually.

Step 5 Detect all of the tampered regions and represent the temper description of authentication.

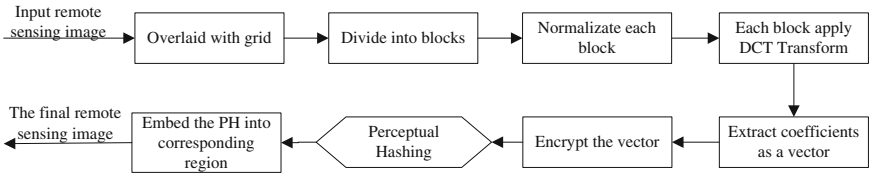


Fig. 45.3 Hash generation and embedding

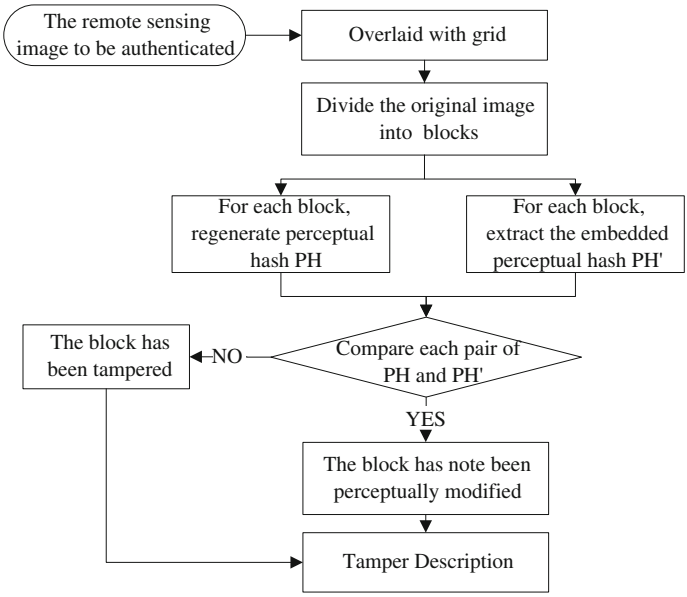


Fig. 45.4 Authentication process

45.4 Experimental Result

In this section we present some experiments to demonstrate the performance of our proposed approach.

Our experiments were performed using a remote sensing image (3000×3000), as shown in Fig. 45.5, the remote sensing image is first shadowed on three different grids with resolution 5×5, 10×10, 20×20 respectively. Then divide the image into blocks according to the grid. For each block, generate the perceptual hash value and embed it into the corresponding region of the remote sensing image.

As shown in Figs. 45.6, 45.7, our proposed model can detect tempering of the remote sensing image, and can represent the tempering. The higher the resolution

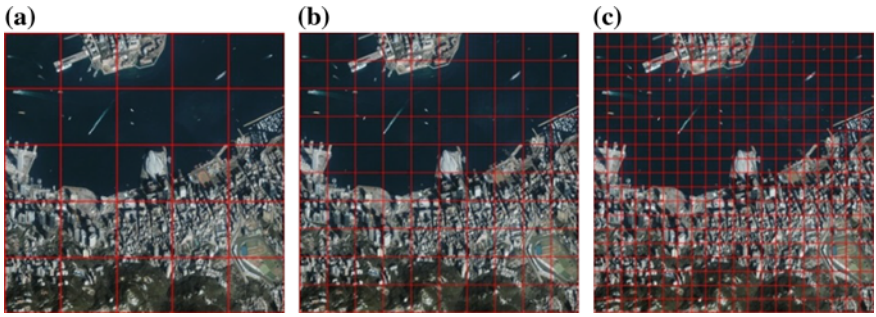


Fig. 45.5 Different resolution grid. **a** 5 × 5 grid, **b** 10 × 10 grid, **c** 20 × 20 grid

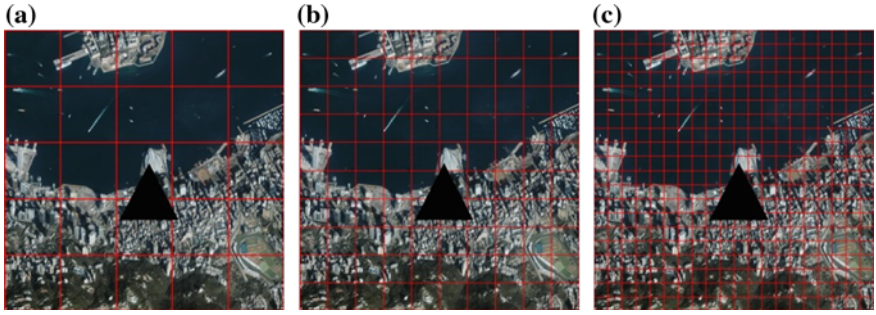


Fig. 45.6 Remote sensing image tampering. **a** 5 × 5 grid, **b** 10 × 10 grid, **c** 20 × 20 grid

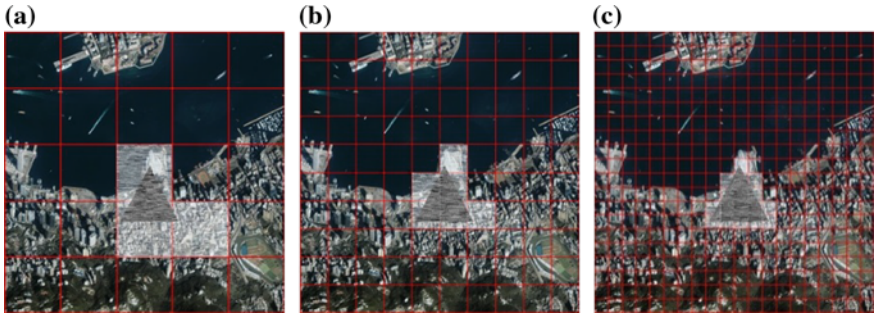


Fig. 45.7 Tampering detection and representation. **a** 5 × 5 grid, **b** 10 × 10 grid, **c** 20 × 20 grid

of the GD, the authentication granularity can be more fine-grained. While at higher resolutions, we need to segment the remote sensing image into more blocks, compute more unit perceptual hash values, and compare the values.

45.5 Conclusions

In this paper, we have described a remote sensing image authentication model based on PH and GD. Compared with other authentication algorithms, the GD-based perceptual hash algorithm can authentically remote sense images with fine granularity. The GD is applied to segment the remote sensing image and represent the tamper of the image. Our proposed authentication model is suitable for remote sensing image with large size and high security demand and can authenticate it with different granularity.

Acknowledgments This research is supported by the National Natural Science Foundation of China, No. 41071245; Research and Innovation Project for Graduates of Jiangsu Higher Education Institutions (CXLX13_378).

References

1. Schneider M, Chang SF (1996) A robust content based digital signature for image authentication. Proc IEEE Int Conf Image Process 3:227–230
2. Chang L, Yan WG, Wang WD (2012) Research on robust image perceptual hashing technology based on discrete cosine transform. In: Business, Economics, Financial Sciences, and Management. Springer, Berlin, Heidelberg, pp 799–809
3. Christophe D, Delaigle J, Macq B (2002) Invisibility and application functionalities in perceptual watermarking: an overview. Proc IEEE 90(1):64–77
4. Zhang D, Lu G (2004) Review of shape representation and description techniques. Pattern Recogn 37(1):1–19
5. Lu G, Sajjanhar A (1999) Region-based shape representation and similarity measure suitable for content-based image retrieval. Multimedia Syst 7(2):165–174
6. Menezes A, van Oorschot P, Vanstone S (1996) Handbook of applied cryptography. CRC Press, Boca Raton
7. Cho SY, Han SY (2003) Wavelet packet based digital watermarking for remote sensing image compression. In: Computer and information sciences (ISCIS 2003). Springer, Berlin, Heidelberg, pp 482–489

Chapter 46
Adaptive Tuning Algorithm Used in Multi-Join Query Optimization

Zhou Jiang, Lianzhong Liu and Zheren Li

Abstract The multi-join query optimization problem is hot and difficult in the data query optimization research field. Based on the study at cost estimation methods and the theory of multi-join queries, this paper gives an improved cost estimation model and a new search algorithm of query execution strategy space. The proposed optimization method uses adaptive genetic algorithm based on cloud theory in searching query strategy space. Simulation results demonstrate the effectiveness of the algorithm.

Keywords Multi-join · Query optimization · Cost model · Adaptive genetic algorithm · Cloud theory

46.1 Introduction

For relational database Multi-join Query Optimization (MJQO), many scholars have conducted a lot of research. Rafi Ahmed studies cost-based query transformation in Oracle [1]. Stanford Arun Swami studies the local random search algorithms (such as iterative improvement II and simulated annealing SA, etc.) in MJQO [2]. Michael Stenbrunn and others have performed theoretical analysis-related topics of join order optimization [3]. Kristin Bennett and others have conducted groundbreaking research on evolutionary algorithm family in the application of query optimization [4]. Guo Liying studied MJQO based on genetic algorithm (GA) [5]. Guo Congli applied ant colony algorithm to MJQO [6]. Lina Song applied GA and simulated annealing algorithm to MJQO and achieved good

Z. Jiang (✉) · L. Liu · Z. Li
Beijing Key Laboratory of Network Technology, School of Computer Science and Engineering, Beihang University (BUAA), Beijing 100191, China
e-mail: jz_19880815@163.Com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_46,
© Springer-Verlag Berlin Heidelberg 2014

results [7]. Xiaofan Zhao combined GA and tabu search algorithm and achieved ideal results [8].

For the traditional GA, the selection of genetic parameters plays a key role. The standard genetic algorithm (SGA) uses fixed crossover probability and mutation probability. Adaptive genetic algorithm (AGA), proposed by Srinivas, significantly improves the convergence than SGA [9]. However, most of these methods only consider the trend of the evolution process, but ignore the randomness of natural evolution.

The join cost estimation model (JCEM) using in fitness function of GA [10] cannot effectively evaluate the join cost of a table with null values and without considering each table cardinality filter. These affect the accuracy of the fitness function.

This paper is mainly divided into the following sections: we propose our JCEM in the second section. We propose AGA based on cloud theory used in multi-join query in the third section. We perform simulation experiment and analyze the results according to the proposed algorithm in the next section. The last section is the reference section.

46.2 Join Cost Estimation Model

In this paper we show how to establish JCEM. The model is based on two assumptions, the attribute value is evenly distributed and the number of tuples in intermediate operation results determines the I/O cost of the query. So the cost of query execution plan is determined by the sum of the number of tuples in intermediate operation results.

46.2.1 The Establishment of a Mathematical Model for Cost Estimation

Suppose there are n relations to participate in the join operation, and the internal nodes of join tree t are t_i, then its cost of execution plan (expressed in cost(t)) is the sum of the number of tuples in intermediate operation results. The cost estimation formula is

$$\cos t(t) = \sum_{i=1}^{n-1} \cos t(t_i),\ i = 1, 2, \ldots, n-1 \tag{46.1}$$

For multi-join query Q, set T is a collection of all possible query plan, in which each member t has relevant cost (expressed in cost(t)). Our purpose is to find t_0 from set T to make it meet the following formula:

$$\cos t(t_0) \approx \min \cos t(t), t \in T \tag{46.2}$$

Cardinality is defined as how many rows are fetched from table. $V(A, r)$ represents the number of unique values which occur in attribute set A in relation r, and $A_j \in A$.

$$V(A, r) = \prod_{A_j \in A} V(A_j, r) \tag{46.3}$$

For internal node t_i of join tree t, l represents the left child node of t_i and r represents the right child node of t_i. $t_i = l$ join r. A is a set of public properties between relation l and relation r, and A_j is one of the public properties. $n(l)$ and $n(r)$ represent the cardinality of relation l and r. null(A_j, r) represents the number of null values which occur in attribute A_j in relation r.

Usually, SQL statement may contain some filtering predicates. These predicates should be applied to the relation first to get filter cardinality. For example, query criteria l. Filter = 2 will reduce cardinality of relation l before join. Suppose the collection of filtering predicates of relation r is R, and $R_k \in R$. $F(R, r)$ represents the filtering probability of attribute sets R in relation r.

$$F(R, r) = \frac{1}{\prod_{R_k \in R} V(R_k, r)} \tag{46.4}$$

Suppose the collection of filtering predicates of relation l is L, The same can be $F(L, l)$. The cost calculation formula is as follows:

$$\begin{aligned} \cos t(t_i) = {} & n(l) \times n(r) \times F(L, l) \times F(R, r) \\ & \times \prod_{A_j \in A} \frac{(n(l) - \text{null}(A_j, l)) \times (n(r) - \text{null}(A_j, r))}{n(l) \times n(r) \times \max(V(A_j, l), V(A_j, r))} \end{aligned} \tag{46.5}$$

Exceptionally, when join operator is Cartesian product.

$$\cos t(t_i) = n(l) \times n(r) \times F(L, l) \times F(R, r) \tag{46.6}$$

46.2.2 Estimate the Number of Different Values in Join Computing Results

We need to estimate the number of different values in properties a from join computing results as $V(a, l \text{ join } r)$, so we can estimate the cost of intermediate join operation. Statistical analysis is as follows:

Suppose property a is from relation l, and not from relation r. If $a \notin L$ then $V(a, l \text{ join } r) = F(L, l) \times V(a, l)$, and if $a \in L$ then $V(a, l \text{ join } r) = 1$. In the same

way, suppose property a is from relation r, and not from relation l. If $a \notin R$ then $V(a, l \text{ join } r) = F(R, r) \times V(a, r)$, moreover, if $a \in R$ then $V(a, l \text{ join } r) = 1$.

If property a is a public property between relation l and relation r. If $a \notin L \cap a \notin R$ then $V(a, l \text{ join } r) = \min(F(L, l) \times V(a, l), F(R, r) \times V(a, r))$, moreover, if $a \in L \cup a \in R$ then $V(a, l \text{ join } r) = 1$.

46.3 Adaptive Genetic Algorithm Based on Cloud Theory Used in MJQO

Dr. Dai Chaohua proposed a new adaptive genetic algorithm called adaptive genetic algorithm based on cloud theory (CAGA) [11]. CAGA uses positive cloud generator in order to deduce adaptive crossover probability and adaptive mutation probability, which are used in AGA. Multi-join queries may be represented as hundreds of equivalent join trees inside the database. The number of equivalent join trees explosively grows along with the growth of the number of joins. We need to use JCEM to compare cost among equivalent join trees and use an effective algorithm to find a join tree which has minimum cost. In our paper, we introduce how to use AGA based on cloud theory in MJQO.

46.3.1 Chromosome Encoding Scheme

The algorithm uses path sequence encoding. First, it forms a one-to-one mapping between each table and a unique integer. Second, it encodes the possible join path as an integer string.

46.3.2 Fitness Function

Individual fitness value in the algorithm is equal to N-table join cost of this individual. It is assumed that there are n relations to participate in the join operation and an internal node of join tree t is t_i. The cost of execution plan (expressed in cost(t)) is the sum of the number of tuples in intermediate operation results, which are expressed in internal nodes t_i. It calculates the individual fitness value according to the cost estimate formula given in JCEM.

$$f_i(t) = \cos t(t) = \sum_{i=1}^{n-1} \cos t(t_i), i = 1, 2, \ldots, n-1 \tag{46.7}$$

46.3.3 Initialization of the Population

Initial population is selected using the heuristic information. This algorithm is based on a method, which combines minimum cardinality criterion heuristics and random algorithm, and avoids Cartesian product. First, we choose each relation in relation set as the first place of the code. Second, we heuristically select a join relation with minimum cardinality from the rest of relations as the second place of the code, and so on. The remaining individuals of Initial population are randomly initialized.

46.3.4 Selection Strategy

We combine the optimal preservation strategy and roulette selection strategy together. The optimal preservation strategy does not use crossover and mutation operations, and copies itself to the next generation directly for the individual with minimum fitness in the population. We combine it with roulette selection strategy and keep the diversity of the population much better. Suppose the size of the population is M and the fitness of individual i is f_i, then the probability of individual i is selected is as follows:

$$P(i) = f_i / \sum_{i=1}^{M} f_i, \quad i = 1, 2, \ldots, M. \tag{46.8}$$

46.3.5 Crossover Operator

We use order crossover method in the algorithm. It generates new individuals from the parent population according to the adaptive crossover probability pc.

1. Randomly select a substring from the first parent.
2. Copy the substring to the corresponding position of an empty string in order to produce an original offspring.
3. Delete existing genes of the substring from the second parent and get order of other genes which are needed by the original generation.
4. Copy these genes to the vacant position of offspring from left to right according to this sequence.

Table 46.1 Hardware and software environments

Hardware and software	Parameter
Compiler	Myeclipse 7.5
Operating system	Windows 7 (32bit)
CPU	Inter Core 2.20 GHz
Memory	2 GB

46.3.6 Mutation Operator

Suppose the number of genes in chromosome is N. Randomly generate two mutation positions from parents and swap their value. Execute such a transform operation $N-1$ times.

46.3.7 Size of Population

Let size of the population be M. If M is small, it could increase arithmetic speed of GA, but reduce the diversity of the population which may cause premature phenomenon of GA. On the contrary, if M is too big, it will reduce the efficiency of the algorithm. In general, we set $M = 100$.

46.3.8 Termination Conditions

The algorithm adopts the method of setting the number of evolutions. It stops the evolution when the evolution reaches a certain quantity. By default the number of evolutions is the size of the population.

46.4 Simulation Experiment and Analysis

CAGA is programmed by Java. The length of chromosome encoding is N. We perform a simulation experiment in $N = 4, 8, 16, 24$. Hardware and software environments used in the algorithm is as follows (Table 46.1):

Table 46.2 shows the experimental data which comes from simulation program.

As can be seen from Fig. 46.1, when the scale of multi-join is small, the average time of SGA and CAGA is greater than the exhaustive search algorithm. However when the scale of multi-join is large (in our case $N > 8$), the average optimize performance of SGA and CAGA is significantly better than the

Table 46.2 Simulation data

Number of joins	Exhaustive search algorithm optimization time (s)	SGA optimization time (s)	CAGA optimization time (s)
$N = 4$	0.33	0.45	0.44
$N = 8$	1.00	1.02	1.02
$N = 16$	5.21	3.50	3.25
$N = 24$	12.53	8.46	6.16

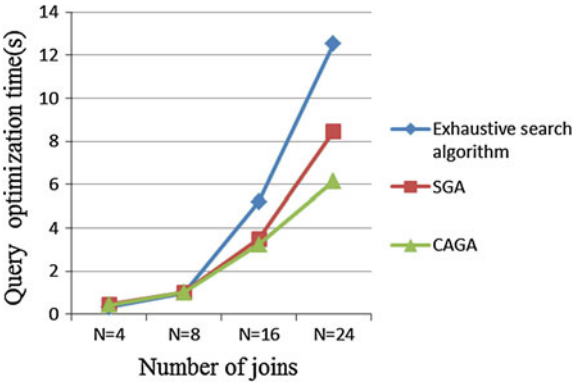


Fig. 46.1 Comparison of various algorithms' optimization time

exhaustive search algorithm. Simulation results show that CAGA converges at a faster rate than SGA and shows the improvement in performance.

Acknowledgments This work was supported by the Co-Funding Project of Beijing Municipal Education Commission under Grant No. JD100060630.

References

1. Ahmed R, Lee A, Witkowski A, et al (2006) Cost-based query transformation in Oracle. Proceedings of the 32nd international conference on very large data bases, VLDB Endowment, 1026–1036
2. Swami A (1989) Optimization of large join queries: combining heuristics and combinatorial techniques, ACM SIGMOD Record, ACM, 18(2):367–376
3. Graefe G (1993) Query evaluation techniques for large databases. ACM Comput Surv (CSUR) 25(2):73–169
4. Bennett KP, Ferris MC, Ioannidis YE (1991) A genetic algorithm for database query optimization. Computer Sciences Department, University of Wisconsin, Center for Parallel Optimization
5. Guo L (2008) Research of query rewriting and multi-join query optimization based on GA of database. Northeastern University (in Chinese)
6. Guo C, Zhu L, Li X (2009) Multi-join query optimization method based on ant colony algorithm. Comput Eng, 35(10):173–175 (in Chinese)

7. Song L (2009) Research and application on multi-join query optimization of database based on genetic algorithm and simulated annealing. Changchun University of Technology (in Chinese)
8. Zhao X (2010) Design and implementation of query optimization algorithm based on genetic tabu search based on stack and chosen the superior. South China University of Technology (in Chinese)
9. Srinivas M, Patnaik LM (1994) Adaptive probabilities of crossover and mutation in genetic algorithms. IEEE Trans Syst Man Cybern 24(4):656–667
10. Liu Y (2008) Application and research on multi-join query optimization of database based on GA. Daqing Petroleum Institute (in Chinese)
11. Dai C, Zhu Y, Chen W (2007) Adapative genetic algorithm based on cloud theory. Control Theory Appl 24(4):646–650 (in Chinese)

Chapter 47
Dynamic Energy-Efficient Virtual Machine Placement Optimization for Virtualized Clouds

Xiaoqing Zhang, Qiang Yue and Zhongtang He

Abstract A virtual machine placement strategy based on the trade-off between energy consumption and SLA is presented. Aiming at dynamical changes of workload requirements, a self-adaptive placement strategy RLWR based on robust local weight regression is presented, which could decide the overload time of hosts dynamically. After detecting overloaded hosts, one virtual machine migration selection algorithm MNM is proposed. The MNM's objective is to get minimal migration number. The migrated virtual machines are deployed using bin-packing algorithm PBFDH. The experimental results show that our algorithm has obvious advantages than other algorithms.

Keywords Cloud computing · Virtual machine placement · Energy consumption

47.1 Introduction

Cloud computing [1, 2] could leverage virtualization technology allowing users to provision resources on-demand on pay-as-you-go. The proliferation of cloud computing has resulted in the establishment of large-scale data centers consuming enormous amounts of energy. The reason for this high-energy consumption is not just in the amount of computing resources used, but rather lies in the inefficient

X. Zhang (✉)
School of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan, China
e-mail: zxqtzy@126.com

Q. Yue
G-CLOUD Technology Co. Ltd., Dongguan, China

Z. He
Chinese Academy of Science, Cloud Computing Center, Dongguan, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_47,
© Springer-Verlag Berlin Heidelberg 2014

usage of resources. Another problem is that idle servers still consume about 70 % of the peak power. Moreover, by using live migration the virtual machines can be dynamically consolidated on the minimal number of physical nodes according to their current resource requirements. However, modern service applications often experience highly variable workloads causing dynamic resource usage patterns. Therefore, aggressive virtual machines consolidation can lead to performance degradation. Ensuring reliable QoS defined via Service Level Agreements (SLA) is essential for cloud computing. Therefore, cloud providers have to deal with the energy-performance trade-off. Recent works on reducing the energy consumption in cloud includes heuristic method for load balance across geographically distributed data centers [3]. The literature [4] presented a three-competitive algorithm for request distribution over the servers of a data center to provide power proportionality. The literature [5] analyzed an online algorithm for resource allocation for a sequence of request arriving to a cloud data center. The literature [6] proposed and proved the competitiveness of an online algorithm for dynamic power management of a physical host with multiple power states.

Our works differ from the above works, in addition to reducing energy consumption, we will also focus on the performance of applications. For reducing energy consumption with meeting SLA, we split the problem of the virtual machine placement optimization into several subproblems, including (1) location strategy, (2) selection strategy, and (3) allocation strategy.

47.2 Power Consumption Model

Comparing with other resources, such as memory and disk storage, CPU consumes more energy in cloud data center. Hence, we will focus on the power consumption of CPU in this paper. The CPU utilization is often proportional to the overall system load, a linear power-to-frequency relationship for a node is applied for simplifying the model. The studies show that an idle node consumes about 70 % of the power consumed by the node running at the full CPU speed on average, which justifies the way of switching idle nodes to the sleep mode to reduce the total power consumption [7]. We apply the following power model,

$$P(u) = k \times P_{\max} + (1 - k) \times P_{\max} \times u \tag{47.1}$$

where, $P_{\max}$ is the maximum power consumed when the physical node is fully utilized, k is the fraction of power consumed by the idle server, and u is the CPU utilization. We will set different $P_{\max}$ in our simulation experiments.

The CPU utilization u is a function of time and is represented as $u(t)$. Then, the total energy consumption E of a physical node can be defined as an integral of the power consumption function over a period of time.

$$E = \int_{t_0}^{t_1} P(u(t))\mathrm{d}t \tag{47.2}$$

47.3 Virtual Machine Migration and SLA Violation

Definition 1 The performance degradation experienced by virtual machine i is defined as (47.3),

$$D_i = \int_{t_0}^{t_0+T_i^m} u_i(t)\mathrm{d}t \times 10\,\%, T_i^m = RAM_i/BW_i \tag{47.3}$$

where, D_i is the total performance degradation by virtual machine i, t_0 is the time when the migration starts, T_i^m is the time taken to complete the migration, $u_i(t)$ is the CPU utilization by virtual machine i, RAM_i is the amount of memory used by virtual machine i, and BW_i is the available network bandwidth.

Definition 2 Overall, SLA violation (SLAV) is defined as a fraction of the difference between the requested MIPS by all the virtual machines $U_{j,r}(t)$ and the actually allocated MIPS $U_{j,a}(t)$ relatively to the total requested MIPS over the lifetime of the virtual machines,

$$\mathrm{SLAV} = \sum_{j=1}^{M} \int_t [U_{j,r}(t) - U_{j,a}(t)]\mathrm{d}t \Big/ \sum_{j=1}^{M} \int_t U_{j,r}(t)\mathrm{d}t \times 100\,\% \tag{47.4}$$

where, M is the number of virtual machines.

47.4 The Design of Algorithm

47.4.1 The Virtual Machine Placement Model

The virtual machine placement can be considered as a bin-packing problem with variable bin sizes and prices, where bins denote the physical nodes, items represent the virtual machines that have to be allocated, bin sizes are the available CPU capacities of physical nodes and prices correspond to the power consumption. The goal is to minimize the price of bin packing. An improved Best-Fit decreasing heuristic based on power aware (PBFDH) is presented to solve the virtual machines placement. The main idea is: we sort all the virtual machines in the decreasing order of current CPU utilizations and allocate each virtual machine to a node that provides the least increase of the power consumption caused by the allocation. PBFDH takes full advantage of the nodes heterogeneity by selecting the most power-efficient ones first. The following is the process of PBFDH

```
Input : Hostlist H = {h_1, h_2, ..., h_m}, VMlist V = {v_1, v_2, ..., v_n}
Output : V → H
SortDecreasing Utilization(V)
For v_i in V
    initialize P_min ← MAX
    HV ← ∅
    For h_j in H
        If h_j has resource to meet v_i
            P(u) = k × P_max + (1 − k) × P_max × u
            If P < P_min
                HV ← h_j
                P_min ← P
    If HV ≠ ∅
        h_j ← HV(v_i)
Return V → H
```

47.4.2 The Detection of Overloaded Hosts

The detection of hot spots is to reduce the workload of overloaded nodes. The traditional algorithm is STA (Static Threshold Algorithm). It is based on the idea of setting upper utilization thresholds for nodes and keeping the total utilization of the CPU by all virtual machines under the threshold [8]. If the CPU utilization of a node exceeds the static threshold, some virtual machines have to be migrated from the node to reduce the CPU utilization in order to prevent a potential SLA violation, as shown in Fig. 47.1.

However, static fixed thresholds are not suitable for dynamic and unpredictable workloads, in which different types of applications can share a physical resource of cloud. Therefore, a new detection mechanism of hot physical node based on robust locally weighted regression named RLWR are proposed, which can adjust the value of the static upper utilization threshold based on a statistical analysis of historical data collected during the lifetime of virtual machines.

For each i let h_i be the distance from x_i to the r-th nearest neighbor of x_i. That is, h_i is the r-th smallest number among $|x_i - x_j|$, for $j = 1, 2, \ldots, n$. For $k = 1, 2, \ldots, n$, let

$$w_k(x_i) = W(x_k - x_i/h_i) \tag{47.5}$$

The following is the process of RLWR.

Step 1. For each i compute the estimate, $\hat{\beta}_j(x_i), j = 0, 1, \ldots, d$, of the parameter in a polynomial regression of degree d of y_k on x_k, which is fit by weighted least squares with weight $w_k(x_i)$ for (x_k, y_k). Thus, the $\hat{\beta}_j(x_i)$ can be computed from the following equation:

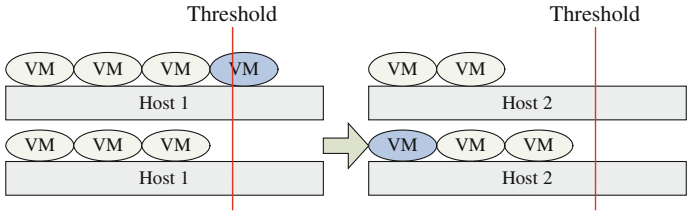


Fig. 47.1 STA

$$\left(\hat{\beta}_0(x_i), \hat{\beta}_1(x_i), \ldots, \hat{\beta}_d(x_i)\right) = \arg\min \sum_{k=1}^{n} \omega_k(x_i)\left(y_k - \beta_0 - \beta_1 x_k - \cdots - \beta_d x_k^d\right)^2 \tag{47.6}$$

When computing the initial fitted value, a triple cube weight function is applied to computing $w_k(x_i)$:

$$W(\omega) = \begin{cases} \left(1 - |\omega|^3\right)^3, & |\omega| < 1 \\ 0, & |\omega| \geq 1 \end{cases} \tag{47.7}$$

The smoothed point at x_i using locally weighted regression of degree d is $(x_i, \hat{y}_i)$, where $\hat{y}_i$ is the fitted value of the regression at x_i. So

$$\hat{y}_i = \sum_{j=0}^{d} \hat{\beta}_j(x_i) x_i^j = \sum_{k=1}^{n} r_k(x_i) y_k \tag{47.8}$$

where, $r_k(x_i)$ does not relay on y_j, $j = 1, 2, \ldots, n$, denotes the coefficient for the y_k that arise from the regression.

Step 2. $B(\omega)$ is defined as the bisquare weight function,

$$B(\omega) = \begin{cases} (1 - \omega^2)^2, & |\omega| < 1 \\ 0, & |\omega| \geq 1 \end{cases} \tag{47.9}$$

Let $e_i = y_i - \hat{y}_i$ be the residuals from the current fitted values. Let s be the media of $|e_i|$. $s = \text{med}|e_i|$. Define robustness weight by

$$\delta_k = B(e_k / 6s) \tag{47.10}$$

Step 3. For each i, computing new $\hat{y}_i$ by fitting a d-th degree polynomial using weighted least squares with weight $\delta_k w_k(x_i)$ at (x_k, y_k),

$$\left(\hat{\beta}_0(x_i), \ldots, \hat{\beta}_d(x_i)\right) = \arg\min \sum_{k=1}^{n} \delta_k w_k(x_i)\left(y_k - \beta_0 - \beta_1 x_k - \cdots - \beta_d x_k^d\right)^2 \tag{47.11}$$

Step 4. Repeatedly carry out Step 2 and 3 for T times. The final $\hat{y}_i$ are robust locally weighted regression fitted values.

$$\hat{y}_i = \sum_{j=0}^{d} \hat{\beta}_j^T(x_i) x_i^j \tag{47.12}$$

Step 5. Using $\hat{y}_i$ from Step 4 to forecast the next $\hat{y}_{i+1}$ and deciding whether the physical node is overloaded through the following system of inequalities.

$$\varepsilon \cdot \hat{y}_{i+1} \geq 1, x_{k+1} - x_k \leq t_m \tag{47.13}$$

where, $\varepsilon \in R^+$ is the safety factor, which defines the risk index of RLWR and how aggressively the system consolidates virtual machines. ε allows the adjustment of the safety of the algorithm, the lower ε, the less the energy consumption, the highest the level of SLAV caused by the consolidation.

47.4.3 The Selection Algorithm of Virtual Machine Migration

Aiming at the cost existing in virtual machine migration, we proposed MNM algorithm, which selects the minimum number of virtual machines needed to migrate from a physical node to lower the CPU utilization below the upper utilization threshold if the upper threshold is violated. Let V_j be a set of virtual machines currently allocated to the node j, $Q(V_j)$ is the power set of V_j. The set R of migrated virtual machines searched by MNM is

$$\begin{aligned} &min|R| \\ &s.t. \begin{cases} R \in Q(V_j) \\ u_j - \sum_{v \in R} u_a(v) < T_u \\ Q(V_j) = \{U | U \subseteq V_j\} \end{cases} \end{aligned} \tag{47.14}$$

where, u_j is the current CPU utilization of the node j, $u_a(v)$ is the fraction of the CPU utilization allocated to the virtual machine v.

MNM sorts the list of virtual machines in the decreasing order of the CPU utilization. Then, it repeatedly looks through the list of virtual machines finds a virtual machine that is the best to migrate from the host.

47.4.4 The Detection of Underloaded Hosts

The idea of the detection of underloaded hosts is: the system finds a host with the minimal utilization compared with the other hosts, and tries to place all the virtual machines from this host on other hosts, while keeping them not overloaded. If such

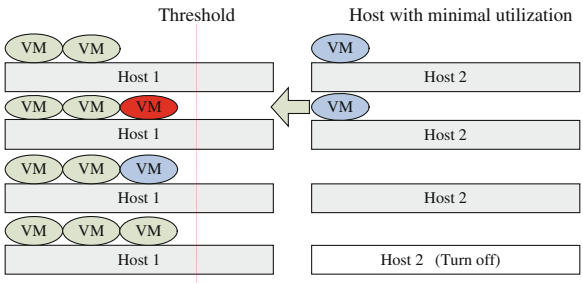


Fig. 47.2 Underloaded host

a placement is feasible, the virtual machines are set for migration to the target hosts. Once the migrations are completed, the source host is switched to the sleep mode to save energy consumption. If all the virtual machines from the source host cannot be placed on other hosts, the host is kept active. Repeat this process iteratively for all non-overloaded hosts. The process is shown in Fig. 47.2.

47.5 Experimental Results

The cloud environment includes a data center consisting of N heterogeneous nodes. According to different process power of CPU, these nodes are divided equally into two types, (1,800 MIPS and 2,600 MIPS). Two types of nodes both include two process elements, 4 G RAM and 1 GB/s network bandwidth. In order to compare the efficiency of algorithms, we use the following metrics to evaluate their performance, including the total energy consumption, the number of virtual machine migration, the overall SLA violation, and the number of shutdown. RC is selected as a comparison algorithm with MNM. RC is a random choice algorithm, which depends on a random selection of a number of virtual machines needed to reduce the CPU utilization by a node below the fixed utilization threshold. According to a uniformly distributed discrete random variable (X), whose values index subsets of V_j. The virtual machine set R that RC algorithm is looking for is:

$$X \stackrel{d}{=} U\left(0, \left|Q\left(V_j\right)\right| - 1\right) \rightarrow R \tag{47.15}$$

Figures 47.3, 47.5, and 47.6 show that the changes of the performance metric of algorithms under different placement scale of virtual machines. In STA, the threshold is set to 0.8. The safety factor $\varepsilon = 1.2$. In RLWR, the fitted polynomial $d = 1$, the number of iterations $T = 2$, smoothing factor $f = 1/3$, $W(x)$ is a triple weight function.

Figure 47.3 shows the changes of the total energy consumption of algorithms. With the expansion of the placement scale, the energy consumption increases. MPA and DVFS obviously consume more energy than other algorithms, because these two algorithms do not consider the migration and consolidation of virtual

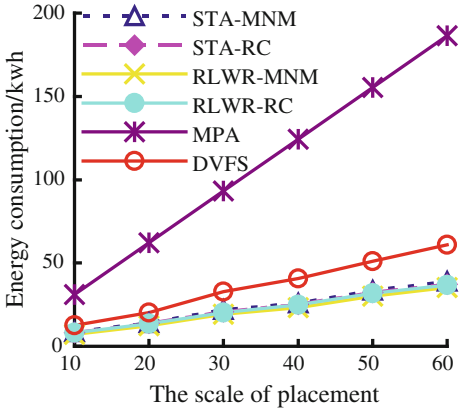


Fig. 47.3 Energy consumption

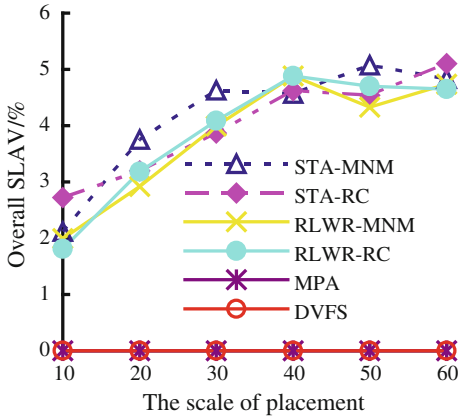


Fig. 47.4 Overall SLA violation

machines. MPA has highest energy consumption due to the CPU utilization of all physical nodes keeps 100 %. Within the algorithms of considering the virtual machines migration, the energy consumption of RLWR is less than STA. Because RLWR can forecast the workloads based on a statistical analysis of historical data collected, while STA cannot reflect the dynamic changes of workloads.

We can see that MPA and DVFS both have no SLA violation, this is because the amount of requested resource equal to the amount of allocated resource in Fig. 47.4. Relatively speaking, the migration number of MNM is minimal, so SLAV is minimal. Meanwhile, the overall SLAV of RLWR is less than STA in general.

Figure 47.5 shows the migration number of virtual machines of algorithms. It seems that the migration number of RLWR is less than STA. This is because STA cannot reflect the dynamic changes of workloads of cloud computing environment. The migration number of MNM is minimal in two hotspot detection mechanism.

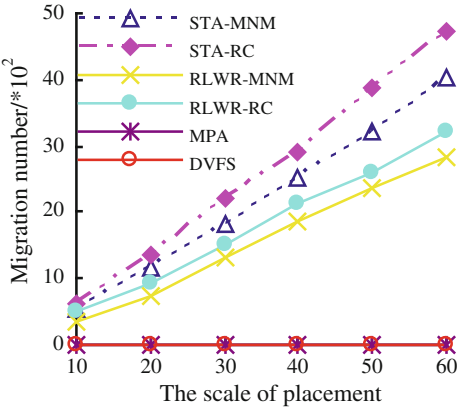


Fig. 47.5 Migration number of virtual machine

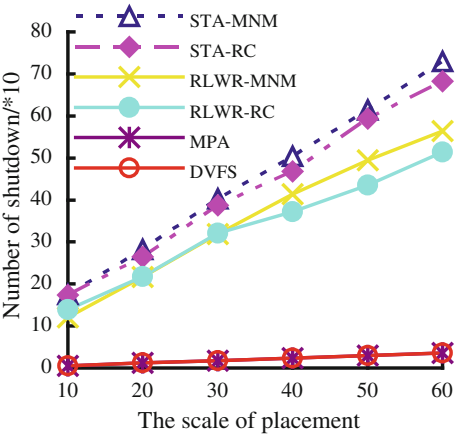


Fig. 47.6 Number of shutdown

From Fig. 47.6 we can see that when using different virtual machines selection algorithm, the number of shutdown of RLWR is roughly approximate to STA. However, RLWR can forecast the workloads, the number of shutdown of RLWR is appreciably less than STA.

47.6 Conclusions

Dynamic virtual machines placement using live migration and switching idle nodes to the sleep mode allows cloud providers to optimize resource usage and reduce energy consumption. But, the obligation of providing high QoS to users leads to the necessity in dealing with the energy-performance trade-off. For solving this problem, a self-adaptive virtual machine placement algorithm is

proposed. Our algorithm can complete dynamic virtual machines consolidation based on an analysis of historical data from the resource usage and greatly reduces energy consumption, while ensuring a high level of adherence to the SLA.

References

1. Vaquero LM, Rodero-Merino L, Caceres J (2008) A break in the clouds: towards a cloud definition. SIGCOMM Comput Commun Rev 39(1):50–55
2. Buyya R, Chee Shin Y, Venugopal S (2008) Market-oriented Cloud computing: vision, hype, and reality for delivering IT services as computing utilities. In: Proceedings of 10th IEEE conference on high performance computing and Communications, pp 5–13
3. Lin M, Wierman A, Andrew LLH (2012) Online algorithm for geographical load balancing. In: Proceedings of 3rd international green computing conference, pp 1–10
4. Lin M, Wierman A, Andrew LLH, Thereska E (2011) Dynamic right-sizing for power-proportional data centers. In: Proceedings of 30th IEEE international conference on computer communication, pp 1098–1106
5. Plaxton CG, Sun Y, Tiwari M (2006) Reconfigurable resource scheduling. In: Proceedings of 18th annual ACM symposium on parallelism in algorithm and architecture, pp 93–102
6. Irani S, Gupta R, Shukla S (2002) Competitive analysis of dynamic power management strategies for systems with multiple power savings states. In: Proceedings of design, automation and test in Europe, pp 117–123
7. Fan X, Weber W-D, Barroso L (2007) Power provisioning for a warehouse-sized computer. In: Proceedings of 34th annual international symposium on computer architecture, pp 13–23
8. Fang Y, Tang D, Ge J (2012) Energy-aware schedule strategy based on dynamic migration of virtual machines in cloud computing. J Comput Inf Syst 8(10):4201–4208

Chapter 48
The Research on Distortion Correction Algorithm of Unwrapping the Cylinder Image of Panoramic Annular Lens

Qiuxiang Liao

Abstract Panoramic Annular Lens (PAL) is a new type of optical imaging system, achieving one-time image formation without scanning. It has more incomparable advantages than the traditional panoramic imaging optical system in large field of view. However, since PAL is based on the principle of Flat Cylinder Perspective (FCP) , so the cylinder image of PAL has inevitably distortion in both tangential and radial image. This article has a profound study on the image formation principle of the PAL system, and unwraps it in tangential image. Because PAL is based on the principle of $f - \theta$ optical lens, there exists the compression in the radial image. This article adopts the least squares estimate algorithm to make revision, and the results of experiments showed that the effect was satisfactory.

Keywords Panoramic annular lens (PAL) · Flat cylinder perspective (FCP) · The least squares estimate algorithm · Distortion correction

48.1 Introduction

Panoramic imaging has incomparable advantage in increasing the field of view, so it has important applications in some nondestructive fields such as robot vision under water and on land [1], pipeline detection and medical endoscopic examination, etc. At the same time, it has a very broad application prospects in many areas including national defense, aviation, medical and civilian use, and more [2].

Traditional optical system follows the central convergence perspective (CCP). In order to obtain panoramic imaging of 360°, pictures must be taken from different angles by scanning devices placed in the imaging system, or by the adoption

Q. Liao (✉)
Wuzhou University, 543002 Wuzhou, Guangxi, China
e-mail: 306804424@qq.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_48,
© Springer-Verlag Berlin Heidelberg 2014

of many imaging equipments, and the pictures are stitched into panoramic imaging [3]. In this way, not only the cost of the imaging is raised, but also the entire field of view cannot be observed in real time. At the same time, the information of image stitching in later stage is also very huge.

Panoramic annular lens (PAL) proposed by P. Greguss [4] in the patent in 1986 is a new kind of design scheme. And with PAL as basis, the PAL imaging system is developed a totally new optical imaging system which is different from traditional optical system, achieving one-time image formation without scanning. It has incomparable advantages in growing the field of view than traditional panoramic optical system. But because the PAL imaging system is based on the flat cylinder perspective (FCP) and follows the principle of $f - \theta$ lens imaging, distortion exists in the tangential and radial images, thus the distortion correction is needed.

In this article, the least squares estimate algorithm is used for distortion correction of PAL's cylinder image. The algorithm is of small amount of calculation, and the experimental results show that visual effect of the revised image is satisfactory.

48.2 PAL Imaging Principle

PAL is based on FCP and follows the imaging principle of $f - \theta$ lens, which will be introduced respectively in the following text.

48.2.1 Flat Cylinder Perspective, FCP

FCP is a method to project the scene within the cylinder field of view around the optic axis of 360° to a circular region on two-dimensional plane [5], and its principle is shown in Fig. 48.1.

PAL is based on FCP. In FCP mapping, the three dimensional space formed by two edges of angle α rotated 360° around the axis z is a range capable for imaging. And this three-dimensional space is projected to a circular region on the two-dimensional image plane after PAL, as the shadow part on the right side chart shown in Fig. 48.1. And the part not available for imaging is the cone formed by two edges of angle 2β rotated 360° around the axis z, which is projected to a round part on two-dimensional plane, corresponding to the blank round part on the right side chart in Fig. 48.1, that is a blind area of PAL. From Fig. 48.1—principle of FCP, it is known that for increasing the imaging field, we can consider increasing α and reducing β, but values of these two parameters are constrained by the existing refractive index of glass [6].

In traditional CCP, parallel lines from different directions will be gathered at different points on a horizontal line. But in FCP mapping, all parallel lines will be focused on one point, so each concentric circle in the two-dimensional circular

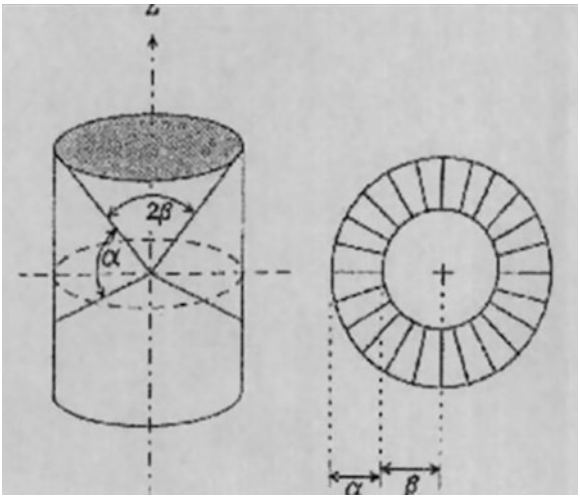


Fig. 48.1 Principle of flat cylinder perspective

imaging plane is a trajectory of image points with the same angle of optical axis. According to the principle of FCP, in fact, the PAL gives up the field of view in front in order to achieve the cylindrical field of imaging around the optical system.

48.2.2 $2f - \theta$ Lens

In general optical imaging system, the ideal image height for imaging is focus length multiplying by tangent of filed angle, which satisfies the equation $y' = f' \tan\theta$. On both sides of the equation, perform the time differential at the same time, there is

$$\frac{dy'}{dt} = f' \sec^2\theta \frac{d\theta}{dt} \tag{48.1}$$

Obviously, the above equation shows the differences between the information recorded by camera and the original information. As seen from the equation, scanning speed, which is obtained in the focal plane from the incident light deflected by the same angular velocity, is a variable rather than a wanted constant, that is to say different surface interval $\delta y'$ will be achieved at the same angle interval $\delta\theta$. In order to solve this problem, we consider adopting $f - \theta$ lens which is a commonly used linear scanning system in laser scanning system to make the record information in accordance with the original information and the light spot on the image surface in accordance time with a one-to-one relationship. Image height of $f - \theta$ Lens and scan angle are in a linear relationship, namely they satisfy the Eq. (48.2).

$$y' = f'\theta \tag{48.2}$$

On both sides of the equation, perform the time differential, there is

$$\frac{dy'}{dt} = f'\frac{d\theta}{dt} \tag{48.3}$$

That is, scanning speed on the focal plane is constant when the incident light is deflected by the same angular velocity, so the record information is consistent with the original information. But distortion will occur in the imaging of $f-\theta$ lens, when the scanning angle θ increases, the image height of $f-\theta$ lens is less than the ideal height determined by geometrical optics, the $\theta/\tan\theta$ times of the ideal height, namely the image compression phenomenon exists.

PAL follows the principle of $f-\theta$ imaging, so all the optical system design on PAL should take into full consideration the specific characteristics of the $f-\theta$ lens.

48.3 Distortion Correction Algorithm of Unwrapping About PAL's Cylindrical Image

Algorithm in this article aims at *r* PAL's cylindrical image, test of cylindrical annular image are chosen from Hou Huijie's (Institute of Information, Zhejiang university) master thesis in 2006—*The unwrapping algorithm research about PAL's cylindrical image*, as shown in Fig. 48.2. Considering there are two aspects for the observed panoramic annular image to unwrapped, one is in tangential, and the other is in radial, so we need to discuss them separately.

48.3.1 Unwrapping in Tangential

Three different patterns are used in test image including square, board, and concentric circles. In order to unwrap it to the familiar rectangular area, the annular image must be scanned according to the principle of PAL imaging. Because the PAL follows the principle of FCP, each concentric circles with the same radius is a trajectory with the same angle of the lens optical axis, namely all the image points on the concentric circles of same radius corresponding to the object points on the same horizontal level of the object plane. According to this feature, circular track is scanned from the circle of minimum radius to the circle of the maximum radius, from inside to outside according to the same radius.

Before scanning, the length and width of a rectangular area must be determined firstly to ensure the image's fidelity as much as possible. The selected length is the circumference of the outside circle with maximum radius to ensure not to lose

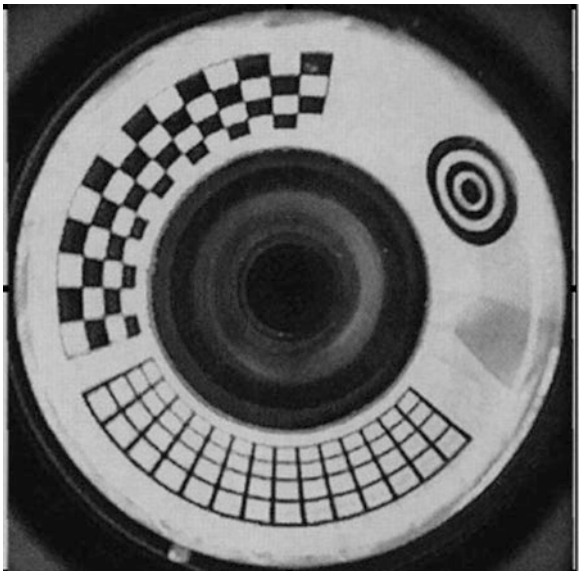


Fig. 48.2 The test of cylindrical annular image

image pixel information uttermost, while the width is the width of the annular image region, namely the difference between the exradius of the maximum circle and the minimum circle in the annual imaging area.

Based on the analysis above, the key for the annular image to be unwrapped is to determine the size of the inner and outer radius to form the annular image. The built-in function of MATLAB is used here, where the size of the inner and outer radius for unwrapped annular image can be chosen by users to realize the man–machine interaction. At the same time, the annular region size to be unwrapped also can be selected according to the requirements, which makes the program more flexible and convenient.

For the image is a standard annular imaging, namely the annular imaging is at the center of the panorama image, suppose the center coordinate is (o_x, o_y), the line height is h and the column width is l, then there is $o_x = 1/2h, o_y = 1/2l$. When the circle center is determined, the size of the circle radius inside and outside can then be calculated by computer through reading the coordinates of two points selected by the user on the inside and outside annulus. And the radius of the circle is the distance between the point on the circle and the center.

In tangential, we select the circumference of the biggest outer circle as the length of the unwrapped rectangular region. And because all the inner circumferences are smaller than the outer circumference and pixel values are read according to the circular trajectory, the pixel points must be insufficient during the unwrapping process. For the principle of equivalent division to all the concentric circles, namely all the concentric circles are divided into the number with the same length of the unwrapped area, so it does not have to consider the complex interpolation problem, thus making the problem easier. In this algorithm, each

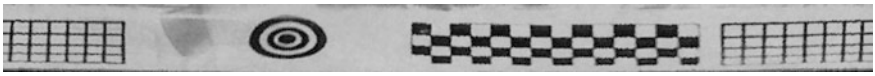


Fig. 48.3 Unwrapped image of the test of cylindrical annular image

interpolation of output pixel value is the value of its most adjacent sampling point in the input image. The experimental results show that this interpolation method is of a small amount of calculation, and the unwrapped image is of high quality. Unwrapped image of the test of cylindrical annular image is shown in Fig. 48.3.

From Fig. 48.3, it is seen that the unwrapped effect of PAL image is good in radial, but with obvious compression phenomenon. The original square becomes a rectangle and the original standard round also becomes flat, so the distortion exists on the image in the radial, which requires distortion correction.

48.3.2 Radial Distortion Correction

For distortion in radial, radial direction should be ensured constant, namely if using coordinates to indicate, x is constant while y changes correspondingly. Suppose the original image coordinates are (x, y), the image coordinates after radial distortion are (x', y'), then Eq. (48.4) is satisfied.

$$x' = x, y' = f(y) \tag{48.4}$$

Therefore, the discovery of the function f is the key in distortion correction. Because the test image adopts the square, the unwrapped image must also be a square, rather than the rectangle in the unwrapped image in Fig. 48.3. Considering the image distortion is not a simple linear relationship, we use the function approximation to perform curve fitting to function f. Curve fitting does not require the value of approximation function to be the same as the function at the node, namely the approximate curve does not need to pass through the known point, but only requires the function to reflect the basic trend of the given data points as far as possible [7]. This article selects the common least squares estimate algorithm with simple calculation, which is an approximation function method commonly used in practice.

Problems in the least squares estimate algorithm are: according to the given data set (x_i, y_i) $(i = 1, 2, \ldots, n)$, the approximate function form is selected, that is a given function class H, and the function $f(x) \in H$is solved to set up Eq. (48.5).

$$\sum_{i=1}^{n} [y_i - f(x_i)]^2 = \min_{f \in H} \sum_{i=1}^{n} [y_i - f(x_i)]^2 \tag{48.5}$$

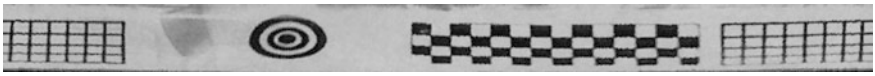


Fig. 48.4 Unwrapped result before the radial distortion correction

Then the function $f(x)$ is named as the least squares function of this set of data. The polynomial fitting, especially the simplest least-squares quadratic estimate polynomial, is selected in this article, and the results show that the effect is ideal.

For a given data set (x_i, y_i) $(i = 1, 2, \ldots, n)$, to solve a quadratic polynomial.

$$f(x) = a_0 + a_1 x + a_2 x^2 \tag{48.6}$$

According to the numerical calculation, coefficient of the least squares estimate polynomial should satisfy the following equation set:

$$na_0 + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 + \ldots + a_m \sum_{i=1}^{n} x_i^m = \sum_{i=1}^{n} y_i$$

.....

$$a_0 \sum_{i=1}^{n} x_i^m + a_1 \sum_{i=1}^{n} x_i^{m+1} + a_2 \sum_{i=1}^{n} x_i^{m+2} + \ldots + a_m \sum_{i=1}^{n} x_i^{2m} = \sum_{i=1}^{n} y_i x_i^m$$

The equation set above is named as regular equation set, which is proved to have the only solution through the linear independence of the function group. Polynomial corresponded to the solution must be the least-squares polynomial fitted by m times of the given data set (x_i, y_i) $(i = 1, 2, \ldots, n)$.

According to the principle above, we need to get a set of data (x_i, y_i) from the unwrapped figures, where x_i is the coordinate y of the original figure, and y_i is coordinate y of the processing figure. We define coordinate values at 6 point through the built-in function in MATLAB. The first and second coordinate values are used to determine the width of the square, and the remaining four points are used in the least squares estimate algorithm. Although the amount of data is small, the fitting result is ideal.

Finally coefficient of square quadratic polynomial function is

$$a_0 = 5.95846, a_1 = 0.61003, a_2 = 0.015027$$

48.4 Experimental Results

Results of the least squares estimate algorithm are shown as below, and the unwrapped result after the radial distortion correction was seen in Fig. 48.5. To facilitate the observation and comparison, unwrapped result before the radial distortion correction is also put forward in Fig. 48.4.

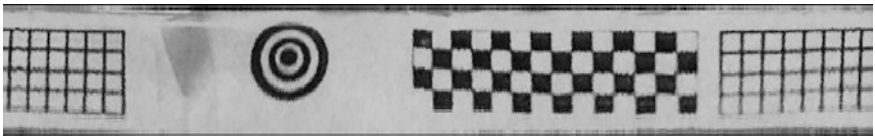


Fig. 48.5 Unwrapped result after the radial distortion correction

From the unwrapped results above, after the radial distortion correction, there is obvious radial amplification of the unwrapped image compared with the image before fitting. And the compression of the image before fitting then can be corrected. From the perspective of revision, the effect is also good. The original square is able to be restored after correction and the oblate in the original image is restored into a standard circle.

References

1. Zhigang Z, Karuppiah DR (2004) Keeping Smart, Omnidirectional Eyes on You: Adaptive Panoramic Stereovision for Human Tracking and Localization with Cooperative Robots. IEEE Robot Autom Mag, 69–78
2. Kopilovic I, Vagvolgyi B, Sziranyi T (2000) Application of panoramic an-nular lens for motion analysis tasks: surveillance and smoke detection, 0-7695-0750-6, IEEE 2000, pp 714–717
3. Xiao X, Guoguang Y, Jian B (2008) Enhancement of panoramic image resolution based on swift interpolation of Bezier surface. Opto-Electron Eng 35(1):105–109 (in Chinese)
4. Huijie H (2006) The algorithm research of unwrapping about panoramic annular lens (PAL)'s image. Zhejiang University Institute of Information Optics Engineering, Zhejiang (in Chinese)
5. Xiao X, Guoguang Y, Jian B (2007) Panoramic image restoration based on optimal parameter cubic spline interpolation. Infrared and Laser Eng 36(5):725–728 (in Chinese)
6. Zhu F, Yang G, Weiyong Y, Cheng H (2001) Study for linearity of panoranmic annular image. Acta Photonica Sinica 30(5):589–592 (in Chinese)
7. Lijuan D, Qiyuan C (2005) Numerical methods (second edition). Beijing Institute of Technology press, Beijing, p 143 (in Chinese)

Chapter 49
A Novel Control Strategy of Permanent Magnet Synchronous Machine Drive Under Field-Weakening Operation

Xiangyang Wu, Kekang Wei and Shihua Bi

Abstract Field-weakening technology is important for permanent magnet synchronous machine (PMSM) control in wide speed range applications. This paper presents a novel field-weakening control method for permanent magnet synchronous motor to improve maximum speed range and power range, based on discrete-time complex current control and voltage control. The effectiveness of the proposed method is confirmed by computer simulations and experiments.

Keywords Permanent magnet synchronous machine · Field-weakening control · Discrete-time complex current control · Voltage control

49.1 Introduction

Recently, permanent magnet synchronous motor (PMSM), because of its higher torque density efficiency, and fast torque speed response, has been widely employed in many industrial applications instead of induction motor. Lately, due to advanced electric machine design methods, PMSM has very wide constant power speed range [1, 2], especially it is reported that a constant power speed ratio (CPSR) as high as 9.5 [3]. But it is very difficult to realize field-weakening operation, for wide speed range operation. So there have been many researches on field-weakening control strategy of PMSM that PMSM can be used in high-speed range application such as traction, robotics, or spindle drive for CNC machine tools.

X. Wu (✉) · S. Bi
School of Aerospace Engineering, Beijing Institute of Technology, Beijing, China
e-mail: weikekang@163.com

K. Wei
Beijing Institute of Specialized Machinery, Beijing, China
e-mail: bjtuwei@qq.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_49,
© Springer-Verlag Berlin Heidelberg 2014

Generally speaking, field-weakening strategies of PMSM are based on Field Oriented Control (FOC) with two current regulators, one for i_q and the other for i_d. The main task of field-weakening control is to design the current trajectory on the i_d–i_q plane. It is proposed that the current reference is calculated by the speed and the torque, which considering the maximum torque per ampere (MTPA) operation, the current-limit circle, and the voltage-limit ellipse in every speed of the machine [4, 5]. But this kind of methods based on machine model is sensitive to the machine parameters. In [6], an algorithm of flux-weakening operation using looking-up table whose values obtained by experimental results is presented. This method resolved the parameter problem, but it requires large experimental work. A method of modifying the current reference and field weakening based on gradient descent of voltage limit according to the ellipse were proposed in [7], which consisted of two parts, one is the determination of the field-weakening region; another is the current reference modification. This method is complexity of the implementation. In [3], a new field-weakening control strategy called single current regulator is presented by Yuan zhang and Xu Longya et al. In the single current regulator the q-axis current regulator is removed, and only the d-axis current regulator is used. As reported with the single current regulator the CPSR can be up to 7:1. But the structure of this method is different from the FOC, the transition between this method and FOC is problem which may increase the system complexity. The most common method is to compensate the d-axis current with a voltage regulator. It well resolved the problems aforementioned and also used in this paper.

All the methods above consist of current regulator (one or two current regulator) in inner loop. And the performance of current regulator effect on the field weakening. So stator current controller plays an important role in field weakening, whose performance effect on the performance of motor control such as quickly response and stability. The synchronous reference frame proportional-integral (PI) current regulator (SRF-PI) has been the industrial standard for current regulation in conventional FOC for more than 20 years, which can be able to regulate AC currents over a wide frequency range with high bandwidth and zero steady-state error. Unfortunately, in the synchronous reference frame with the fundamental excitation, the transformation of electric variables such as voltage, current, creates cross-couple between d- and q-axes that is proportional to the fundamental excitation frequency, led to degrade the performance of the current regulator as the fundamental excitation frequency increased [8, 9].

In this paper, a novel flux-weakening control method of an PMSM based on voltage regulator and discrete-time complex current regulator is proposed. The voltage regulator can automatically generate d- and q -axis current commands together with some constraints to keep the operating point close to the voltage-limit curve in the field-weakening region, especially considering DC-link voltage varying. And the discrete-time complex current regulator increases the performance of current inner loop. It is verified by experimental result that the speed range of PMSM is enhanced by the proposed field-weakening algorithm.

49.2 Modeling and the Control System of PMSM

The voltage equations of PMSM in the synchronous reference frame can be written as follows:

$$v_{sd} = R_s i_{sd} + L_d \frac{\mathrm{d}i_{sd}}{\mathrm{d}t} - \omega L_q i_{sq} \tag{49.1}$$

$$v_{sq} = R_s i_{sq} + L_q \frac{\mathrm{d}i_{sq}}{\mathrm{d}t} + \omega L_d i_{sd} + \omega \lambda_m \tag{49.2}$$

where

v_{sd}, v_{sq}	d–q components of stator voltage;
i_{sd}, i_{sq}	d–q components of stator current;
R_s	stator resistance;
L_d, L_q	d–q components of stator self-inductance;
ω	Electrical rotor angle velocity
λ_m	Flux linkage of the permanent magnet.

According to (49.1, 49.2), the PMSM control system based on FOC is designed as shown in Fig. 49.1.

The PMSM is fed by a three-phase voltage-source inverter. The basic control strategy is FOC and the PWM signal is produced by space vector modulation (SVPWM), which incorporating speed control, current vector control, and flux-weakening control. The output of the speed regulator could be the torque reference or the current reference (amplitude with sign) as a function of the adopted flux-weakening algorithm. The speed regulator is a PI controller. The current regulators and the field-weakening control will be elaborated later.

For a practical motor drive system, the operating states should stay within the system limits, considering both motor and inverter ratings. They can be expressed in terms of the current and the voltage constraints, respectively, as follow:

$$i_{sq}^2 + i_{sd}^2 \le I_{sm}^2 \tag{49.3}$$

$$v_{sq}^2 + v_{sd}^2 \le V_{sm}^2 \tag{49.4}$$

where

I_{sm}	the maximum line-current;
V_{sm}	the maximum phase voltage amplitudes.

With the resistance and the transient terms ignored, Substituting (49.1) and (49.2) into (49.4), one can also obtain the voltage constraint which can be represented by an elliptical region as follow:

$$\left(\omega L_q i_{sq}\right)^2 + \left(\omega L_d i_{sd} + \omega \lambda_m\right)^2 \le V_{sm}^2 \tag{49.5}$$

Figure 49.2 shows two voltage-limit ellipses on i_{ds}—i_{qs} plane corresponding the current-limit circle, and two constant torque curves corresponding. As the speed

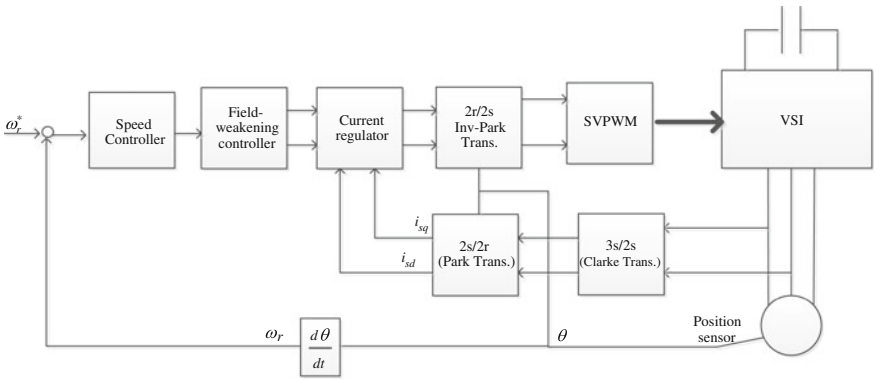


Fig. 49.1 Block d diagram of PMSM drive system

rises, the ellipse becomes smaller and smaller, resulting in smaller operation regions. To produce the maximum output power, the current trajectory is divided into three regions, according to the field-weakening operation state [10]:

Region I The speed is low and field-weakening is not necessary. In this region, MTPA (Maximum Torque per Ampere) or $i_{sd} = 0$ control is used with two current regulator, whose current trajectory is OA or OD line;

Region II The current trajectory moves along the current-limit circle from point A to B as the speed goes up.

Region III As the speed become very high, the voltage ellipse become very small, the current trajectory stays on voltage ellipse. As shown in Fig. 49.2, current vector moves from B to C. here, C is defined as the center of the ellipse.

In all the three regions, SRF-PI is applied as the current regulator.

49.3 Field-Weakening Control

In PMSM drive system is fed by a three-phase voltage-source inverter, whose maximum output-phase voltage value is limited to maximize the utilization of the capacity of DC-link voltage. According to the principle of SVPWM, the maximum phase voltage amplitude is described as follow:

$$\sqrt{v_{sq}^2 + v_{sd}^2} \leq V_{sm} = v_{dc} \Big/ \sqrt{3} \tag{49.6}$$

From (49.5) and (49.6), some conclusion can be deduced as follow:

(1) When the motor speed is low and the voltage is not limited, d- and q-axes current can be determined by respectable axes voltage independently.
(2) When the motor speed is high, as operate in the Region II, the voltage magnitude is fixed, and the current follows the (49.3), or Line AB in Fig. 49.2. And i_{sd} can

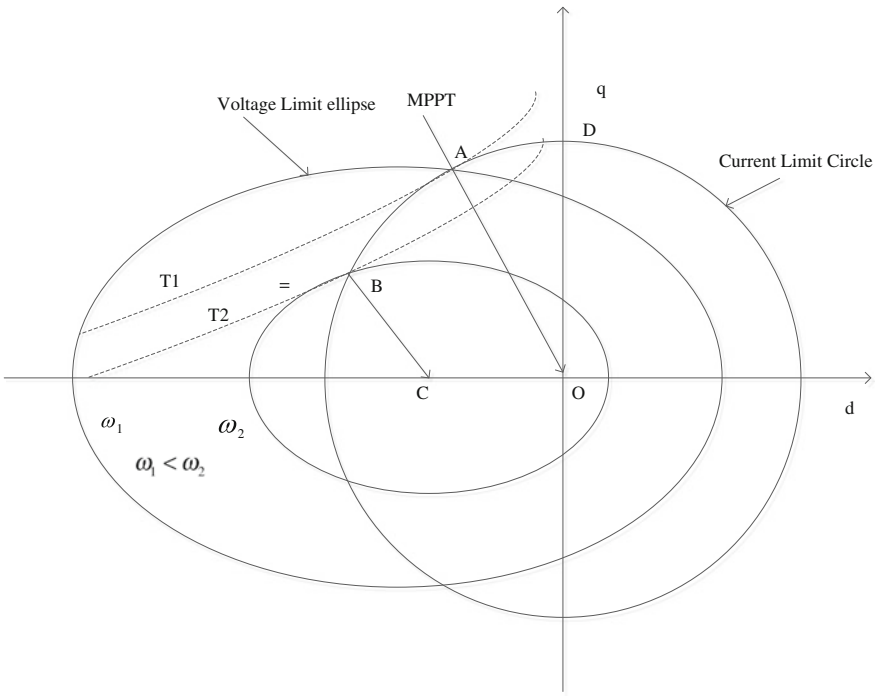


Fig. 49.2 The current and voltage-limit ellipse

be more negative to reduce the magnetic flux more. In this case, i_{sd} and i_{sq} are connected with each other and one of them is determined by the other.

(3) When the motor speed reaches to Region III, the i_{sd} should not be negative more. In this paper, to simplify, i_{sd} is fixed as Eq. (49.7) and just i_{sq} command vary reflect to speed.

$$i_{sd} = -\frac{\lambda_m}{L_d} \tag{49.7}$$

As described above, the control block diagram of field-weakening is shown in Fig. 49.3. The command value of phase voltage is come from the DC-link voltage feedback value, described in (49.6). The value of phase voltage reconstructed by switch function of inverter, as follow:

$$\begin{cases} u_{an} = \dfrac{u_{dc}}{2}\left(\dfrac{2}{3}s_a - \dfrac{1}{3}s_b - \dfrac{1}{3}s_c\right) \\ u_{bn} = \dfrac{u_{dc}}{2}\left(\dfrac{2}{3}s_b - \dfrac{1}{3}s_a - \dfrac{1}{3}s_c\right) \\ u_{cn} = \dfrac{u_{dc}}{2}\left(\dfrac{2}{3}s_c - \dfrac{1}{3}s_a - \dfrac{1}{3}s_b\right) \end{cases} \tag{49.8}$$

where $s_{x,x=a,b,c}$, is the duty of upper power switch in three-phase inverter.

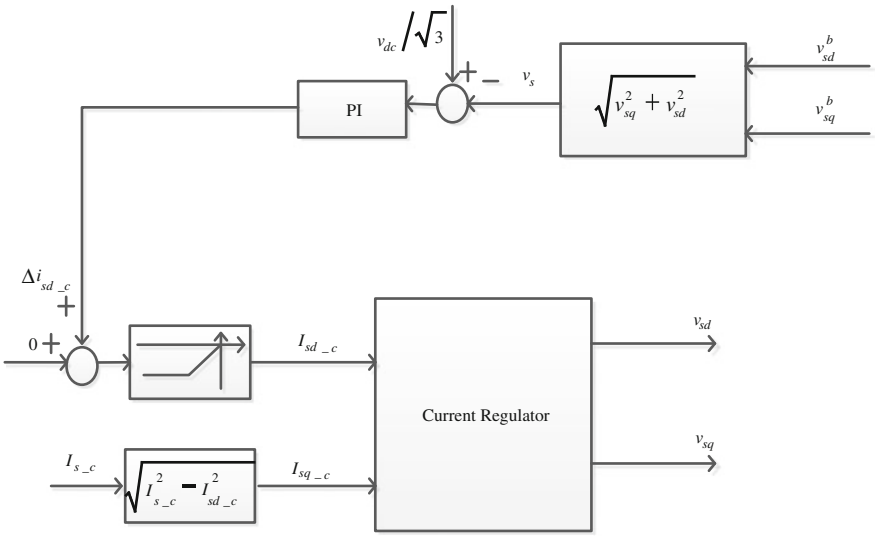


Fig. 49.3 Field-weakening control

49.4 Discrete-Time Complex Current Regulator

Current regulator plays an important role in field-weakening control. As mentioned above, the performance of SRF-PI is degraded with fundamental excitation frequency increased due to cross-couple between d and q axes. Unfortunately, when motor speed is very high, operating at field weakening the effect of cross-couple is obvious. What is more, the delay of digital SVPWM control system, as shown in Fig. 49.4, enhances the influence of the cross-couple. The delay of digital SVPWM control system can be modeled in synchronous reference frame as follows:

$$D(z) = z^{-1} \cdot e^{(j\omega T_s)^{-1}} \tag{49.9}$$

In our system, surface mounting permanent magnet synchronous motor (SPMSM) is chosen, which the d–q components of stator self-inductance are equal, as follows:

$$L_d = L_q = L_s \tag{49.10}$$

So discrete-time motor model can be described in complex form, as follows [11]:

$$G_P(z) = \frac{I_s(z)}{V_s(z)} = \frac{1 - e^{-(R_s/L_s)T_s}}{R_s\left(ze^{j\omega(k)T_s} - e^{-(R_s/L_s)T_s}\right)} \tag{49.11}$$

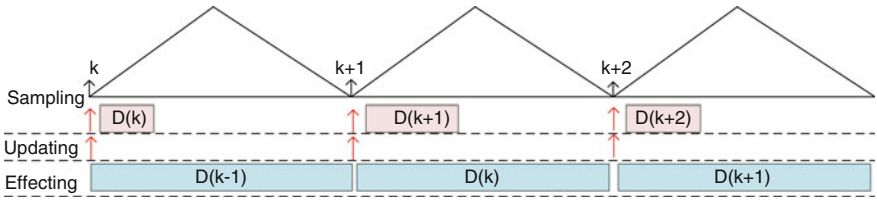


Fig. 49.4 Time sequence of current sampling calculation and PWM update

According to pole-zero point's compensation method and direct design PI control methods, the complex PI current controller can be written as follows:

$$G_c(z) = \frac{k_p\left(ze^{j\omega_s T_s} - e^{-k_i T_s}\right)}{z-1} \tag{49.12}$$

And then the compensation of digital control can be described as follows:

$$G_d(z) = z \cdot e^{(j\omega T_s)} \tag{49.13}$$

Ignoring the switching harmonics is appropriate in this situation since the ripple current created by these harmonics is eliminated from the feedback signal by the current sampling method and filtering, the closed-loop transfer function of control system is

$$G_{cls}(z) = \frac{I(z)}{I^*(z)} = \frac{G_c(z)G_d(z)G_P(z)D(z)}{1 + G_c(z)G_P(z)G_d(z)D(z)} \tag{49.14}$$

The performance characteristics of this discrete-time complex PI current regulator were analytically investigated using eigenvalue migration as shown in Fig. 49.5. Excellent pole-zero cancellation for the pole migrating from A to B is achieved, allowing the control performance to be decoupled from the speed.

49.5 Experiment Result

Figure 49.6 shows the experimental setup to verify the effectiveness of the proposed method. Table 49.1 gives the parameters of the tested PMSM. This PMSM is coupled to Magnetic powder dynamometer, and the PMSM is in the torque control mode.

Figure 49.7 shows experimentally measured maximum torque versus machine speed of the proposed and the conventional method

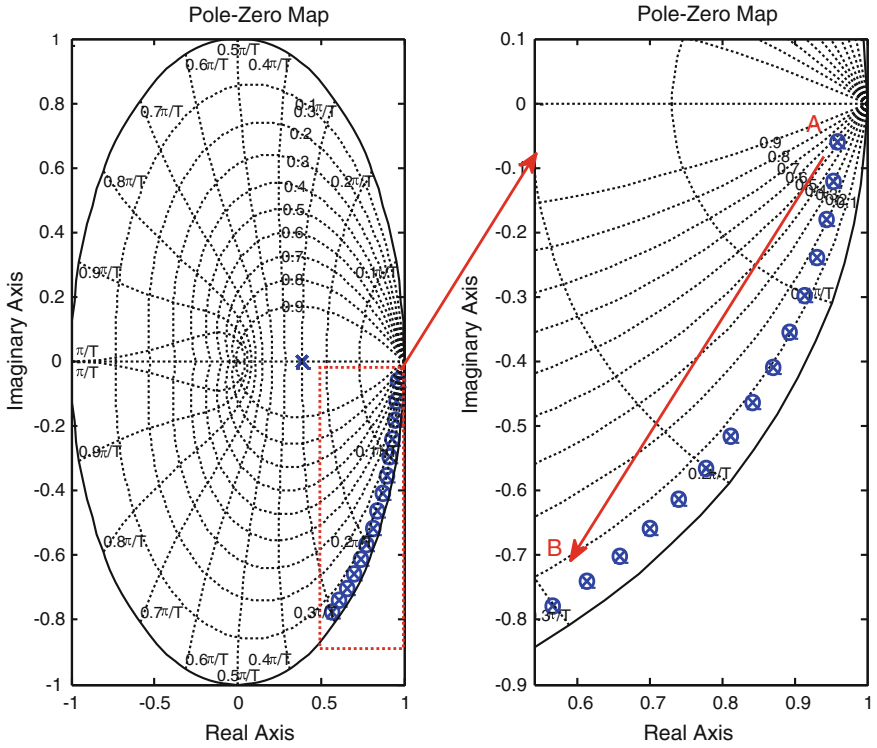


Fig. 49.5 The pole-zero map of close loop



Fig. 49.6 Experimental setup

Table 49.1 Parameter of PMSM

Variable name	Value	Variable name	Value
Poles	8	Rated speed	3000 rpm
Rated current	29.5 A	Back EMF constant	122.4 V/krpm
Inductance	3.2 mH	Resistance	0.22Ω

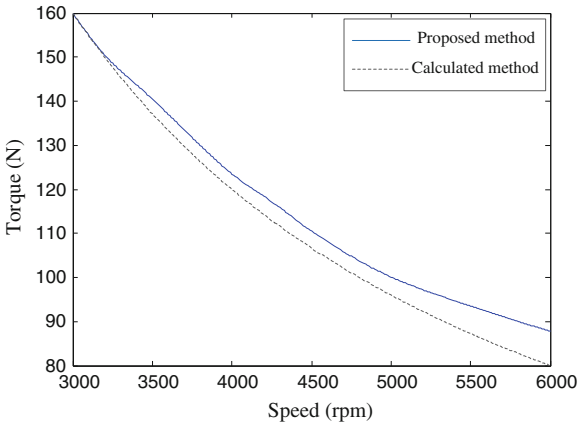


Fig. 49.7 Maximum torque versus speed

49.6 Conclusion

In this paper, a novel flux-weakening control method of an PMSM has been proposed. With DC-link voltage is sampling in real-time and output-phase voltage reconfigured by inverter of switch function, voltage regulator can generate appropriate current command, and discrete-time complex current regulator, with discrete-time complex PI allowing the control performance to be decoupled from the speed, and make performance of current regulator very high. The effectiveness and superiority of the proposed method have been verified by the experimental results.

References

1. Soong WL, Miller TJE (1993) Theoretical limitations to the field-weakening performance of the five classes of brushless synchronous AC motor drive. In: Sixth international conference on electrical machines and drives, (Conf. Publ. No. 376) pp 127–132 IET
2. Soong WL, Miller TJE (1994) Field-weakening performance of brushless synchronous AC motor drives. IEEE Proc Electr Power Appl 141(6):331–340
3. Zhang Y, Xu L, Guven MK, Chi S, Illindala M (2011) Experimental verification of deep field weakening operation of a 50-kW IPM machine by using single current regulator. IEEE Trans Ind Appl 47(1):128–133

4. Morimoto S, Takeda Y, Hirasa T, Taniguchi K (1990) Expansion of operating limits for permanent magnet motor by current vector control considering inverter capacity. IEEE Trans Ind Appl 26(5):866–871
5. Pan CT, Sue SM (2005) A linear maximum torque per ampere control for IPMSM drives over full-speed range. IEEE Trans Energy Convers 20(2):359–366
6. Lenke RU, De Doncker RW, Kwak MS, Kwon TS, Sul SK (2006) Field weakening control of interior permanent magnet machine using improved current interpolation technique. In: 37th IEEE conference on power electronics specialists, 2006, PESC'06. pp 1–5
7. Yifa S, Shouyi Y, Weihua G et al (2010) Field weakening operation control strategies of permanent magnet synchronous motor for railway vehicles. Proc CSEE 30(9):74–79 (in Chinese)
8. del Blanco FB, Degner MW, Lorenz RD (1999) Dynamic analysis of current regulators for AC motors using complex vectors. IEEE Trans Ind Appl 35(6):1424–1432
9. Holtz J, Quan J, Schmittt G, Pontt J, Rodriguez J, Newman P, Miranda H (2003) Design of fast and robust current regulators for high power drives based on complex state variables. In: 38th IAS annual meeting IEEE conference on industry applications 2003. Conference record of the Vol. 3, pp 1997–2004
10. Morimoto S, Takeda Y, Hirasa T, Taniguchi K (1990) Expansion of operating limits for permanent magnet motor by current vector control considering inverter capacity. IEEE Trans on Ind Appl 26(5):866–871
11. Kim H, Degner MW, Guerrero JM, Briz F, Lorenz RD (2010) Discrete-time current regulator design for AC machine drives. IEEE Trans Ind Appl 46(4):1425–1435

Part IV
Rail Transportation and Automatic Control

Chapter 50
Research on the Balise Uplink Signal Process Method Based on the Noise Feature Extraction and Adaptive Noise Cancellation

Yanning Zhang and Huibing Zhao

Abstract In this paper, the Balise Uplink signal noise feature and noise process method are researched. The model of the transmission channel for Uplink signal is established, and the AR power spectrum model is used to extract the noise feature. Then an adaptive noise cancellation is proposed based on the noise feature extraction. This method improves the SNR of the Uplink signal, so that Balise Uplink signal transmission is more reliable.

Keywords Balise uplink signal · Noise feature extraction · Adaptive noise cancellation

50.1 Introduction

Balise, which is based on the principle of electromagnetic coupling, is a large-capacity equipment used for wayside-onboard information transmission. As a subsystem of the CTCS, Balise is used to provide position, grade conversion, wireless communication, and other information to CTCS-3 Train Control System, while providing line speed, line slope, track circuits, temporary speed restriction line parameter information to CTCS-2 Train Control System [1]. Because of its low cost, high performance, and other advantages, Balise technology is successfully and widely used in our high-speed railway and urban rail traffic control system.

However, the environment of railway scene is complicated. The problem of Balise Data loss would happen in information transmission because of the influence of track current noise, electromagnetic interference noise, and excitation

Y. Zhang (✉) · H. Zhao
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Room 811, Si Yuan Building, No.3 Shang Yuan Cun, Beijing 100044 Haidian, China
e-mail: ynzhang0528@163.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_50,
© Springer-Verlag Berlin Heidelberg 2014

inrush current's radiation coupled noise [2]. This phenomenon had occurred in the Wuhan-Guangzhou PDL. Under the high-speed condition, how to eliminate the potential safety hazard of Balise Data loss has become a problem to be solved in Balise application and development. Currently, there are no researches about noise characteristics and noise processing of Balise Uplink signal. Therefore, relevant research is necessary.

50.2 Balise Uplink Channel Modeling and the Noise Feature Extraction

50.2.1 Balise Uplink Channel Modeling

The Balise Transmission System consists of the (wayside) Balise and the On-board Transmission Equipment [3]. When the train passes over the Balise, coupling coil induces a 27 MHz magnetic field then converts it to electricity, making the Balise circuit work. Balise control module modulates telegram information to FSK signal, which is transmitted cyclically by the coupling coil. The On-board Antenna receives FSK signal and transmits it to On-board Equipment [4]. The center frequency of Uplink signal is 4.234 MHz, The frequency deviation is 282.24 kHz and the mean data rate is 564480 bit/s.

According to the Test Specification for Eurobalise FFFIS [5], interface "A" is used for transmitting telegrams between the Balise and the On-board Equipment, as shown in Fig. 50.1. Interface "A" contains sub-interfaces "A1" and "A4." Interface "A1" is used for transmitting Balise telegrams from the Uplink Balise to the Antenna Unit. Interface "A4" is used for transmitting the required power (Tele-powering) from the Antenna Unit to the Balise.

Figure 50.2 shows Uplink signal transmission channel model. The Uplink signal transmit terminal is Balise, the receiving terminal is Antenna Unit. Transmission channel is equivalent to an amplifier. The amplification factor M is related to the magnetic coupling coefficient. Transmission channel also includes the input of environmental noise. From Fig. 50.2 we can get Eq. (50.1).

$$y(k) = Mx(k)+n(k) \tag{50.1}$$

where $x(k)$ is the FSK signal transmitted from the Balise, $n(k)$ is a noise signal, $y(k)$ is a signal received from Antenna Unit. When the train does not pass over the Balise, the On-board Equipment receives noise signal.

Figure 50.3 shows the relationship between Balise horizontal position and Tele-powering signal transmitted from Balise, Tele-powering signal received from the On-board Equipment [6]. From Fig. 50.3 we can see when the horizontal position is equal to 0 (x = 0), Tele-powering signal received from the On-board Equipment and the peak of Uplink signal are both the largest. Considering the BTM threshold,

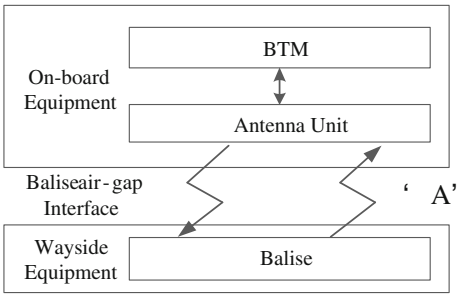


Fig. 50.1 Interface ‘A’ of balise

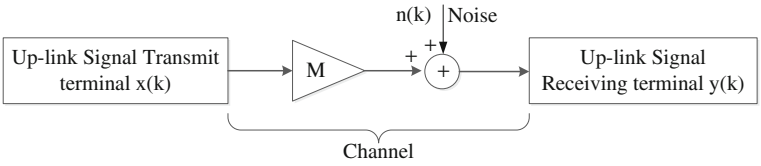


Fig. 50.2 Uplink signal transmission channel model

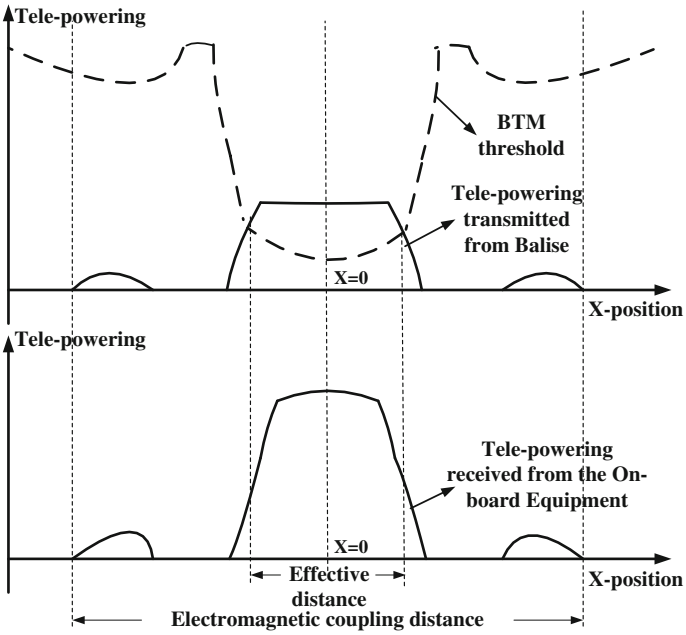


Fig. 50.3 Relationship between balise horizontal position and tele-powering signal transmitted from balise, tele-powering signal received from the on-board equipment

the reception of energy must be greater than a certain value before the BTM starting demodulation and decoding. The length of X which corresponds to BTM demodulation and decoding is called the effective distance.

50.2.2 Uplink Noise Signal Feature Extraction

Uplink Noise Signal can be regarded as stationary random signal and it can be estimated by the modern power spectrum, whose basic principle is that the input is white noise, the output is a signal to be studied, then use the known signal to calculate model parameters and use the model parameters to estimate the power spectrum of the signal. Modern power spectral estimation is based on the parameters model of the stochastic process, generally using the AR model, whose transfer function is described in Eq. (50.2).

$$H(z) = \frac{\sqrt{e}}{A(z)} = \frac{\sqrt{e}}{1 + a_2 z^{-1} + \ldots + a_{(p+1)} z^{-p}} \tag{50.2}$$

where p is the order of the polynomial, e is amplitude variance.

Figure 50.4 shows the power spectrum estimation model of Uplink noise signal.

AR model is used to estimate the Uplink noise signal. According to the AR model parameters we can calculate the power spectrum feature of Uplink noise signal. If we use the parameters of Uplink noise sample as the parameters of AR model, and use the standard Gaussian noise as the input, the Uplink noise signal can be simulated.

Figure 50.5 shows a noise signal sample received from the On-board Equipment when passing over the Balise and its power spectrum. We set the AR model order 32 and use the Burg algorithm to calculate model parameters.

50.3 Uplink Signal Adaptive Noise Cancellation

50.3.1 Basic Principle of Adaptive Noise Cancellation

Adaptive noise cancellation is an extension of adaptive filtering. It can extract useful signals from the noise signals and inhibit or attenuate noise interference [7]. Figure 50.6 is a block diagram of the adaptive noise cancellation. Input 1 is a superimposed signal of noise signal and desired signal, input 2 $x(k)$, $e(k)$ is the estimation error. The filter coefficients of the adaptive filter are controlled by the error signal $e(k)$. It can be adjusted automatically according to $e(k)$ and the adaptive algorithm. Error signal can be expressed by the Eq. (50.3).

$$e(k) = s(k) + n(k) - x(k) \tag{50.3}$$

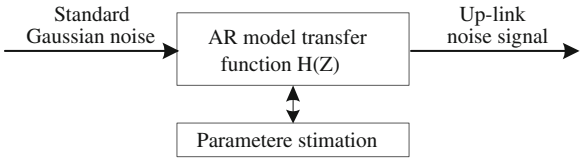


Fig. 50.4 Uplink noise signal power spectrum estimation model

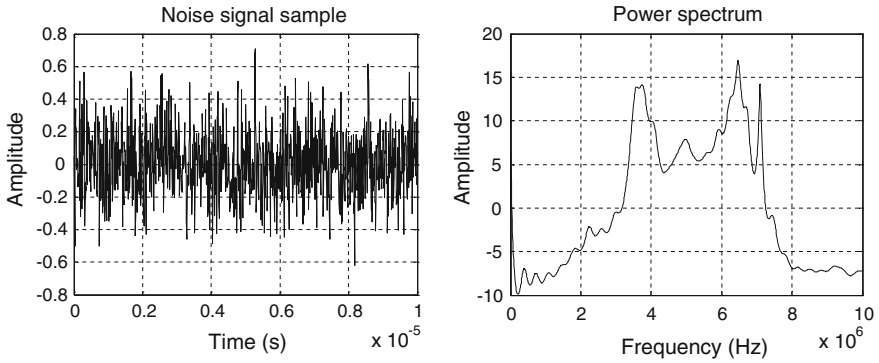
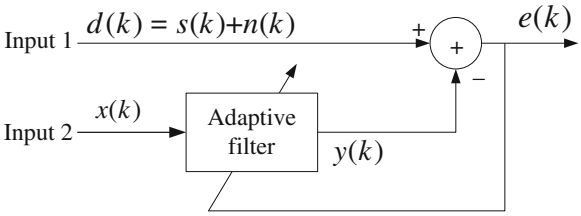


Fig. 50.5 Noise signal sample and its power spectrum. *Left* Noise signal sample. *Right* power spectrum of the noise signal sample

Fig. 50.6 Adaptive noise cancellation principle



Squaring both sides of Eq. (50.3) and taking the expectations, we can get Eq. (50.4).

$$E\{e^2(k)\} = E\{s^2(k)\} + E\{[m(k) - y(k)]^2\} - 2E\{s(k)[m(k) - y(k)]\} \quad (50.4)$$

Assuming that $s(k)$ has nothing to do with noise signal $n(k)$ and $x(k)$, the last item of Eq. (50.4) is 0. And because $s(k)$ and filter coefficients are independent with each other, when the adaptive algorithm converges, $E\{[n(k) - y(k)]^2\}$ takes the minimum value. This indicates that the error signal $e(k)$ is equal to the estimate of desired signal $s(k)$, so that it can achieve noise cancellation.

Adaptive algorithm can adopt NLMS algorithm, whose principle is simple and convergent rate is fast. The iterative formulas are given in Eqs. (50.5, 50.6, and 50.7).

$$y(k) = w^T(k)x(k) \tag{50.5}$$

$$e(k) = d(k) - y(k) \tag{50.6}$$

$$w(k+1) = w(k) + \frac{\alpha}{\|x(k)\|^2 + \beta} e(k)x(k) \tag{50.7}$$

where $w(k)$ is the weight vector at time k. α controls convergent rate of the algorithm, $0 < \alpha < 2$. β is a small positive number. When the input signal is small, it is used to solve the numerical instability problem that caused by the division.

50.3.2 Realization of the Uplink Signal Adaptive Noise Cancellation

Figure 50.6 shows that the adaptive noise cancellation method requires two inputs, one is a useful signal containing noise, and the other is a noise signal. For the Balise system, the On-board Equipment receives Uplink signal with noise when passing over the Balise, otherwise it only receives noise signal. The whole process has only one input. Therefore, it is need to modify the method of traditional adaptive noise cancellation.

When a train passes over a Balise or before, the environment of the Balise system is almost the same, so the noise distribution can be considered identical. The signal that the On-board Equipment received when passing over the Balise is regarded as input 1 of the adaptive noise cancellation. When the train does not pass over the Balise, we select a sample of the noise to estimate AR model parameter. It is regarded as input 2. Entire process of Uplink signal adaptive noise cancellation is shown in Fig. 50.7.

In addition, the parameters of adaptive noise cancellation include the AR model order$n$, the filter order N. To ensure the performance of noise cancellation, it is need to analyze on-site data to determine these parameters.

50.4 Simulation

Figure 50.8 shows a simulated noise signal produced by the parameters estimation and the power spectrum of the noise sample. We can see the noise sample contains single-frequency interference. Figure 50.9 shows an Uplink signal with noise and the output signal through adaptive noise cancellation. The filter order is 8, $\alpha = 0.1$, $\beta = 5$.

Sampling rate is 564480×10^3 bit/s.

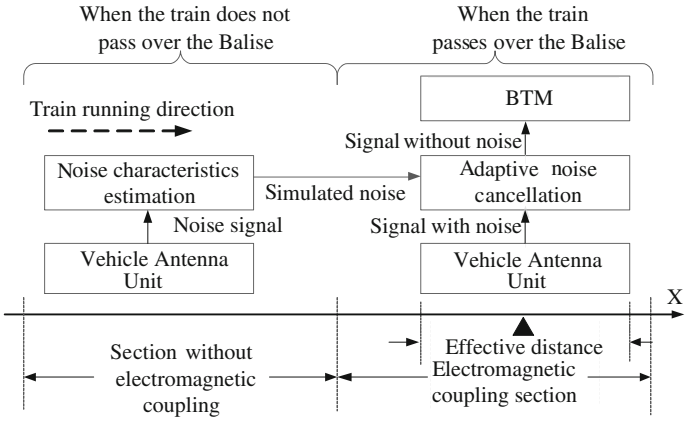


Fig. 50.7 Uplink signal adaptive noise cancellation process

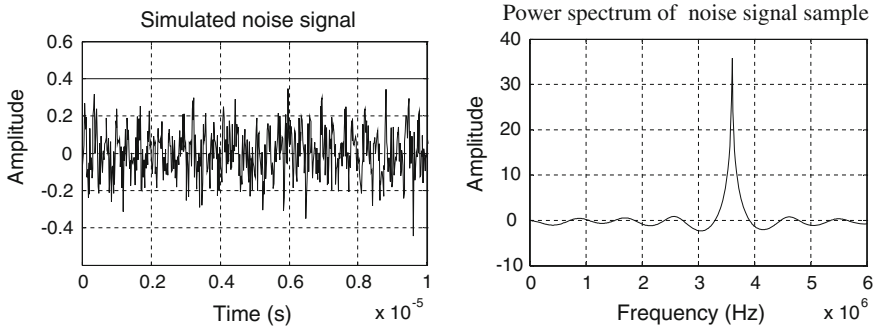


Fig. 50.8 Simulated noise signal and the power spectrum. *Left* Simulated noise signal, *Right* power spectrum of the noise signal sample

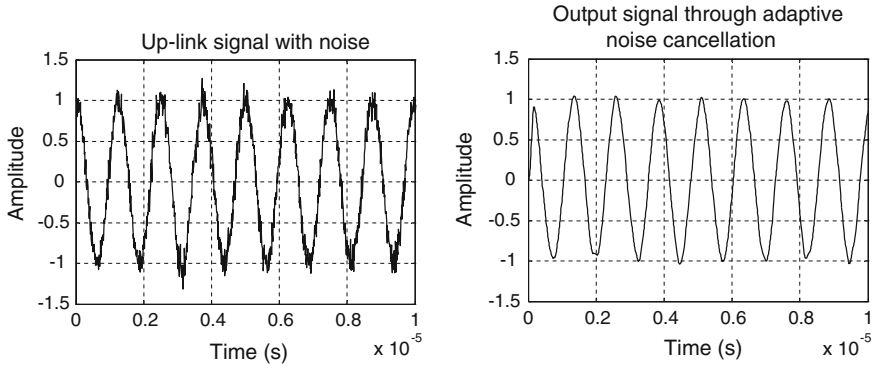


Fig. 50.9 Uplink signal with noise and the output signal. L*eft* uplink signal with noise. *Right* output signal through adaptive noise cancellation

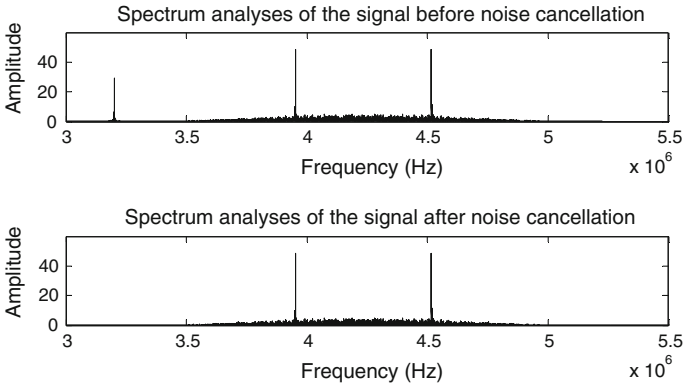


Fig. 50.10 Spectrum analyses of the signals before and after noise cancellation. *Top* spectrum analyses of the signal before noise cancellation. *Bottom* spectrum analyses of the signal after noise cancellation

The spectrums of signals in Fig. 50.9 are shown in Fig. 50.10. We can see the interference component has been eliminated effectively.

50.5 Conclusion

This paper proposes the process method of the Balise Uplink signal based on the noise feature extraction and adaptive noise cancellation. The Uplink transmission channel model is also established. The AR power spectrum estimation model is applied to extract the noise feature. It can analyze the noise characteristics and simulate the occurrence of noise through the parameters estimation. The simulation results show that the method can improve the quality of Balise Uplink signal transmission.

Acknowledgments This work was financially supported by the High-tech R&D project of The Ministry of Railways of the People's Republic of China (Grant No. 2011 × 021-c). The authors acknowledge this support.

References

1. (2008) CTCS-3 Train control systems overall technical program Ministry of Railways of the people's Republic of China. Science and Technology No. 34. China railway publishing house, Beijing, (in Chinese)
2. Dawei C (2011) research on performance of balise equipment under condition of high—speed, Jiaotong University, Beijing (in Chinese)

3. (2003) ERTMS/ETCS-Class1: Test specification for eurobalise FFFIS Ref: SUBSET-036, V2.2.1
4. Guo Rui, Li Yin, Wang Yongchao, Liu Jiuwen (2010) Design of BTM receiver module in Intermittent Balise of CTCS. Electron Measur Tech 33(7):19–22 (in Chinese)
5. (2003) ERTMS/ETCS-Class1: Test specification for eurobalise FFFIS Ref: SUBSET-085, V2.1.2
6. Houjin C, Xue J, Hu J (2008) Digital signal processing. Higher Education Press Beijing (in Chinese)
7. Binfang C (2007) Research on adaptive noise cancellation technology research, Hunan University Changsha (in Chinese)

Chapter 51
Research on the Relation Between the Wind Direction of Vehicle Air Condition and Cooling Efficiency of the Cab

Bin feng Guo, Hong chao Wang and Yan ling Wang

Abstract Applying the computer-aided optimization (CAO) platform Isight to integrate computational fluid dynamics software FLUENT with Design of Experiment (DOE) components. The DOE components automatically call the values of fluid simulation, and treat the heat flux of front and back seats as the cooling efficiency of the cab. Fluent adopts the Pressure-based Couple Solver (PBCS) algorithm and switches the RNG $\kappa-\varepsilon$ two equations model as the three-dimensional turbulence model of the cab. Through DOE it can analyze how the direction of the air-condition inlet impact the heat flux of the front and back seats significantly.

Keywords Design of experiment · Isight · Fluent · Direction of inlet air · Cooling efficiency

51.1 Introduction

The cab of the vehicle parked in the sunshine in summer will be heated up to 50 °C in a short time, to increase the comfort when driving in summer, it is essential for the air-condition system to decrease the temperature of the cab as soon as possible [1]. At present, the research on the temperature field of the cab at home and abroad is mainly utilizing the Computational Fluid Dynamics (CFD) technology analyses and make conclusions by comparing a few different working conditions. However, the design factors of the working conditions divided by manual work only include several levels, and the factor levels distribute so unequally in the design space, thus it is unable to reflect reasonably the relationship between the input factor and its response.

B. f. Guo · H. c. Wang (✉) · Y. l. Wang
Department of Mechatronic Engineering, Zhengzhou Huaxin College, Zhonghua North Road, Xinzheng, Henan, China
e-mail: whcjordan123@163.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_51,
© Springer-Verlag Berlin Heidelberg 2014

Isight possesses a powerful Computer-Aided Optimization (CAO) functionality and extensive CAD/CAE accesses so that the users can quickly establish a complex process of Design of Experiment (DOE), setup or modify the design variables and automatically carry out the multi-cycles analysis [2]. This paper, based on Isight, achieves the connection and data transfer between Fluent and DOE, and combines with relevant theories to analyze the response relationship between the direction of air-condition inlet and the heat flux of seats in the cab.

51.2 Computational Fluid Process

51.2.1 Geometrical Model

The inner structure of the cab is so complex that it will greatly increase the computational cost if directly establish the cab model, and the air flow in the cab belongs to the low-speed incompressible flow, in addition, some components in the cab have little effect on the air flow. Thus, this paper need to simplify the geometrical model as follows:

1. Simplify the geometrical shape of solid walls and seats in the cab;
2. Ignore the steering wheel and the instrument panel;
3. Ignore the effect of defrosting in the process [3];
4. Treat the cab as a symmetry to decrease the computational cost.

The simplified cab model is shown in Fig. 51.1, the cold air flow is passed into the cab from the front four inlets and passed out from the outlets below the control panel.

51.2.2 Meshing

This paper applies the software ICEM to mesh the geometrical cab model, due to the structured mesh owning lots of merits such as rapidly generating, high quality, and simple date structure, etc. [4]. By continuously balancing the computational cost and the mesh quality, the number of final generated mesh is 117668.

51.2.3 Simulation Process

This paper mainly focuses on the process when the temperature drops in the cab, thus has no need to consider the air flow between the cab and outside, assumes the cab is a completely sealed space. Because of treating the air flow as a incompressible unsteady flow, it belongs to the turbulence, thus it adopts the RNG $\kappa-\varepsilon$ two equations model to simulate the air flow. As the process when the temperature

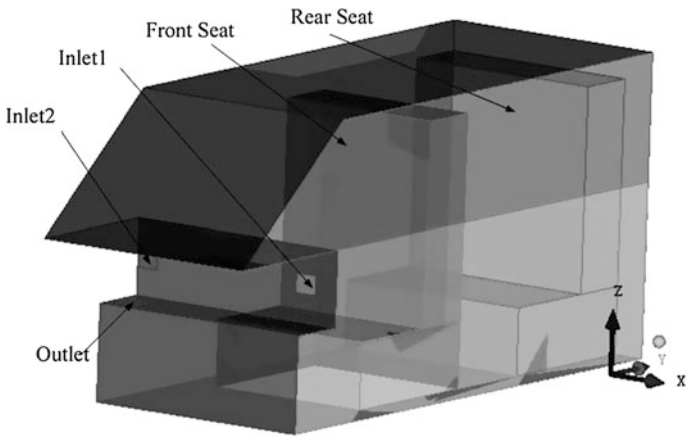


Fig. 51.1 Geometrical model of the cab

in the cab drops is so fast that it can ignore the radiation heat transfer, mainly focus on the heat convection between the cold air and solid walls [5]. The relevant parameter settings are as follows:

1. Active Fluent Start Journal command to start recording a new script and name this script file as input.jou, this step is to create a command-flow file for integrating Fluent into Isight;
2. Adopt the unsteady Pressure-Based Couple Solver (PBCS) [6];
3. Adopt the RNG $\kappa-\varepsilon$ two equations model;
4. Setup the boundary conditions, for the four inlets, it chooses the velocity inlet condition, setup its magnitude as 2 m/s, its direction is represented in the form of Cartesian Coordinate, setup the inlet temperature as 16 °C, the outlet chooses outflow condition and setup the symmetry plane;
5. Setup the initial temperature as 53 °C, the time step as 0.05 s, the number of steps as 200, the maximum iterations as 50. After the computation converged, it obtains the heat fluxes of front and rear seats which are seemed as the response of DOE.
6. After the computation, active Start Transcript, select Report/Fluxes/seat, then close Transcript and save as output.dat.
7. Close Fluent.

51.3 Softwares Integration

The Simcode component in Isight is a access that achieves the integration of the input, process, and output of program, it mainly contains three parts: (1) Rewrite the input file; (2) Carry out the program; (3) Read the output file [7]. Fluent is

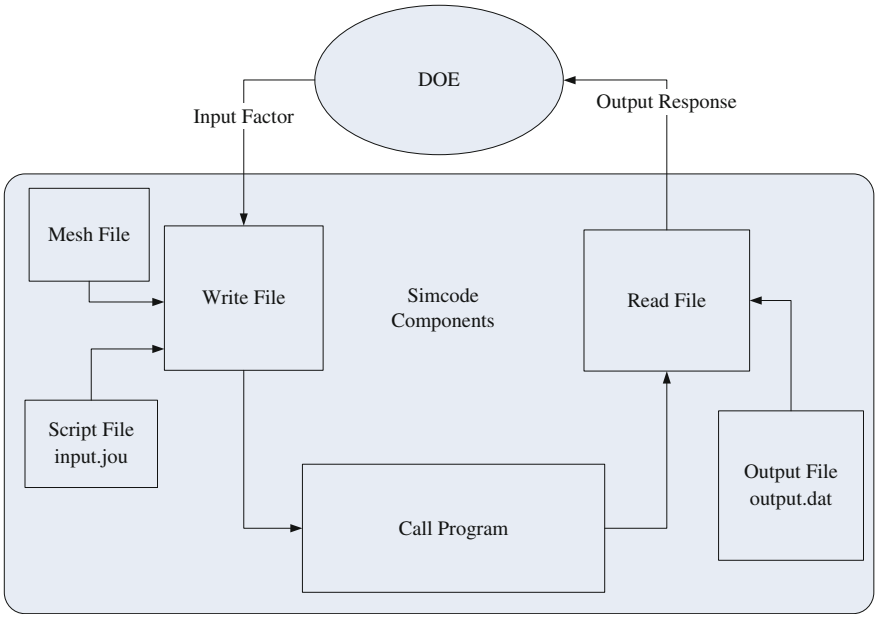


Fig. 51.2 Flow chart of simulation

integrated with Isight by means of Simcode component. Based on Simcode, compile the batch processing file(fluent.bat) and call Fluent, the input files are the mesh file and the script file, the DOE factors are the projections of inlets on the Cartesian coordinate, thus there are six factors in sum. The air inlet close to the door is defined as Inlet1, the components in each direction are represented as Inlet1-x, Inlet1-y, and Inlet1-z. The air inlet below the control plane is defined as Inlet-2, corresponding components are represented as Inlet2-x, Inlet2-y, and Inlet2-z, in this process, the responses of DOE are the heat fluxes of front and rear seats, the heat flux of front seat is defined as Output1, the other is defined as Output2. The concrete flow of simulation is shown in Fig. 51.2.

51.4 DOE Process

51.4.1 Selection for the DOE Algorithm

This paper adopts the Optimal Latin Hypercube Design (OLHD), the principle of which is to divide evenly each dimensionality into m-parts in a n-dimensional space, each space can be represented as $[x_k^{i-1}, x_k^i]$, $i \in [1, m]$. Randomly select m-points, make sure that each level of the factor is analyzed only once, thus it forms a m-dimensional space OLHD including m samples. OLHD improves the uniformity

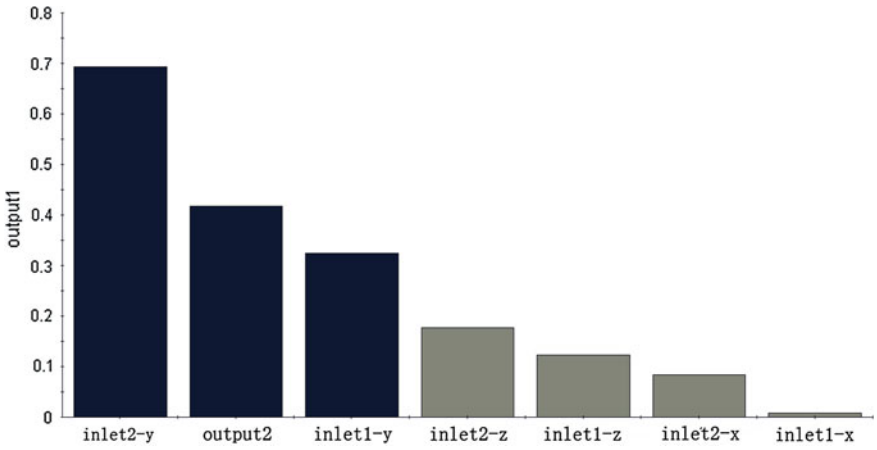


Fig. 51.3 Correlation graph of output1

of RLHD (Random Latin Hypercube Design) and the fitting precise between the factor and its response. Setup the level of each factor as 100, then generate the design matrix by OLHD [8].

51.4.2 DOE Results Analysis

51.4.2.1 Correlation Analysis

The correlation graph can obviously show the correlation between each parameter and the output variable as shown in Figs. 51.3 and 51.4, the computational formula of correlation coefficient is shown as below:

$$r = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sqrt{\sum (X - \overline{X})^2}\sqrt{\sum (Y - \overline{Y})^2}} = \frac{S_{XY}}{S_X S_Y} \tag{51.1}$$

In this formula: $S_{XY} = \sum (X - \overline{X})(Y - \overline{Y})/(n-1)$; the sample standard deviation of X is $S_X = \sqrt{\sum (X - \overline{X})^2/(n-1)}$; the sample standard deviation of Y is $S_Y = \sqrt{\sum (Y - \overline{Y})^2/(n-1)}$.

Firstly, if ***r*** is between one and zero it belongs to positive correlation which means the two variables are in direct proportion, and it is represented as dark bars as is shown in Fig. 51.3, on the contrary, the light color bars represent the negative correlation if ***r*** is between minus one and zero. The more the absolute value of ***r*** is close to one, the stronger the degree of correlation is, the more the absolute value of ***r*** is close to zero, the weaker the degree of correlation is.

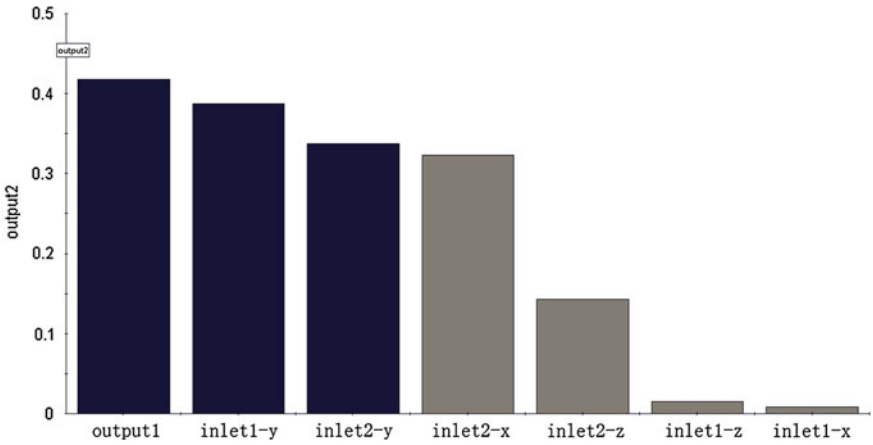


Fig. 51.4 Correlation graph of output2

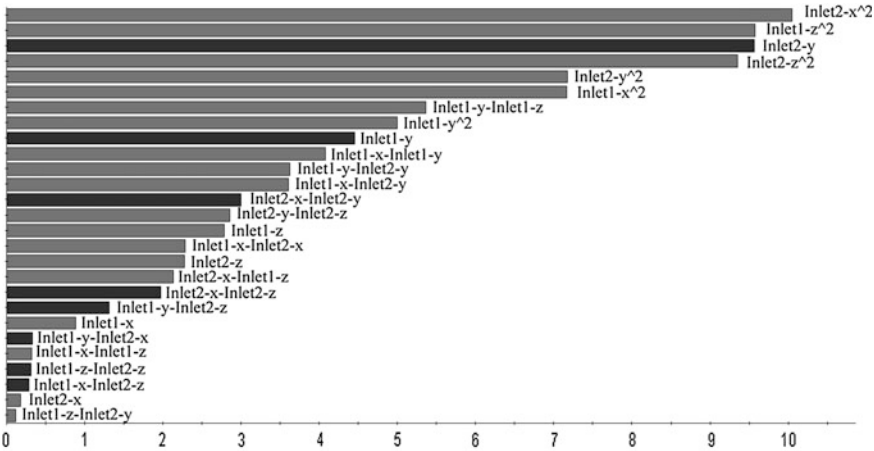


Fig. 51.5 Pareto graph of output1

By analyzing the correlation graphs, it clearly indicates that Output1 presents a positive correlation with Inlet1-y and Inlet2-y, what is more the former positive correlation is weaker than the later, and presents a negative correlation with the other factors; Similarly, Output2 presents a positive correlation with Inlet1-y and Inlet2-y, the former positive correlation is stronger than the later, and presents a negative correlation with the other factors; It also shows that the correlation between Output1 and Inlet1-x is nearly close to zero, the same as the correlations between Output2 and Inlet1-x, Inlet-z.

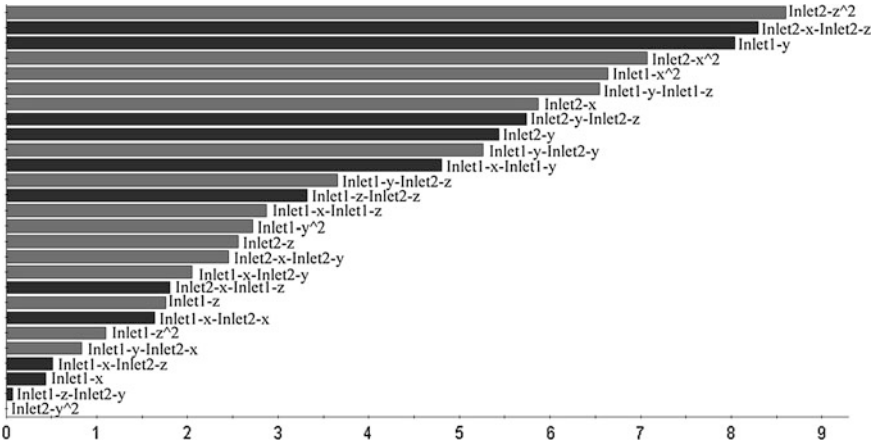


Fig. 51.6 Pareto graph of output2

51.4.2.2 Contribution Rate Analysis

To analyze the relation between the cross-terms of factors with their responses, it need to make use of the Pareto graphs which reflect the contribution rate of all the factors in the model to each response after the samples fitted, as shown in Figs. 51.5 and 51.6, the dark bars represent the positive effect, the light color bars represent the negative effect. From the two graphs, the cross-terms Inlet2-x-Inlet2-y and Inlet2-x-Inlet2-z present the positive effect to Output1, and the cross-terms of Inlet1 mainly present the negative effect, similarly, the cross-terms Inlet2-x-Inlet2-z, Inlet2-y-Inlet2-z, and Inlet2-y-Inlet1-y present the positive effect to Output2, and the cross-terms of Inlet1 mainly present the negative effect.

51.5 Conclusion

By DOE it can focus on the key factors and exclude the disturbances, thus simplify the problem and decrease the workload, based on Isight, integrating all the necessary components, it is able to carry out the repeated work automatically so that the computational cost can be greatly reduced. This paper analyzes the correlation and contribution rate among the six components of air inlet and corresponding cross-terms with the heat flux of front and rear seats, it indicates that Inlet2 has the most impact on the cooling rate of front and rear seats, for Inlet1it should be modified to decrease the negative effect on the heat flux of front and rear seats, when Inlet1-x and Inlet1-z are equal to zero, the negative effect can be reduced to the minimum.

References

1. Wang F (2011) Computational fluid dynamics analysis, vol 1. Tsinghua University Press, Beijing, pp 34–39 (in Chinese)
2. Lai Y (2012) Parameter optimization theory and examples explanation about isight, vol 10. Beijing University of Aeronautics and Astronautics Press, Beijing, pp 101–111 (in Chinese)
3. Xiao H (2011) Numerical simulation of the cab flow field considering the body heat. Comput Simul 6:121–126 (in Chinese)
4. Ji B, Chen J (2012) Meshing technology and examples explanation, vol 6. China WaterPower Press, Beijing, pp 160–168 (in Chinese)
5. Wang D (2011) Analysis on the temperature field in cab based on the thermal balance model, vol 5. Hunan University, Changsha, pp 33–39 (in Chinese)
6. Mo Z, Yang Y, Deng F (2008) CFD numerical simulation on the arrangement of vehicle air-condition. Build Energy Environ 2:102–106 (in Chinese)
7. Zhang J, Chen J (2008) Research on the three-dimensional flow and heat transfer of vehicle air-condition. Tianjin Cars 2:21–35 (in Chinese)
8. Liu J, Chen J (2003) Numerical analysis on the air flow distribution of carriage. J Shanghai Jiaotong Univ 37(7):36–41 (in Chinese)

Chapter 52
Vocational Skills Comprehensive Evaluation Method of Track Maintenance Workers

Dongyao Jia and Shengxiong Zou

Abstract The traditional personnel quality evaluation currently are subjective or objective evaluation method. However, objective evaluation methods for investigating dynamic indicators are mainly restricted to equipment maintenance capabilities, and static evaluation methods cannot represent a subjective evaluation of the long-term changes in personal qualities of the track maintenance workers. This paper proposes a comprehensive evaluation method for vocational skills of maintenance personnel based on the static and dynamic evaluation results, mechanic skills evaluation can be divided into comprehensive quality evaluation and the level of job evaluation, the ratio of static and dynamic evaluation results is proposed based on the two-stage comprehensive evaluation method and evaluation procedure, so the comprehensive evaluation method is more diverse and flexible. Experimental results show that this method can overcome the static or dynamic unity of evaluation result, the accuracy of the evaluation method is effectively improved for considering the subjective and objective factors, which makes the evaluation results more reasonable.

Keywords Comprehensive evaluation · Static and dynamic evaluation · Two-stage method

52.1 Introduction

Comprehensive evaluation refers to the evaluation of the object system for multiattribute, multisystem description, so as to reach the integrity and overall evaluation. There are many traditional evaluation methods such as AHP, Delphi

D. Jia · S. Zou (✉)
Advanced Control Systems Laboratory, Beijing Jiaotong University, No.3 Shang Yuan Cun District, Beijing, China
e-mail: bjtu1989@126.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_52,
© Springer-Verlag Berlin Heidelberg 2014

method, fuzzy comprehensive evaluation method, and the matter-element analysis method. Samanta and Sarkar [1] use analytic hierarchy process (AHP) to select the mine maintenance personnel based on the different attributes, But the data statistics is large and difficult to determine the weights when exiting too many indicators. Gupta and Sachdeva [2] use fuzzy Delphi method to select the best logistics provider based on Fuzzy TOPSIS protocol, However, the subjective factors are relatively strong and the lack of accurate quantitative analysis. Sun and Wang [3] used to establish a human factors evaluation index system and evaluation model based on Matter Element Model of matter element analysis method. However, the index values and the correlation calculation of this method based on matter-element analysis theory are much more complex. The proposed evaluation method is currently widely used in the management class comprehensive quality evaluation of employees, and for rail transport mechanic vocational skills assessment involves very little.

This paper based on the static and dynamic evaluation results of the track maintenance workers to evaluate the comprehensive vocational skills, using two stages of evaluation methods and evaluation procedures and importing static, dynamic results importance ratio λ which reflects the different emphasis degree, simulation results of the evaluation of the method is consistent with the analytic hierarchy process, However, the evaluation system, evaluation methods, and evaluation results of this method have a great advantage of the diversity, flexibility of results analysis, making the comprehensive evaluation method more diverse and flexible.

52.2 Comprehensive Evaluation Principles

In this article, the professional quality evaluation data is divided into two major categories of static and dynamic, the static data include basic personal situation, professional quality, skill level, etc.; dynamic data include maintenance efficiency, maintenance satisfaction, and equipment stability after repairing [4].

Track maintenance workers vocational skills evaluation system is the collection of static and dynamic evaluation system and the evaluation factors don't have the correlation, as the dynamic evaluation system characteristics, vocational skills maintenance system is a hybrid system of the tree structure which is shown below:

Static evaluation investigates the recently overall quality level of the track maintenance workers, but it can't represent a long-term change in track repairman personal qualities and the evaluation results are stationary and idealized [5]. Meanwhile, the results of the indicators are mostly relying on experts and repairman scoring to determine. The evaluation results are still subjective though introducing the concept of entropy to determine the integrated weights. Therefore, the evaluation results can't fully represent the overall level of track maintenance workers vocational skills Fig. 52.1.

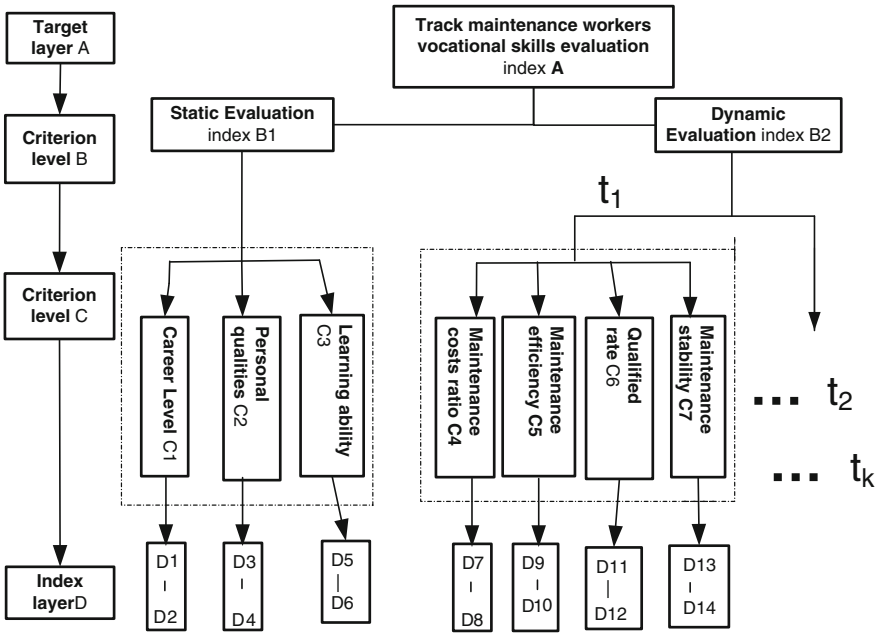


Fig. 52.1 The mixed-tree structure of the occupation skill evaluation

Dynamic evaluation investigates the maintenance skill levels of track maintenance workers, the evaluation results are dynamic and real-time, this method indicates the equipment maintenance level comprehensively in period of time [6]. Meanwhile, the evaluation data all based on maintenance work orders which makes the evaluation results objective. However, the examined index mainly confined to equipment maintenance capabilities that can't fully reflect the overall quality of vocational skills.

Comprehensive evaluation method combines static evaluation and dynamic evaluation, it takes both the static and dynamic factors into account that makes the evaluation results more comprehensive; meanwhile, considering the influence of subjective and objective factors which makes the evaluation results more reasonable, so the comprehensive evaluation method is scientific and practical [7].

52.3 Comprehensive Evaluation Method

This paper uses two-stage evaluation methods based on the static and dynamic evaluation results, two-stage comprehensive evaluation method comprises the following two steps:

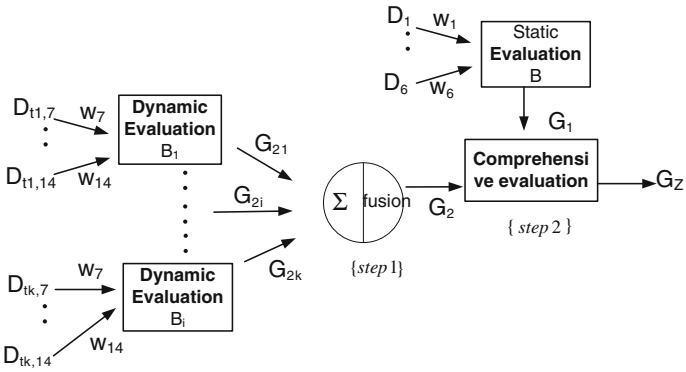


Fig. 52.2 The method of the second phase of comprehensive evaluation

Step one: G_{2i} represents dynamic evaluation value of vocational skills at time t_k which merged into dynamic comprehensive evaluation value G_2.

Step two: The final comprehensive evaluation value G_Z is combined static evaluation value G_1 and dynamic comprehensive evaluation value G_2.

In Fig. 52.2, $D_1 \sim D_6$ represent the observed index value of static evaluation, $w_1 \sim w_6$ represent corresponding weights of each index of Static evaluation.

The Static evaluation value G_1 as follows:

$$G_1 = f_1(D_1 \cdots D_6, w_1 \cdots w_6) \tag{52.1}$$

$D_{tk,7} \sim D_{tk,14}$ represent the observed index value of dynamic evaluation at time t_k, $w_7 \sim w_{14}$ represent corresponding weights of each index of dynamic evaluation.

Dynamic evaluation value G_{2k} at time t_k as follows:

$$G_{2k} = f_2(D_{tk,7} \ldots D_{tk,14}, w_7 \ldots w_{14}) \tag{52.2}$$

Dynamic comprehensive evaluation value G_2 is merged by dynamic evaluation value G_{2i} at time t_k.

$$G_2 = \sum w_{2i} G_{2k} \tag{52.3}$$

The final comprehensive evaluation value G_Z is combined static evaluation value G_1 and dynamic comprehensive evaluation value G_2.

$$G_Z = w_1 G + w_2 G_2 = w_1 G + w_2 \sum_{i=1}^{k} w_{2i} G_{2i} \tag{52.4}$$

The method is simple and the evaluation process is clearer. The results of dynamic comprehensive evaluation will fully reflect the characteristics of dynamic evaluation. The final evaluation result G_2 is combined G_1 and G_2, comprehensive

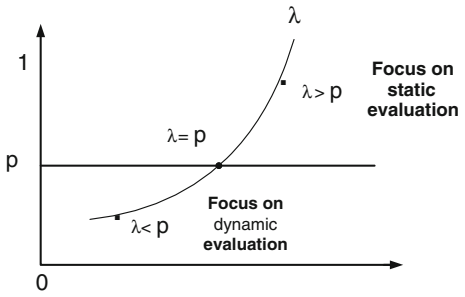


Fig. 52.3 The relationship of λ and p

evaluation results can reflect the different emphasis degree of static and dynamic evaluation by adjusting the weighting coefficients.

Track repairman comprehensive evaluation value G_z is determined according to the obtained static evaluation value G_1 and dynamic evaluation value G_2 [8].

$$G_z = \theta_1 G_1 + \theta_2 G_2, (\theta_1, \theta_2 > 0, \theta_1 + \theta_2 = 1) \tag{52.5}$$

θ_1, θ_2 reflect the importance of the static and dynamic evaluation value, so they also reflect the purpose of the evaluators.

We define static and dynamic results importance ratio $\lambda = \theta_1/\theta_2$, the static, dynamic evaluation fixed ratio p is given [9]. Meanwhile, p on behalf of the degree of importance under static and dynamic evaluation results. In general, the professional maintenance capability of comprehensive evaluation is more relatively important and the large dynamic evaluation data, so $p < 1$.

The different results of comprehensive evaluation can be achieved by comparing λ and p, the obtained results can provide a wide range of proposals Fig. 52.3.

When $\lambda > p$, the static evaluation factors in the evaluation focus more strongly than the dynamic evaluation in the comprehensive skills evaluation, which means that the evaluators more value the comprehensive quality of track maintenance workers in the comprehensive evaluation. In this case, the evaluation results can provide the basis for the following decisions: the unit of moral education model and assessment of outstanding communist party member, etc.

When $\lambda < p$, the dynamic evaluation factors in the evaluation focus more strongly than the static evaluation in the comprehensive skills evaluation, which means that the evaluators more value the equipment maintenance ability of track maintenance workers in the comprehensive evaluation, in this case the evaluation results can provide the basis for the following decisions: business backbone selection, mechanic performance appraisal, etc.

When $\lambda = p$, which means the evaluators comprehensively consider the comprehensive quality of track maintenance and professional maintenance ability level.

The value λ is a debugging process in evaluation process. Assuming a special case, the comprehensive sorting is consistent with dynamic evaluation results,

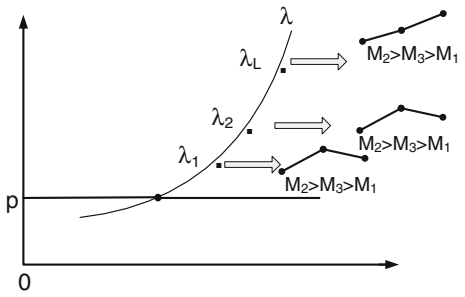


Fig. 52.4 The schematic of λ selected and sort of maintenance worker

which means the evaluation results more fully reflect the impact of dynamic evaluation, but the results are inconsistent with the static evaluation. When the evaluator requires to focus on static evaluation impact, we choose $\lambda > p$, but the sort of evaluation results did not change with the increasing of λ. so λneed to find a "tipping point" λ_L, which make comprehensive evaluation results more reflect the influence of static evaluation Fig. 52.4.

52.4 The Evaluation Result Analysis

Experiment is carried in Wuhan Metro Station Qiaokou, six evaluation experts including Wuhan metro chief engineer, maintenance workshop director, and maintenance foreman were selected to evaluate six track maintenance workers through questionnaires and interviews, etc.

52.4.1 Comprehensive Evaluation Method

When $\lambda = p = 0.67$, comprehensive evaluation results and the static, dynamic evaluation results comparison chart were shown as follows:

In Fig. 52.5, the comprehensive evaluation results $M4 > M1 > M2 > M3$, Among which " $>$ " Means much better. The comprehensive evaluation results with static and dynamic evaluation results of track maintenance workers *M4* are the same high scores, which means the overall quality and operational maintenance capability of repairman *M4* in the maintenance is the best. Noticing track repairman *M2*, *M3* Evaluation Orders, $M2 > M3$ in Static Evaluation, $M2 < M3$ in dynamic Evaluation, $M2 < M3$ in comprehensive evaluation, but the gap is not big, that means when $\lambda = p = 0.67$, comprehensive evaluation results have showed the influence of dynamic evaluation result for the static evaluation results, the gap of comprehensive evaluation of *M2* and *M3* is narrowing. If evaluators continue to emphasize the dynamic assessment of impact on the consolidated results of the evaluation, $\lambda > p$ by increasing the value of λ, the sorts of comprehensive

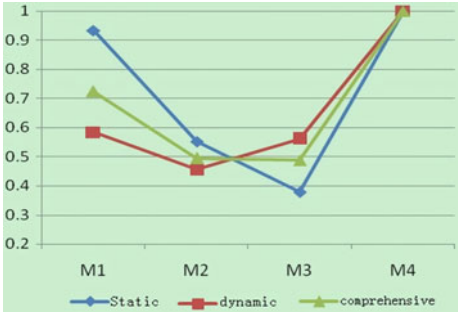


Fig. 52.5 The contrast of evaluation results

evaluation results of M2, M3 will be reflected to be consistent with dynamic evaluation results.

52.4.2 Analytic Hierarchy Process

AHP evaluation index architecture uses the same structure of comprehensive evaluation, and scores are used to indicate all evaluation levels, evaluation scores of the individual index are given in Table 52.1:

The comprehensive evaluation of track maintenance workers using linear comprehensive evaluation model is as follows:

$$I = \sum_{i=1}^{n} W_i \times Y_i \tag{52.6}$$

In formula (52.6), I represents comprehensive evaluation value of vocational skills, which Indicates the level of track maintenance workers vocational skills, W_i is the i-index related weight, Y_i is the i-index score.

According to formula (52.6), the results of AHP evaluation results are as follows:

$$G_z = [81.194 \quad 76.206 \quad 74.1795 \quad 90.0960 \quad 68.8015 \quad 70.5360]$$

$$M4 > M1 > M2 > M3 > M6 > M5 \tag{52.7}$$

The comprehensive evaluation results are compared with the traditional AHP, the results are given in Table 52.2. Experiments show that the proposed track repairman vocational skills evaluation results are consistent with AHP, and this method can effectively reflect the track repairman actual level of vocational skills. At the same time, comprehensive evaluation method has a great advantage of evaluation system, evaluation methods, and evaluation results diversity and evaluation analysis flexibility compared to traditional evaluation methods, and it can be more comprehensive, objective, and flexible to evaluate track maintenance professional skill and has great practical significance.

Table 52.1 The score table of index

	M1	*M2*	*M3*	*M4*	*M5*	*M6*	W_i
Career level	81.4	75.75	73.55	91.4	68.3	68.5	0.2
Personal qualities	80.3	78.9	77.55	92.4	68.25	71.6	0.15
Learning ability	83.9	76.9	74.9	91	66	71.5	0.15
Maintenance costs rate	80.5	76.5	75.6	89.5	74.2	65.3	0.12
Maintenance efficiency	82.2	74	73	87.4	67.3	74	0.12
Maintenance qualified rate	79	74.3	71.6	90.2	70	72.3	0.15
Maintenance stability	81	77.1	73	86.8	68.4	71	0.11

Table 52.2 The contrast of comprehensive evaluation method and AHP

	Evaluation architecture	Evaluation method	Evaluation results	Analysis
Comprehensive evaluation method	Mixed tree	Static Dynamic comprehensive	Static Dynamic comprehensive	Multiple
AHP	Tree	Comprehensive	Comprehensive	One kind

52.5 Conclusion

The evaluation method of track maintenance comprehensive vocational skills based on static and dynamic evaluation results is put forward in this paper. First, mechanic skills evaluation can be divided into the overall quality appraisal and the job evaluation level. Second, the importance of static and dynamic results ratio λ is introduced, which reflects the different focused degree of evaluators for static, dynamic evaluation results. Finally, by using two evaluation methods of comprehensive evaluation method and AHP to evaluate the level of vocational skills, evaluation results show that two kinds of evaluation methods are basically the same, but comprehensive evaluation method has a great advantage of evaluation system, evaluation methods and evaluation results diversity and evaluation analysis flexibility compared to AHP, track maintenance workers vocational skills evaluation results are in line with the mechanic themselves, which confirmed the comprehensive evaluation method is scientific and effective.

References

1. Samanta B, Sarkar B, Mukherjee SK et al (2004) An application of AHP for selection of maintenance personnel in mine. J Mines Metals Fuels 52:420–424
2. Gupta R, Sachdeva A, Bhardwaj A (2012) A framework for selection of logistics outsourcing partner in uncertain environment using TOPSIS. Int J Ind Syst Eng 12(2):223–242
3. Sun BJ, Wang L (2008) Application of material-units analysis method to the human resources allocation in enterprises. Autom Instrum 23(7):1–3 (in Chinese)

4. Xu ZS, Zhang L (2003) An overview of operators for aggregating information. Int J Intell Syst 18(9):953–969 (in Chinese)
5. Darema F (2005) Grid computing and beyond: the context of dynamic data driven applications systems. Proc IEEE 93(3):692–697
6. Damarla TR, Pham T, Lake D (2004) An algorithm for classifying multiple targets using acoustic signature. In: Proceedings of SPIE signal processing, sensor fusion and target recognition, Vol 5429. pp 421–427
7. John H (1975) A adaptation in natural and artificial system. University of Michigan Press, Cambridge
8. Bian X, Tian HP, Guo YJ (2004) Suppliers with stimulating dynamic evaluation method. Nankai Bus Rev 7(5):87–90 (in Chinese)
9. Yi PT, Zhang N, Guo YJ (2009) Study on dimensionless methods in dynamic comprehensive evaluation. J Northeast Univ 30(6):889–92 (in Chinese)

Chapter 53
Middleware-Based Distributed CTCS-3 Simulation Platform

Lianbao Yang, Tianhua Xu and Zhenxian Wang

Abstract In recent years, with the rapid development of the high-speed railway, the safety and efficient operations are increasingly higher requirements. Railway accidents are caused by many factors, such as human factors, technical factors, natural factors, and sometimes are the result of several factors working together. This paper aims to build a scalable CTCS-3 simulation platform in the open environment for fault analysis and safety verification, using JMS and RMI heterogeneous interconnect middleware. Finally, given in the mudslides, rain, snow, and other natural environmental factors, with the application of Wulongquandong Station-Xianningbei Station actual line data, we simulate three trains. The results indicate that the platform can easily extend and obtain the safe operation of trains.

Keywords CTCS-3 · Middleware · Distributed · Simulation-platform

53.1 Introduction

Distributed simulation technology [1] because of its similarity with the actual structure of the system as well as the simulation fidelity and reliability has been widely used in practical engineering. For the analysis of parallel discrete event systems (PDES), distributed simulation has more irreplaceable advantages. CTCS-3 [2] is a typical parallel discrete event system. Distributed simulation during the design should ensure that all distributed information exchange between computers in real-time and with high accuracy. Middleware technology [3, 4] can play a shielding effect of different pervasive computing devices and provide a unified interface specification for different applications.

L. Yang (✉) · T. Xu · Z. Wang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 12120288@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_53,
© Springer-Verlag Berlin Heidelberg 2014

JMS (Java Messaging Service) is a Java platform for MOM (Message Oriented Middleware) technical specifications. JMS facilitates the messaging system of java applications for message exchange, and by providing a standard produce, send, receive messages interface simplifies enterprise application development. RMI (Remote Method Invocation) [5, 6] provides the transfer of information between objects simultaneously, with high reliability and real-time characteristics.

In this paper, CTCS-3 simulation platform [7, 8] is developed by Java with the application of two technologies, JMS and RMI heterogeneous interconnect. JMS and RMI can be used in heterogeneous, distributed environment for data transmission. JMS technology focused on delivering reliable, asynchronous messaging, and support for loosely coupled; while RMI technology is for tightly coupled systems to achieve synchronization method between java objects intermodulation. According to different application purposes, we have selectively used two techniques to make the system function module component-based design, based interface calls, allowing the module to maintain high internal cohesion between modules remain loosely coupled, which built a scalable simulation platform.

53.2 Design of CTCS-3 Simulation Platform

53.2.1 The Design Framework of CTCS-3 Simulation Platform

Through the analysis of the CTCS-3, we design a framework of CTCS-3 simulation platform, as shown in Fig. 53.1.

As shown in Fig. 53.1, CTCS-3 Simulation Platform contains line data processing module, the simulation data preprocessing module, and train simulation module. Each main module can be subdivided into different modules. Among them, the line data processing storage module's main function is to load the simulation need to use the line data file and the data in an Excel table corresponding processing storage; simulation data preprocessing module is mainly to preprocess the data sent to other modules so that the simulation module can be invoked on demand; train simulation module as the core module of the system is responsible for all trains simulated to generate simulation data. Now a brief introduction of each module:

(1) Line data processing module: This module reads actual high-speed rail line data stored in Excels through JAVA under the appropriate paths.
(2) RMI/JMS communication module: This module provides the use of JAVA RMI distributed development API, and provides the interface to interact with other subsystems. The system receives information and sends information through this module.
(3) Train parameter processing module: This module will initial train interface settings parameters and store them in local memory for the simulation calls.

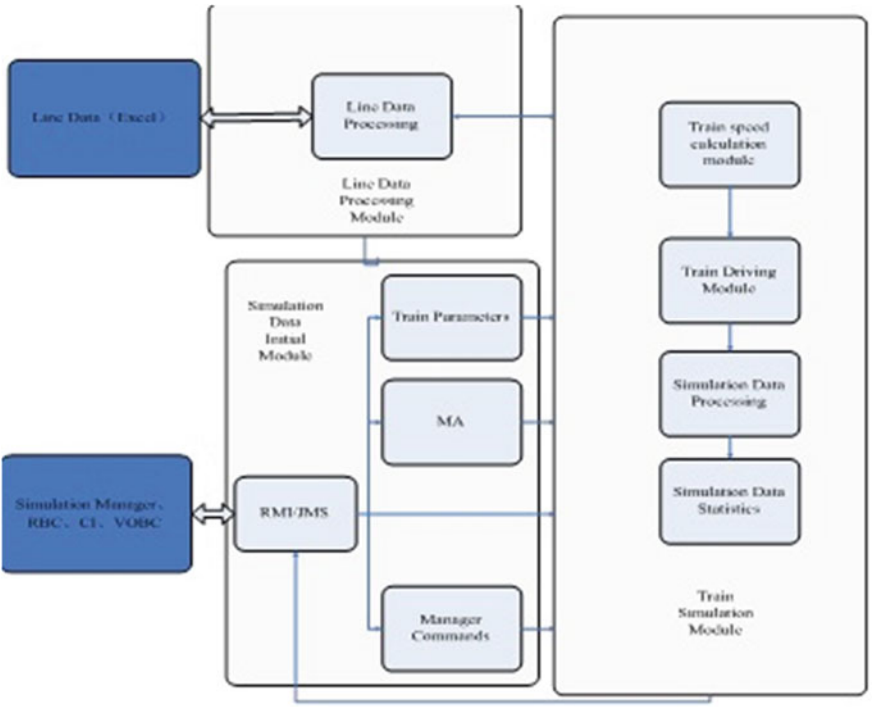


Fig. 53.1 The framework of CTCS-3 simulation platform

Meanwhile, for each cycle the module can provide different trains with different parameters.

(4) MA information processing module: This module will receive MA transmitted from RBC and store it in local memory for the simulation. Meanwhile, the module is capable of performing simple checks to ensure the correct use of simulation MA, combined local data with MA content. In addition, the module is also responsible for matching the different trains' MA for each cycle.
(5) Manager Information Processing Module: This module generates different control commands, to ensure operational arrangements in accordance with the simulation. In addition, the module is also responsible for matching different train operators control information for each cycle.
(6) Train speed calculation module: This module will provide real-time calculation of speed for different train, according to received MA, in line with local database information for each cycle.
(7) Train driving simulation module: This module will control train speed in efficient operation within the allowable range, according to the train speed.
(8) Simulation data calculation module: This module based on the train driving simulator information, calculates train running status, such as train speed, train position updates for each cycle.

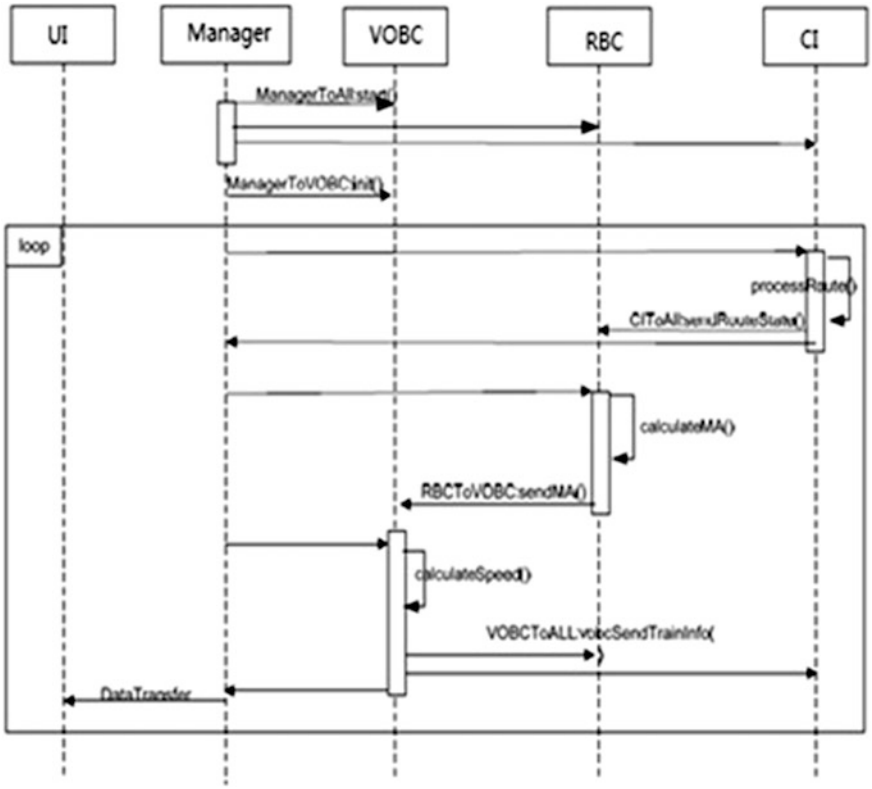


Fig. 53.2 CTCS-3 subsystem timing diagram

(9) Simulation data statistics module: This module is based on simulation results and get some statistical information, such as the occupancy of each section, route of occupancy, train running total mileage, and other information.
(10) Data recording module: This module for each cycle of the simulation data recording, saves the data in the database. This file can be used as emulator log file for later viewing and can also be used as a way to find the cause of failure during the simulation process.

The simulation function is performed by a distributed computing system structure, including user interfaces (UI), simulation Test Manager (Manager), computer interlock (CI) subsystem, RBC Center (RBC), subsystem and vehicle (VOBC), subfive module systems. The five modules simulate the actual function of the corresponding system, including calculation, evaluation, and validation of the data needed to complete the simulation needs. The CTCS-3 subsystem timing diagram [9] is shown in Fig. 53.2.

Various functional modules together constitute a distributed system where the train control functions between modules using RMI (Remote Method Invocation)

for remote communication, so that the distributed system control function of the columns between modules to facilitate data communications and procedure calls effectively improve distributed system efficiency. Train control functions to achieve synchronization between the various modules require a higher degree of coupling is large, and not suitable for use in the JMS J2EE framework technology for communication. Platform simulation interface display module and submodule synchronization between other less demanding, and visualization of the simulation process requires a lot of data transfer support, in order to improve the stability and flexibility of a communication system that will achieve the related control functions and columns degree small display interface module and Simulation test Manager modules communicate using JMS technology, so as to minimize the degree of coupling systems while improving system stability and flexibility.

53.3 CTCS-3 Simulation Platform Design Goals

Middleware-based Distributed CTCS-3 Simulation Platform, from the high-speed rail transportation system integrated perspective, ensures safety, improves transport efficiency point of view, a combination of design, simulation, fault tolerant algorithms to evaluate and verify the function to ensure train control system in the event of a fault tolerant state for safe and reliable operation. CTCS-3 Simulation platform will be a process that involves automatic train control, security systems and fault-tolerant algorithms, and other professional integrated design and verification evaluation. Based on integrated and unified engineering high-speed rail line data (Excel table form), through the simulation of a variety of dynamic performance, the platform provides users with a real safe and efficient train control system simulation platform when a fault occurs in a fault-tolerant state operating conditions.

CTCS-3 Simulation Platform building has three overall objectives as shown in Fig. 53.3.

53.4 The Analysis of Simulation Results

ATP (Automatic Train Protection) is the basis and key to ensure the safety equipment mainly in charge of the train operation interval control and overspeed protection task, ensuring driving safety (Fig. 53.4).

Through simulation affected by natural disasters (rain, snow, landslide, debris flow, etc.), we design a module to recalculate the MA and control train running speed, to verify the flexibility, reliability, and security of the platform (Fig. 53.5).

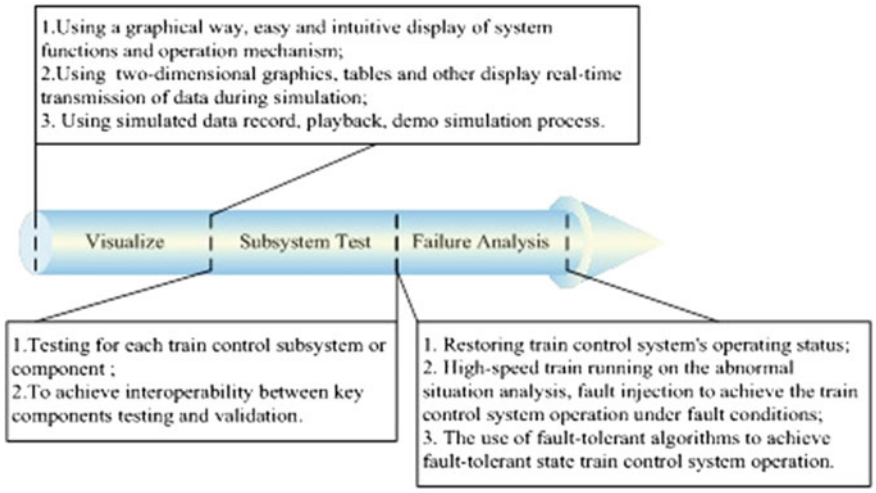


Fig. 53.3 CTCS-3 simulation platform design goals

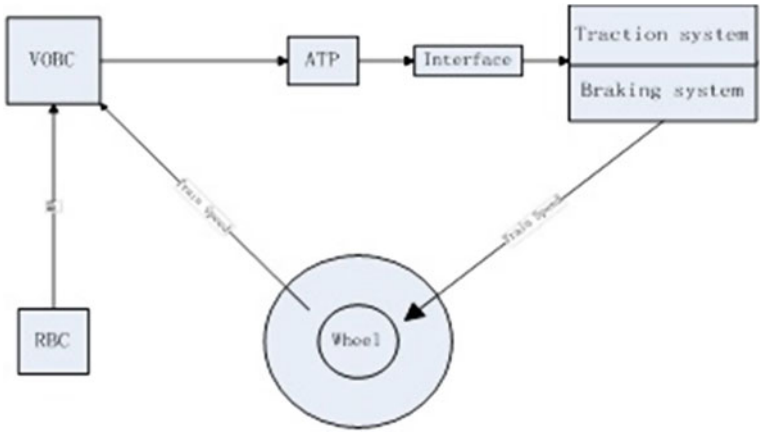


Fig. 53.4 ATP in train running speed control

The simulation results are as follows:

From Figs. 53.6 and 53.7 we can see that the train speed during operation in certain rail segments will be reduced to zero, or to very low speed. This shows that trains have received a front rail segment of natural disaster, and according to the different natural disasters, has made the corresponding decision control and achieved the purpose of safe operation.

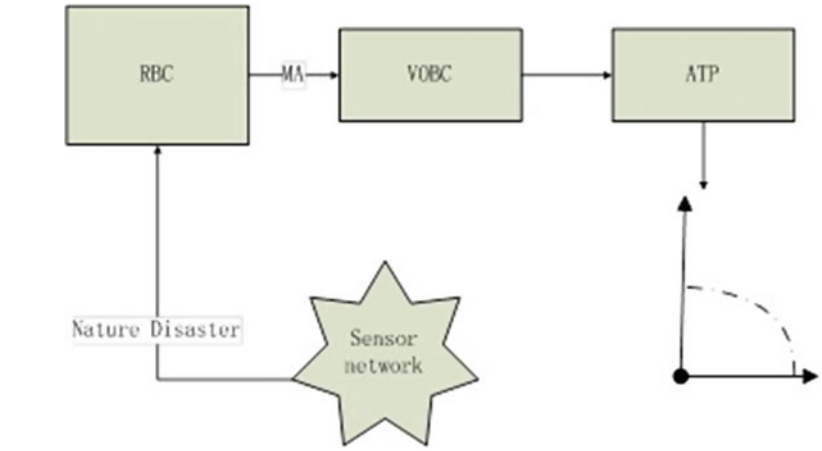


Fig. 53.5 ATP speed protection under the influence of natural disasters

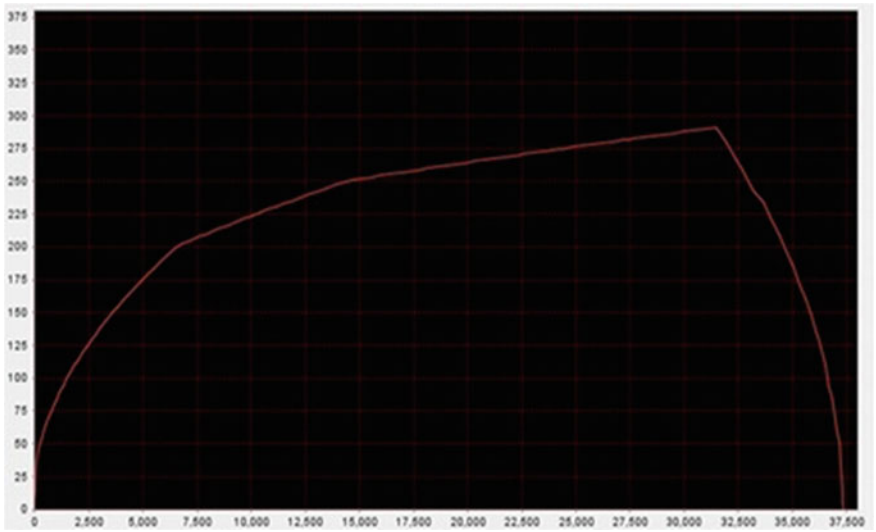


Fig. 53.6 The target–distance curve of the train running normal

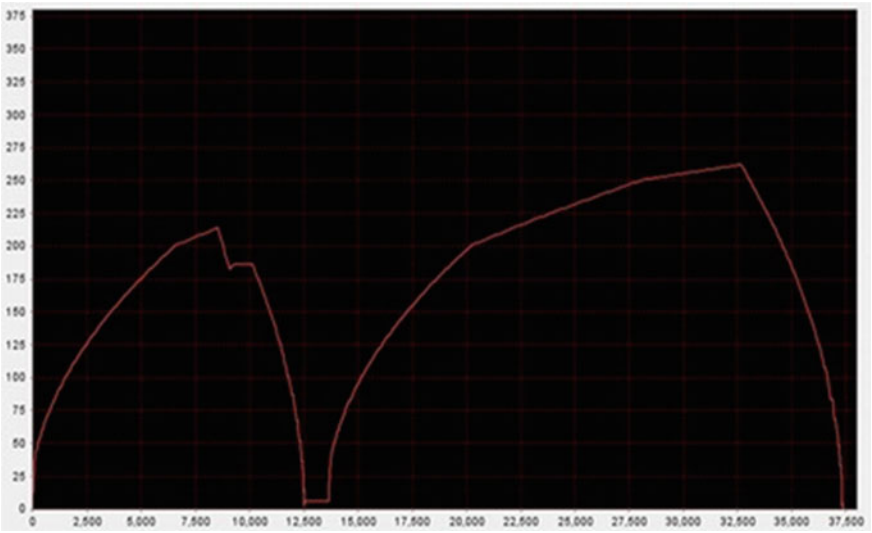


Fig. 53.7 The target–distance curve of the train affected by natural disasters

53.5 Conclusion

In this paper, we design a scalable CTCS-3 simulation platform in the open environment for fault analysis and safety verification, using JMS and RMI heterogeneous interconnect middleware. The platform will play a very important role in the safe and reliable operation and help researchers to perform fault analysis and diagnosis.

Acknowledgments This work is supported by the National 863 plan projects (No. 2011AA010104), International Science & Technology Cooperation Program of China under Grant No. 2012DFG81600, the State Key laboratory of Rail Traffic Control and Safety of Beijing Jiaotong University within the frame of the project (No. RCS2012ZT005).

References

1. Liu C, Du Y-h (2006). The distributed simulation of train group operation model based on object—oriented petri subnet. J China Railway Sci 27(1):109–113 (in Chinese)
2. Technology Innvovation General Scheme of CTCS-3[S]. Beijing: Ministry of Railways Transport Bureau, 2008.(in Chinese)
3. Dumitraş T, Narasimhan P (2013). A study of unpredictability in fault-tolerant middleware. Comput Netw 57(3):682–698
4. Al-Jaroodi J, Jawhar I, Al-Dhaheri A, Al-Abdouli F, Mohamed N (2010) Security middleware approaches and issues for ubiquitous applications. Comput Math Appl 60(2):187–197
5. Alt M, Gorlatch S (2005). Adapting Java RMI for grid computing. Future Gener Comput Syst 21(5):699–707

6. Thiruvathukal GK, Thomas LS, Korczynski AT (1998) Reflective remote method invocation concurrency: practice and experience 10(11–13):911–925
7. Xu W, Xin YW, Lu GZ (2007). Research and development of pervasive computing middleware technology. J Comput Sci 34(11):1–5 (in Chinese)
8. Cai B, Shangguan W, Li X, Wang J (2010). Research on supporting technology for simulation CTCS-3 based on multi-resoulution modeling. J Beijing Jiaotong Univ 34(2):5–10 (in Chinese)
9. Berkenkötter K (2008) Reliable UML models and profiles. Electron Notes Theoret Comput Sci 217:203–220

Chapter 54
Study on Express-Slow Routes Mode of Foreign Subway

Senhui Li, Xingchen Zhang, Junhua Chen and Jixuan Liu

Abstract By taking the routes of express–slow train mode in New York and Tokyo as main research object, express line only and express–slow routes are selected as two different express–slow train schemes. Making a comparison about applicable characteristics between two express–slow train design schemes, considering features of infrastructure of urban rail transit in our country, and the demand of passenger flow, as well as learns from related construction projects abroad and operating experience. Based on this, inspiration of operation plan of express–slow train in the regional urban rail line of our country is drawn.

Keywords Urban rail transit · Express–slow train · Transfer and joint · Crossing stations setting · Features of passenger flow

54.1 Introduction

With rapid development of urban rail transit, many long lines which operate over 50 km come up in urban rail transit network, connecting the central city and suburban areas. For this kind of more than 50 km route, frequency of train stops will lower the speed and prolong the passenger's traveling time. In the long term, the efficiency of line operation and the attractiveness to the passenger flow will reduce gradually. In order to reduce the adverse impact on the efficiency of line operation and the attractiveness to the passenger flow which brought by the stretching of the line and the frequency of train stop, some foreign cities adopt the

S. Li (✉) · X. Zhang · J. Chen · J. Liu
School of Traffic and Transportation, Beijing Jiaotong University, Haidian Distrit, Jiaoda East Street, 100044 Beijing, China
e-mail: 12120878@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_54,
© Springer-Verlag Berlin Heidelberg 2014

express–slow train mode, and it achieved good application effect. This paper will use the advanced experience of New York and Tokyo subway for reference and put forward some suggestions for the construction of Beijing subway.

54.2 Overview

54.2.1 Conception of the Express–Slow Train Mode

The express–slow train operation mode refers to a combined scheme which two or more kinds of speed kind trains working on the same route. The slow train stops at all stations, but not the express. The express improves the running speed by reducing the stop [1]. It can give consideration both to the commuting passengers' demand for shorten the travel time in long distance and the short distance passengers' demand for daily travel.

54.2.2 Running Mode of the Express–Slow Train

According to whether the train running on the same track or not, the divided highway lines operation can be divided into two different kinds, one is designed to have a special line for express, the other is designed that the slow train and the express run on the same track.

The method of setting up a special line for the express make the slow train and the express run on their own track, and it can be divided into single special line(three tacks mode) and double special line(four tracks mode). The express can overtake the slow train at any position without extra overtaking lines setup [2]. This way is the main pattern of the subway in New York.

The other method means that express and slow train running on the same track, when the express is about to overtake the slow train, the slow need to stop on the dodging line at the next station waiting for the express going out first [3].This is also the main pattern of subway in Tokyo and Beijing subway line 6. It saves the underground space and the building cost.

54.2.3 Features of the Express–Slow Train Mode

The express–slow train mode can effectively reduce the negative effects brought by the frequent stop, improve the train speed, shorten the traveling time, and provide a better service for the passengers. Also it improves the running efficiency and decreases the number of running train.

But this combined running mode also brings some negative effects: the passengers on the overtaken train will have more waiting time because of the crossing station stop. In the case of higher density of trains, the express will overtake the slow train lowering the passing capacity. In addition, excessive crossings will cause more difficulty for the building and increase the cost. However, too few crossings will lower the passing capacity and affect the departure balance [4].

Both New York subway and parts of Tokyo metro is the model for the express–slow train mode in foreign cities. So, study on their subway system can provide enlightenment for the development of the express–slow train in our country.

54.3 Express–Slow Train in New York

54.3.1 Broadway-Seventh Avenue Line

Broadway-Seventh Avenue Line is a main line in the New York subway system. There are New York Line 1, Line 2, Line 3 running on it. From 96 Street in the middle of Manhattan to Chambers Street, IRT runs in the express–slow train mode.

54.3.1.1 Line and Station Overview

IRT runs in four tracks mode from 96 Street to Chambers Street, the line and station schematic diagram is shown in Fig. 54.1.

As in Fig. 54.1, the express of Line 2 and Line 3 runs on the inner two tracks while the slow train of Line 1 runs on the outer two. The highest speed of both express and slow trains can be 56 km/h. There are 18 stations in this interval, 6 of them as express station. The express can save 30–40 s by abandoning one stop.

Figure 54.1 shows that most express station layout is double-island platform, except 34 Street Penn Station. The layout is adopted because of the increasing passengers. According to the New York metropolis, the passenger flow value at 34 Street Penn Station is 27010176 in 2012. Using this layout can lead the passengers to transfer at the neighboring station—Times Square.

54.3.1.2 Transfer Interface

New York subway provides passengers a detailed timetable. The departure interval of Line 1 is 3 min, 4–6 min for Line 2, and 5–7 min for Line 3. The rate of the express and the slow train is about 1:1. If a passenger wants to transfer at Times Square station, his waiting time will be less than 1 min in the peak period.

To take a travel at 8 a.m. in the peak period for example, the superiority of the express over the slow train can be seen. From 50th Street to Chamber Street, if Line 1 is chosen, there are 12 stations and the total time is 18 min. If Line 3 is

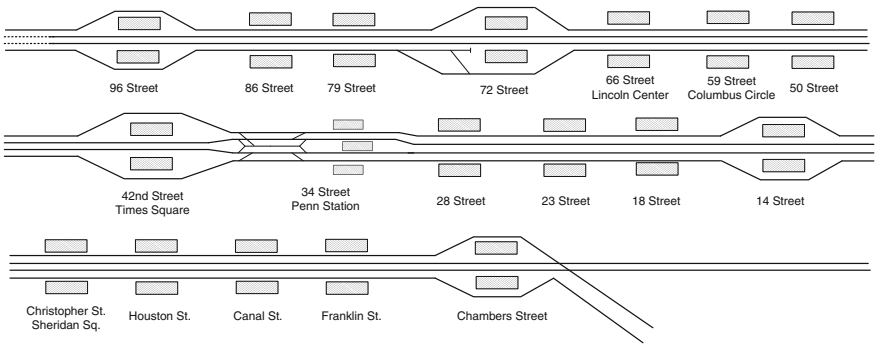


Fig. 54.1 Party-line from 96 Street station to Chambers Street station

chosen, it will cost 2 more minutes for transfer, but the stop stations decrease to 5, the total travel time is only 13 min.

For the long-distance passengers, the express not only saves the traveling time, but also reduces the burden causing by the frequent stop. If there is express–slow train in a line, the passenger flow of the express is usually 2–3 times more than the slow train.

54.3.1.3 Running Organization

Figure 54.2 shows the running status from 72 Street to 34 Street in 25 min in the morning peak period. The express of Line 2 and Line 3 reach alternately. The departure interval of the express and the slow train is 2–3 min, the ratio of the express and the slow train is 3:2, but the passengers of express is 4–5 times of slow train in the peak period, leading to the crowding in the express and the low taking rate in the slow train [5].

New York subway works 24 h a day. It brings great convenience for people traveling at night. The main difficulty of 24 h running is the track maintenance. The collinear of Line 1, Line 2, and Line 3 sets 3 stop lines and 16 crossing lines. Although the number of stop line is few, bidirectional running can be continued when one track goes wrong because of the four tracks mode. As there are enough crossing lines on which trains can transfer to normal tracks, the late arriving decreases in evidence [6].

54.3.2 Flushing Line

54.3.2.1 Line and Station Overview

Flushing Line connects Queens and Manhattan, and it is a part of the IRT system. Like Broadway-Seventh Avenue Line, Flushing Line also have special line for express. But from 33 Street to Main Street it adopts three tracks mode, the express

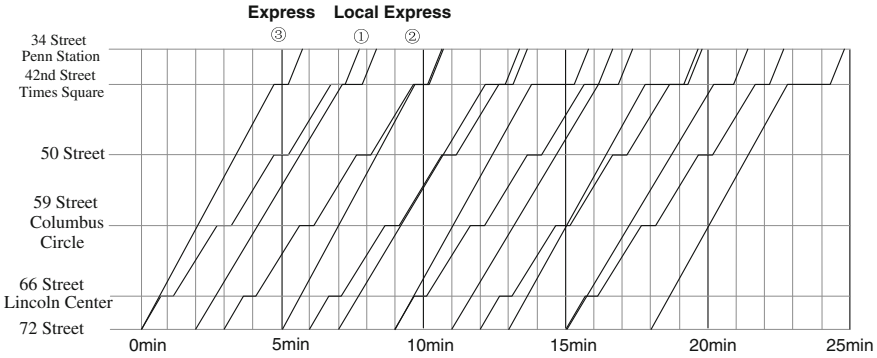


Fig. 54.2 Train operation diagram between 72 Street Station and 34 Street Station

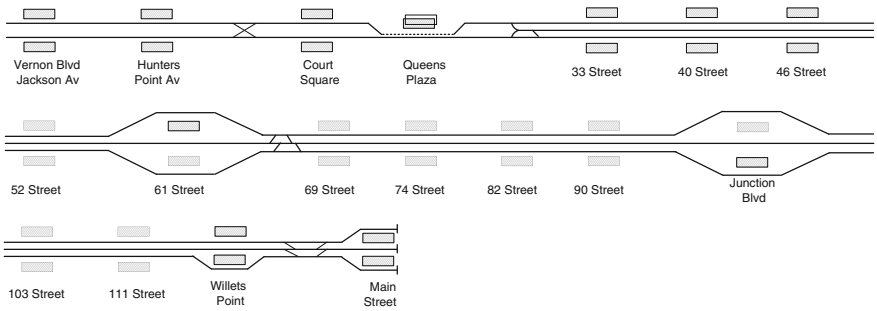


Fig. 54.3 Lines from Vernon Blvd Jackson Avenue station to Main Street station

runs on the middle single track in the peak period. Figure 54.3 shows parts of Flushing Line and its stations.

The express–slow train of Line 7 serves for the whole Flushing Line. The slow train of Line 7 runs for 24 h day. Since the restriction of the number of tracks, some expresses run from Main Street to Times Square between 6:20 and 10 a.m., some run from Times Square to Main Street between 2:50 and 10 p.m.

54.3.2.2 Contrast of Express Single-Line and Double-Line

According to Table 54.1 to summarize, special express single-line is more suitable for the line connecting the city center and the residence with a tidal passenger flow. Special express double-line provide express–slow train in the same service interval, it is suitable for great double-direction passenger flow in 24 h.

Table 54.1 Contrast between two mode

Lines	Geographical layout	Economy	Capacity
IRT Broadway-Seventh Avenue Line	Three wide and deep branches in two ends	High engineering and maintenance costs	Enough tracks for large flow
	Crossing CBD Manhattan	Wasting space	Services two directions at the same time
Flushing Line	Connecting Manhattan (CBD) and Queens (residential district)	Saving cost and space	One direction only
			Lower passenger capacity
			Simpler flow inside the station

54.4 Express–Slow Train in Tokyo

Tokyo metro has a long history since it has various kinds including subway, urban railway, linear motor, and so on. The express–slow train on Tsukuba Express share the same tracks.

54.4.1 Tsukuba Express

54.4.1.1 Line and Station Overview

Tsukuba Express connects the Akihabara Station and Tsukuba Station. In total, it is 58.3 km in length and the average length between each station is 3.1 km. There are 20 stations in the whole line among which there are only 6 stations equipped with wiring. And the wiring is very simple. Among the 20 stations there are 3 stations which do not have scissors crossover and 2 stations setup preventing collision line because of the overtaking [7]. These two stations are Yashio station and Nafareyama-otakanomori station as shown in Fig. 54.4.

54.4.1.2 Running Organization

In terms of the urban railway, since the requirement of passenger flow is quite much, the high-speed car collinear operation organization is very complex. And the signal control, scheduling systems, and vehicles are put forward higher requirements. Therefore, Tokyo speed collinear configuration only in the ground or the ground line.

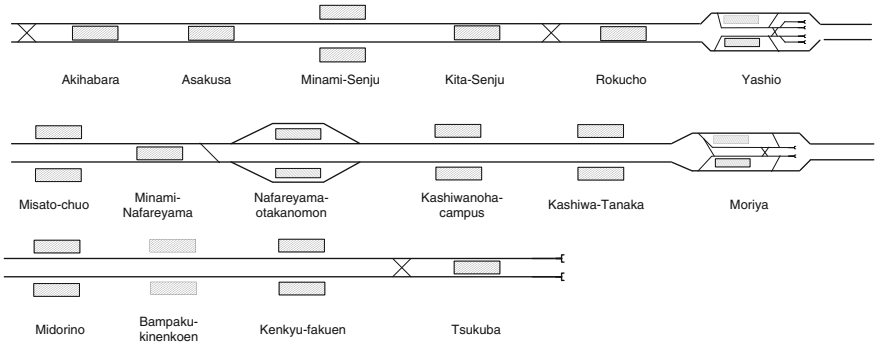


Fig. 54.4 Lines of Tsukuba Express

Tsukuba Express has three kinds of cars, namely, express, shuttle bus, and slow. The express stops at seven stations, including stations before Rokucho station, Minami-Nagareyama, Nagareyama-otakanomon, Moriya, and Tsukuba statation, travels 45 min; shuttle bus stops at 16 stations, travel time 52 min while the slow train 57 min. In contrast, the express saves 12 min, and the shuttle 5 min in contrast of the slow train.

54.4.2 Location of the Overtaking Station Factors

Choosing the location of overtaking station depends on characteristics of passenger flow, time-efficient, cost-benefit, and facilities and equipment of the line.

54.4.2.1 Characteristics of Passenger Flow

In general, stations that the express stops at has a large passenger flow and the parking time of slow train will be longer. In addition, the express train needs to stop to handle the passenger flow in the overtaking station and the slow train needs to wait until the express train departs. So that is the reason why it takes the slow train a long time waiting there. Therefore, we should not set the stations which have large passenger flow as overtaking stations.

54.4.2.2 Time Efficiency of Passenger

Selecting the mode of express–slow trains will increase the interval of the departing time of the train in the overtaking station. And the time efficiency of passenger will be reduced. As a result, passengers' satisfaction of this line will decrease. So spacing of the express stations should be large enough to improve travel speed effectively, otherwise it will affect passenger benefits.

54.4.2.3 Cost-Benefit of Operation

The combination of express trains and slow trains may lead to the result that in the same time the number of traffic decreases. In general, the operational benefits and socioeconomic benefits will decrease. So as a result, the express–slow train need to be operated in the section with enough capacity and the overtaking station should be set up in this section [8].

54.4.2.4 Facilities and Equipment on the Line

In order to set up overtaking stations, enough space is necessary. And the cost of setting up overtaking stations is quite large. As a result, it is better to choose to setup overtaking stations on the over-ground station.

54.5 Enlightenment

By the study on the construction and running of foreign express–slow train, the followings can be found:

1. Express–slow train is feasible in technic for our country, but as most of Beijing Subway routes are underground, it will cost too much for the construction of the special line for the express. Considering features of infrastructure of urban rail transit in our country and the demand of passenger flow, the second stage of Beijing Subway Line 6 project can learn from the Tokyo metro in express–slow train running mode.
2. New York Subway running mode can be referenced by cities like Beijing, Shanghai, Guangzhou in which are famous for separation of workplace and residence. In the designing stage, the investigation of the passenger flow should be done first.
3. Our subway system should draw lessons from New York Subway in the issue of the rate of express and slow trains. As the demand for express is obviously increased in peak period, the rate of express and slow trains should be adjusted, improving service levels.
4. Learning from the experience of timetable joint in New York, the operation diagram should be drawn in detail, and used flexibly according to the fact. As long as the running schemes are appropriate, the better running density and efficiency can be achieved.

Acknowledgments The work was supported by the National Basic Research Program of China (No. 2012CB725406), and Beijing Municipal Education Commission scientific research and graduate education projects (No. T11H100010).

References

1. Pan X, Wei Q, Qiu L, Cheng W (2012) Mode of express and slow routes applied in the design Beijing subway new lines. Urban Rapid Rail Transit 5:33–37 (in Chinese)
2. Yang Y (2008) Development of suburban rail transit in foreign Metropolis. Urban Mass Transit 2:17–21 (in Chinese)
3. Shi R (2012) Optimization study on operation organization of urban rail transit. Southwest Jiaotong University (in Chinese)
4. Wang B, YU D, Li D (2012) Analysis on operation and organization of Tokyo Metro. Urban Rapid Transit 25(1):111–115 (in Chinese)
5. Jun R, Ping A (2011) New York City Express Train. Traffic Transp 5:69–70 (in Chinese)
6. Jiang Y, Ye X (2007) Track layout methods of typical URT in Foreign cities. Urban Mass Transit 12:35–39 (in Chinese)
7. Xiong Y (2012) Research on the express/slow train of the reginal urban rail line. Southwest Jiaotong University (in Chinese)
8. Guo Y (2009) Study on optimal problems of stop schedule plan for urban rail transit. Beijing Jiaotong University (in Chinese)

Chapter 55
Optimized Design of Urban Rail Vehicle Grounding System

Ling Gao, Ruichang Qiu, Lei Wang and Haijie Jia

Abstract For most of the metro operation within the narrow underground space, personnel and equipments highly dense, a large number of electrical and electronic equipments and their strong and weak electrical lines cross layout, electrical engineering safety issues have drawn increasing attention, and the grounding is an important protective measure to ensure the normal working of electrical equipment and personal safety, so do a good job in metro weak current system, distribution system and grounding protection is an important measure to ensure the personal safety. Based on grounding system in practice, combined with Shenzhen Metro problems in actual operation, this paper expounds the metro electrical equipment grounding types, principles, grounding device settings, and so on, puts forward the improvement program of the grounding system, and make a reasonable verification.

Keywords Urban rail vehicle · Grounding system · Grounding protection · Shenzhen metro

55.1 Introduction

With the rapid development of rail transit, reliability requirements for vehicle electrical grounding system are increasing. Vehicle electrical grounding system should not only guarantee the operation safety of the train, but also to ensure the safe and stable operation of the equipment, effectively resist electromagnetic interference from the outside, and provide reliable protection for passengers, etc [1]. High-voltage equipments are mainly composed of traction inverter device (VVVF), auxiliary power unit (SIV), and other equipments. These devices from catenary

L. Gao (✉) · R. Qiu · L. Wang · H. Jia
School of Electrial Engineering, Beijing Jiaotong University, No.3 Shang Yuan Cun, Hai Dian District, Beijing, China
e-mail: 12125843@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_55,
© Springer-Verlag Berlin Heidelberg 2014

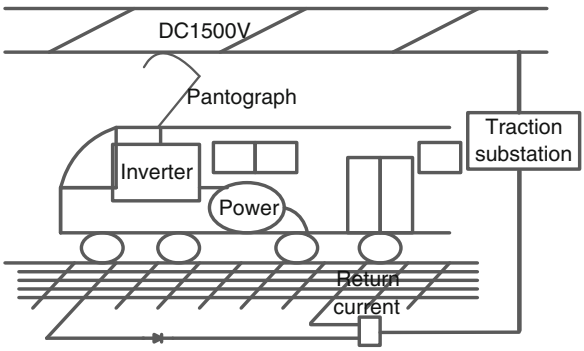


Fig. 55.1 Metro power supply

through pantograph (usually DC1,500 V high voltage) or through the collector shoes from the third rail (usually DC750 V or DC1,500 V high voltage) by current collection, after a series of power conversion and output to the control motor or other loads, the negative side by installed in the car grounding device to realize circumfluence to substation negative side, thus ensuring power equipment work smoothly and reliably [2]. In the metro, power supply system return path as shown in Fig. 55.1: the positive of traction substation → catenary → pantograph → vehicle load → wheel → track → underground return line → the negative of traction substation. Or the positive of traction substation → shoegear → vehicle load → wheel → track → underground return line → the negative of traction substation [3].

Metro vehicle is not a fixed system, so the fixture with the ground in different places is the vehicle "earth" is not the earth, but only the relative zero potential benchmark, car-body.

In this paper, based on the Shenzhen Metro problems in the actual operation and technical changes made as an example, the importance of the metro vehicle ground is introduced, and the metro grounding design principles are analyzed.

55.2 Grounding System

55.2.1 Protective Grounding

The power system or a part of electrical equipment by grounding wire connected to earth is called "ground." The metro refers to the relative zero potential benchmarks car-body. In order to make all kinds of electrical equipments on the vehicle to work, and ensure the safety of passengers, to provide high-quality services to passengers, all equipment must be adequately grounded. Classification by special function, grounding is divided into protective grounding and working grounding. Working grounding is divided into earth shielding and return grounding.

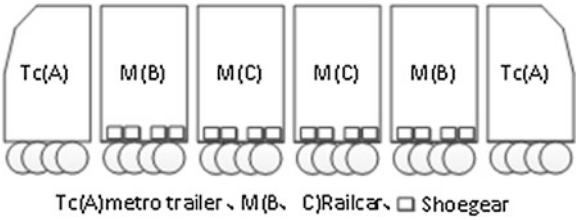


Fig. 55.2 Mode of train formation

To keep the stabilization of electronic equipment and release the charges to avoid discharging inside the equipment, and ensure the safety of workers and operators, protective grounding is essential and crucial. That all of the vehicle parts that conductive and can be raised, such as bogie, traction equipment box, auxiliary power supply box, traction motor, and so on, They may carry dangerous high voltage when they are in an irregular situations, and therefore must be connected to the body by short and thick ground wires.

55.2.2 Earth Shielding

Equipotential connection of the whole train will improve the signal ratio of the communication equipment at work, can effectively improve the quality of communication. Car-body as equipotential connection, provides a good reference for useful information. If there is a short circuit or lightning currents in the earthing device, shielding layer two grounding cable ends at different potentials, the shielded wire is current flows, shielding layer itself will form a big source of interference. Equipotential bonding of all trains, may prevent current through the shielded wire. Interference will not occur in the signal transmission.

55.2.3 Return Grounding

Metro vehicle of Shenzhen Metro line 3 uses 6 yards, divided into two power unit, each of the 3 sections of the car as a unit, grouping method is as follows: +Tc − M − M = M – M − Tc+ as shown in Fig. 55.2.

Tc cars equipped with auxiliary power supply device, and M in the middle with traction inverter equipment; Each M car has four sets of shoegears at both ends in order to achieve the flow from the third rail to high-voltage plug. When trains are running on the running lines or using the library parking, all through the vehicle's shoegears at the same side to work, providing traction system or the auxiliary inverter 1,500 V power required.

In practical work, the main work is as ground return line for power supply. As shown in Fig. 55.3, the DC1,500 V backflow of A car is connected to B car by cable and terminal flow back into the carbon brush, and then to the rail. The same

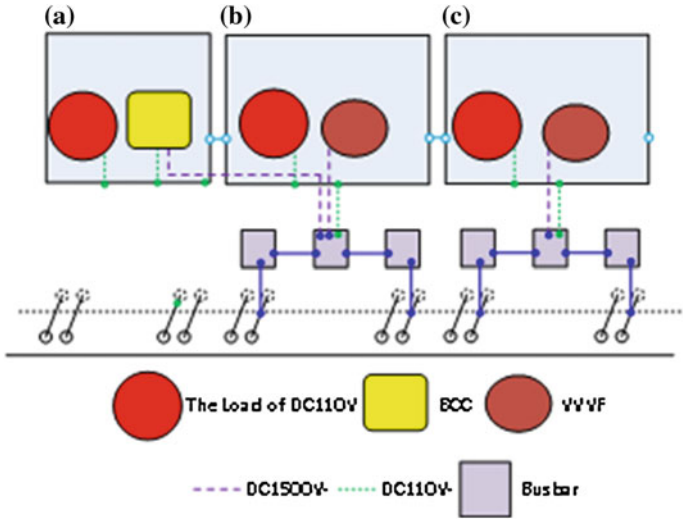


Fig. 55.3 The wiring diagram of the shenzhen metro

as B car.The DC1,500 V, and DC110 V circuit through the rail of metro vehicle to finish the reflux [4].

55.2.4 Grounding Principles

To separate return grounding and protective earthing and to separate the earthing of high-voltage circuit and grounding of low-voltage circuits, Bogie ground nearby to grounding terminal station from ground terminal to the grounding device of return line impedance consistent as far as possible. The ground should be set on the body central; Equalizer is arranged between cars,eliminate potential difference, and the low-voltage electrode wire together. The vehicle electrical equipment shell, cannot rely on a mechanical contact, must be reliably connected to body, otherwise it will cause instability of the system. Grounding point must be close contact with solid, such as brazing. Different metal welding together, to prevent the effect caused by the corrosion of chemical battery reaction. If use fastening contact, it must ensure that the contact surface is not coated with paint [5].

55.3 Problems and Improvement of Shenzhen Metro

The traction system IPM module exploded, IPM drive unit damaged, train DC arrester damaged, train air conditioning system work abnormal, train AC380 V, and DC110 V power system fault occur frequently in Shenzhen Metro Line 3 in the actual operation.

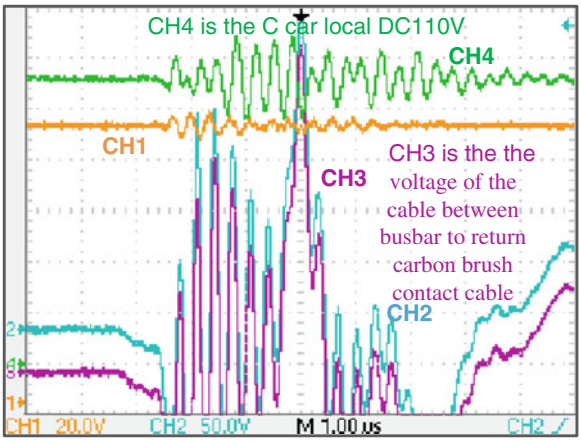


Fig. 55.4 The voltage oscillogram of C car

The main components of metro such as traction, braking, train network have the microcomputer control units. The DC110 V power supply has two kinds: there is external power supply (such as: the pantograph is raised or third rail power) and the auxiliary converter can work normally, the charger of the auxiliary converter for vehicle storage battery, at the same time the DC110 V power supply for the whole car; no external power supply, the vehicle battery for the whole car DC110 V power supply. Therefore, the quality of DC110 V is very important, if the DC110 V fluctuation is too large, it may cause damage to the traction and braking control units.

Through the scene test, record the change of traction converter DC110 V power quality and research on the voltage of the cable between busbar to return carbon brush contact cable.

The quality of C car DC110 V power input port and the voltage of the cable between busbar to return carbon brush contact cable shown as Fig. 55.4.

After analysis, the output voltage of DC110 V power is between 95–115 V and fluctuates frequently. The direct reason which causes the system to malfunction is the DC110 V instability. Due to the low-voltage DC110 V and high-voltage DC1,500 V have same grounding, DC1,500 and DC110 V grounding public impedance is large. Effect of 1500 V reflux has been the main reason which causes the DC110 V fluctuate frequently. When the large current flows from transformer substation to vehicles, causing DC110 V fluctuations, resulting in the loss of power drive module which burned IPM module.

55.3.1 The Design of the Main Grounding System of Vehicle

Before changed, the DC1,500 V and DC110 V grounding devices are together, each car connected through connection line; Tc vehicle equipment through the special grounding device connected with the grounding of M car then return to the track.

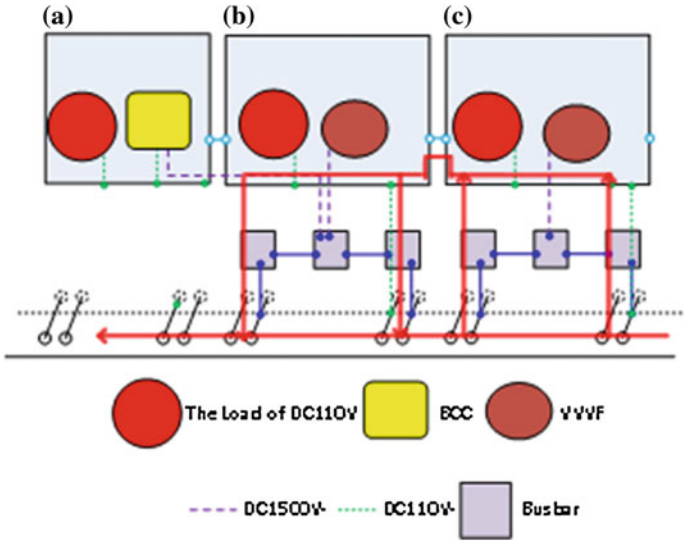


Fig. 55.5 The wiring diagram of first method

First method: DC110 V grounding directly connected to the vehicle bogie as shown in Fig. 55.5, allowing DC110 V earthing device and DC1,500 V grounding device part, thereby reducing the impaction on DC110 V.Due to the resistance of aluminum alloy car-body less than rails, the stray currents of rails might flow from C car-body to B car-body. The presence of reflux has an influence on DC110 V.

Second method: in order to reduce the stray current through the body to the other car-body, A car DC110 V negative is connected with the B car-body, and through the resistance R(50 mOhms) links with brush as shown in Fig. 55.6, then the DC110 V grounding device and DC1500 V earthing device are isolative, making DC110 V negative relative carbon brush is in a high impedance state, to prevent the current from the C car bogie → C car → B car → B car bogie for reflux, avoids DC110 V negative potential drastic change.

55.3.2 Inspect the Improved Project

Through field testing, analyzing the quality of C car DC110 V after series resistance shown as Fig. 55.7.

Comparing the two test, before C car local DC110 V voltage fluctuation peak can be more than 30 V, and the modified peak voltage is about 12 V. So, the DC1,500 V grounding and DC110 V grounding completely separated and access the grounding resistance, played a significant improvement.

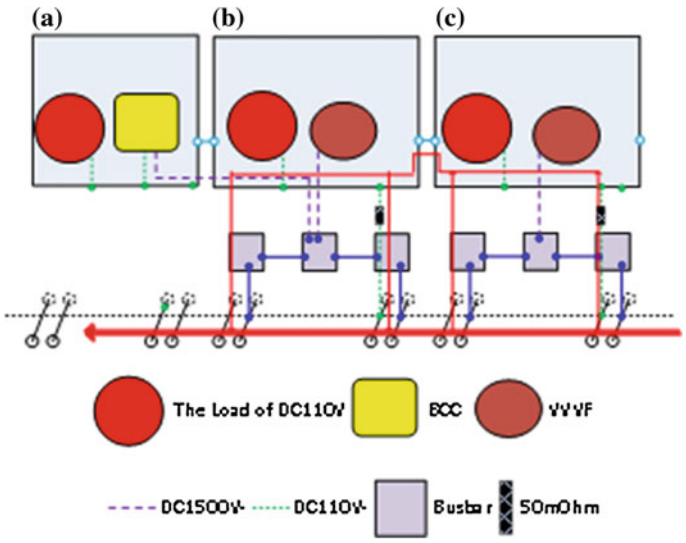
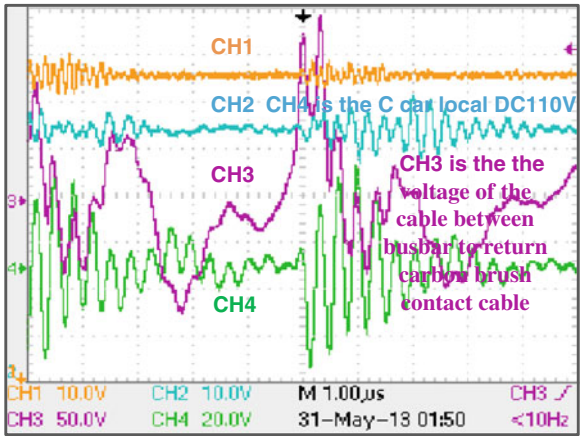


Fig. 55.6 The wiring diagram of first method

Fig. 55.7 The voltage oscillogram of C car with series resistance



As the sample points were set per 2,000 m, an interpolation per 200 m was set in both longitude and latitude direction. Given three base points $A(x_1,y_1)$, $B(x_2,y_2)$, $C(x_3,y_3)$ which are not in one straight line in the plane, and set their related.

55.4 Conclusion

With the development of power electronic technology, as the heavy current and weak current integrated integration system, metro vehicle electromagnetic environment becomes more and more complicated. The metro vehicles grounding

system provide a path for leakage current, lightning current, electromagnetic interference in the system to into the earth, thus ensuring the device is working properly.

This paper mainly attempts to analyze the grounding design through introducing Shenzhen Metro line 3. By changing the structure of grounding system can improve the vehicle's reliability and for emergency anti risk ability, especially can effectively avoid the burning problems caused by the occurrence of abnormal voltage fluctuation, such as IPM equipment. The study is a reference to the existing line train and subsequent new trains [6].

Acknowledgments This work was supported by the China National Science and Technology Support Project "The Key Technology and Engineering Application Demonstration of Rail Transit Train Energy-saving" under Grant 2013BAG24B01. We would like to thank the anonymous reviewers for their comments and suggestions.

References

1. Zhou C, Cao MS, Liu Q (2009) Discussion on metro station grounding design. Urban Rapid Rail Transit 22(5):87–91 (in Chinese)
2. Zhu J, Wang JQ, Li LS (2009) Design of grounding system for urban railway vehicles. Rolling Stock 47(1):26–28 (in Chinese)
3. Zhong BY (2008). Analysis on earth methods of metro vehicles. Electric Locomotives Mass Transit Vehicles 31(4):55–57 (in Chinese)
4. Gou ZT, Lv Z (2012). Earthing design of metro train DC110 v circuit. Urban Mass Transit 15(9):94–97 (in Chinese)
5. Cao HK, Zhang A, Jiang FJ (2008) Treatment of electromagnetic interference in the subway trains. Raliwany Locomotive Car 2008:525–529 (in Chinese)
6. He BG, Zhang ZQ (2001). Seismic response of metro tunnel in Beijing stratum. In: 2011 International conference on computational and information sciences (ICCIS), pp 893–896

Chapter 56
Analysis of the Property of Heavy Haul Railway's Traffic Flow Based on Hybrid Cellular Automaton

Wentan Deng and Huibing Zhao

Abstract According to the property of heavy haul transportation on Da-Qin line, this paper, based on hybrid cellular automaton, proposes a traffic flow model to simulate the operation of the train on Da-Qin line and to analyze the property of traffic flow. It investigates the relationship between the original speed and the braking distance of 20,000 t heavy haul trains on horizontal line, and analyses the relationship between the slope and the braking distance of the trains on sloping line. The simulation results are compared with the theoretical results to verify the feasibility and availability of the model.

Keywords Heavy haul railway · Hybrid cellular automaton · Traffic flow

56.1 Introduction

Because of the advantages of the huge capacity, high efficiency, low cost, energy saving, heavy haul railway, which is considered as the future of international freight transportation, develops rapidly around the world. The studies of the heavy haul railway's operation rules and traffic flow features have both theoretical significance and practical value for protecting of traffic safety, improving the efficiency of transportation and reducing energy consumption.

Railway transportation is a complex system which presents the nonlinear, stochastic, and dynamic features. Using the traffic flow theory to build a model to describe the general characteristics of the actual train traffic flow is the effective method to research the features of heavy haul railway. Cellular automaton (CA) model which is discrete in time, space, and state is a useful tool to describe

W. Deng (✉) · H. Zhao
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
e-mail: 10111051@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_56,
© Springer-Verlag Berlin Heidelberg 2014

nonlinear phenomena and study the behavior of complex systems. It has incomparable superiority compared with the traditional mathematical model for its simplicity, parallelism in description, and evaluation of the traffic flow [1].

In 1950s, cellular automaton was first proposed by John von Neumann to describe the self-replication mechanism of biological [2]. S. Wolfram proposed the rule No.184, and modeled the traffic flow by one-dimensional cellular automaton for the first time in 1980s [3]. In 1990s, the cellular automaton developed rapidly and used widely to traffic simulation. In 1992, Nagel and Schreckenberg improved the rule No. 184 and proposed the famous NS model [4]. The model rules were simple and easy to calculate, which could describe some of the actual traffic phenomena, such as the stop–go phenomena and congestion. In 2005, Li Keping, Gao Ziyou, Ning Bin proposed the railway cellular automaton model based on the NS model, which analyzed the characteristics of traffic flow [5–8]. Zhou Hualiang, Li Feng, Fu Yinping, Xun Jin improved railway cellular automaton on the basis of the NS model and the CA model in Ning [8] and simulated the train operation in fixed block, moving-like block, and moving block signal system and investigated feature of traffic flow, such as the delay propagation of train, train following model, the ability of railway, etc. [9–12].

With the limits of the homogeneous, synchronization, and the discrete state space of cellular automaton and without considering the differences of the line conditions (curves, ramps, bridges, and station), train case (locomotive performance, train length, passenger, or freight train), and people behavior (drivers, dispatchers), research efforts just modeled a simple horizontal and straight rail transport system. This paper presents a hybrid cellular automaton model with attributes, based on the NS model and the CA model in Ning [8] and simulates the characteristics of heavy haul railway. We apply this model to simulate the characteristics of traffic flow and analyze time–space graph and time–velocity curves. The simulation results have verified the effectiveness and availability of models, which could be used to improve transport efficiency.

56.2 Heavy Haul Railway

The heavy haul railway is a line which can transport at least 40 million tons annual, the length 150 km, the axle weight of train no less than 27 tons, and the weight at least 8,000 tons [13]. The freight organization, overloading, running safety, and operations management of heavy haul railway are different from ordinary freight trains and passenger trains. Heavy haul railway is noted for its large load, long braking distance, slow starting, complex operation, and other characteristics. According to a single heavy haul railway's role in the network and the juncture condition with other railway network, it can be classified into closed mode, semi-closed mode, and full open mode [14]. After years of development of heavy haul railway in China, the 5500–5800 tons unit trains operate in Beijing-Shanghai, Beijing-Guangzhou, Beijing-Harbin, and other existing busy main lines

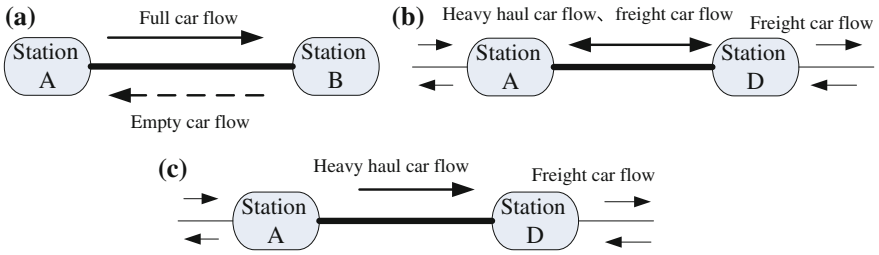


Fig. 56.1 The modes of heavy haul railway. **a** Closed mode. **b** Full open mode. **c** Semi-closed mode

to improve the transport capacity of these lines; 10,000 and 20,000 tons combined and unit trains operate in Datong-Qinhuangdao and Shuozhou-Huanghuagang coal lines (Fig. 56.1).

56.3 Hybrid Cellular Automaton

Hybrid cellular automaton (HCA) was first proposed by Canadian W. Pries in 1986, which applied different updating rules in different cells or by the external parameters to overcome the limitations of homogeneity in original cellular automaton [15].

This paper proposes a hybrid cellular automaton model with attributes, which is defined as follows:

$$A = (L_m, M, N, S, f_n) \tag{56.1}$$

where A is hybrid cellular automaton with attributes; L_m is a m long cell; $M = \{m_1, m_2 \ldots m_k\}$ is the collection of cell properties; $N = \{l_1, l_2 \ldots l_k\}$ is the collection of cell neighbors; S is collection of cell status; f_n is cellular state transition function set (Fig. 56.2).

In the time t, the characteristic polynomial of cellular automaton is:

$$A^{(t)}(x) = \sum_{i=0}^{m-1} s_i^{(t)} x^i \tag{56.2}$$

where $s_i^{(t)}$ is the state of cell i in time t and $s \in [0, 1 \ldots S_{\max}]$.

The state of cell i in time $t+1$ depends on its own state in time t and its neighbor states:

$$s_i^{t+1} = f_k\left(M, N, s_i^t\right) \tag{56.3}$$

where $f_k \in f_n$ is the state transition function; N is its neighbors; $M_i, M_{i+1} \ldots M_N$ are the attributes of its neighbors.

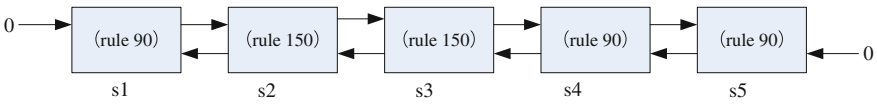


Fig. 56.2 The hybrid cellular automaton

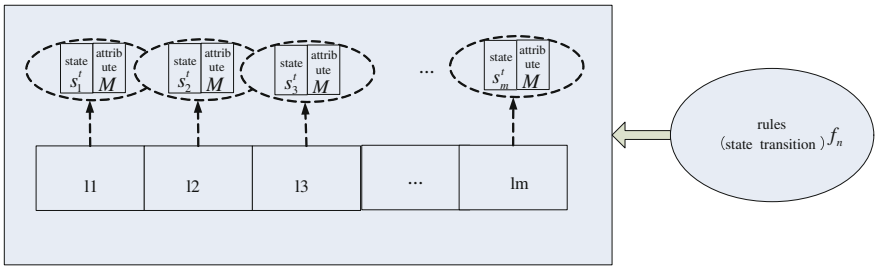


Fig. 56.3 Hybrid cellular automaton model with attributes

Hybrid cellular automaton model with attributes shown in Fig. 56.3 has cell attributes. The cell state updating rules depend on the properties of cellular and the dynamical selection of external parameters.

56.4 Hybrid Cellular Automaton Model of Heavy Haul Railway

Based on the NS model and the CA model in Ning [8], this paper establishes the hybrid cellular automaton model of Da-Qin line. The railway tracks are divided into L cells with different length and different updating state rules, $i = 1, 2 \ldots L$. The train speed is $v \in 0 \sim \text{int}(V(t))$. Each cell has its own attributes, which defined as: $M = \{\text{type, state, signal, length, gradient, } V_{L_{\max}}\}$. Station model has multi-platforms to stop a number of trains and the behind train can overtake the front train at the station. Train must stop outside the station if the station platforms are all occupied. The model is four-aspect fixed block system, m cells compose a block section and the protection lights are set in the entrance of station. $V_g, V_{gy}, V_y, 0$ are the limiting speeds of green, yellow-green, yellow, red lights. Heavy haul trains employ cyclic braking, normal braking, emergency braking, and other modes under different conditions of lines; but in the specific long-big ramp segment, they must brake many times to ensure safety (Fig. 56.4).

The dynamic limiting speed function is defined as:

$$V(t) = F(s, \text{signal}, V_{T_{\max}}, V_{L_{\max}}) \tag{56.4}$$

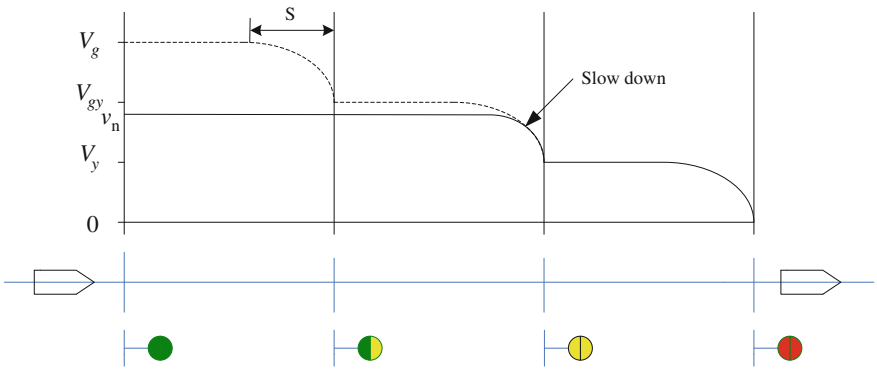


Fig. 56.4 Limiting speed curve of train

where S is the distance of train from the front signal light; signal is the color of the signal light; $V_{T_{\max}}$ is the maximum speed of train; $V_{L_{\max}}$ is the maximum speed of line.

Model algorithm steps are:

1. According to the schedule of trains, the model starts trains from starting stations directly with open boundary condition and sets the number, type, length, maximum speed, acceleration, deceleration, stop time, and other attributes of trains. The trains leave the model when they reach the boundary.
2. According to the state transition function, update each cell state, the train speed and position one by one.

$$\begin{aligned} &if\ v_{n-1} < V(t) \\ &v_n = v_{n-1} + a \\ &elseif\ v_{n-1} > V(t) \\ &v_n = v_{n-1} - b \\ &else \\ \text{Speed update}: \quad &v_n = v_{n-1} \end{aligned} \tag{56.5}$$

$$\text{Location update}: \quad X_n = X_{n-1} + v_n \tag{56.6}$$

3. Update the entrance signal lights of each block section. If the block section is occupied by a train, signal light sets to red, followed with yellow one, green-yellow one, and green one.

System updating interval is 1 s. Each cell is 1 m. When the maximum speed of the train is $V_{T_{\max}} = 30\,\text{Cell/s}$, the corresponding speed of the train is 108 km/h.

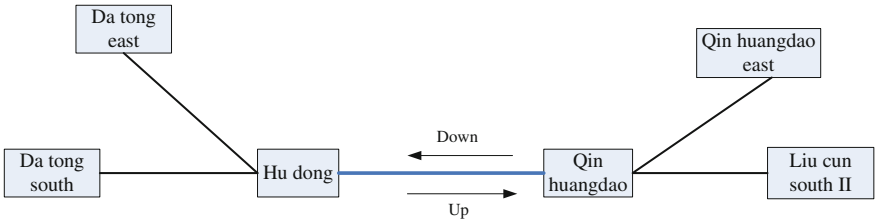


Fig. 56.5 A railway network

56.5 Simulation and Analysis

A railway network is modeled by Fig. 56.5, which is a simple model of Da-Qin line. Trains stop few or stay less to do some technical operations or pass the stations directly between Hu dong station and Qin huangdao station. To reduce the complexity of the model, a railway line is used to replace the stations in middle of Hu dong and Qin huangdao station. Railway lines are all the double-track and there are a total of six stations of Da tong South, Da tong East, Hu dong, Qin huangdao, Liu cun South II, Qin huangdao East station. Each station has 10 cells long. Da tong South-Hu dong line has 350 cells; Hu dong-Qin huangdao line has 750 cells; Qin huangdao-Liu cun South II line has 250 cells; Da tong East-Hu dong line has 500 cells; Qin huangdao-Qin huangdao East has 500 cells. 10,000 tons and 20,000 tons combined trains and unit trains start from Da tong South and Da tong East station. 10,000 tons train has six cells long, $V_{T_{\max}} = 30\,\text{Cells/s}$, acceleration and deceleration $a = b = 1\,\text{Cell/s}^2$; 20,000 tons train has 12 cells long, $V_{T_{\max}} = 25\,\text{Cells/s}$, acceleration and deceleration $a = b = 1\,\text{Cell/s}^2$. Qin huangdao East and Liu cun South II are unloading stations. Each 10 cells compose a block section; the initial signal lights are green; the 5–10th block sections between Hu dong- Qin huangdao line are long-big ramp sections, gradient $= -12‰$. Limiting speeds of green, green-yellow, yellow signal lights are $V_g = 25, V_{gy} = 18, V_y = 8$. Simulation time is $T = 1000$.

56.5.1 Train Operation Simulation under Mixed Running Condition

Figure 56.6 is a time–space and time–velocity curve diagram. About 10,000 and 20,000 tons combined trains start from Da tong south station in up direction and empty trains return from Liu cun South II station in down direction. The train accelerates from the starting station and runs at the highest limiting speed as it reaches the maximum speed. The stopping train decelerates and stops at the station; the passing train slows down the speed, passes the station, and then accelerates to leave. The horizontal axis is time and the vertical axis is distance as

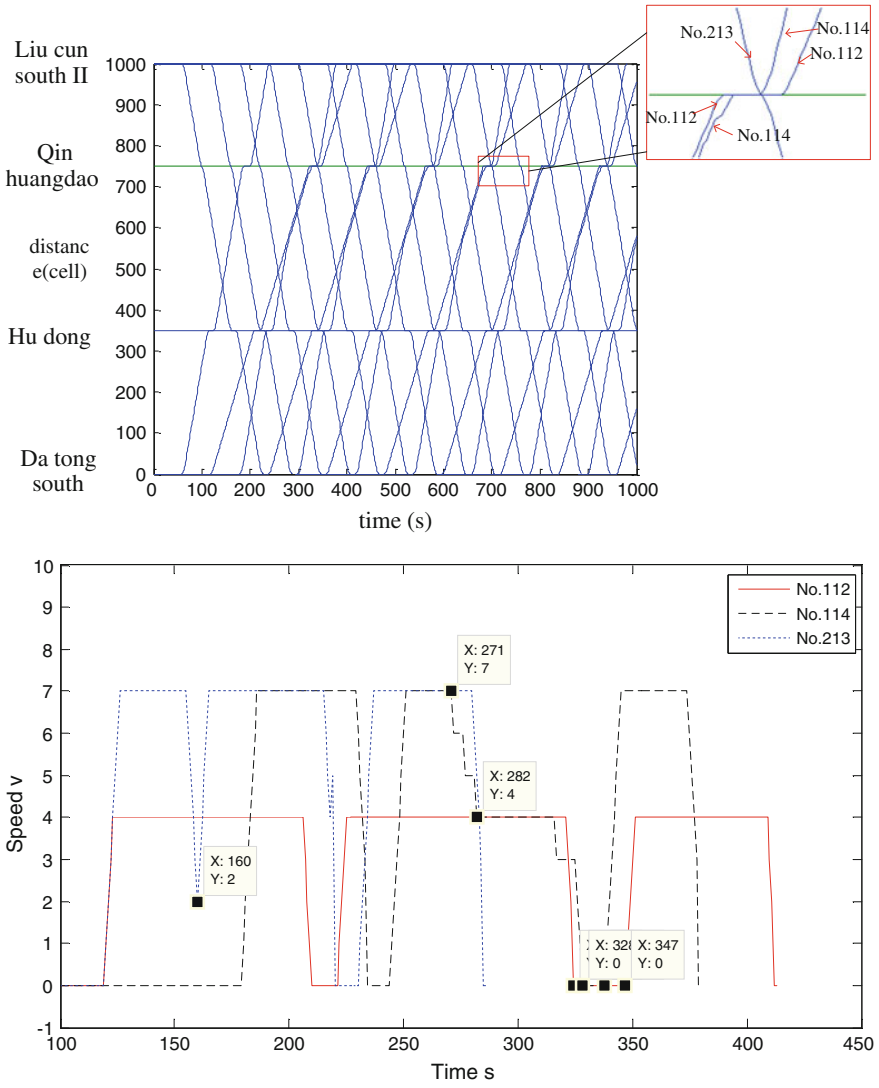


Fig. 56.6 Time–space diagram of train operation simulation

shown in Fig. 56.6a. The up-bias lines are the time–distance relationship of the trains starting from Da tong south station in up direction; the down-bias lines are the time–distance relationship of the trains return from Liu cun south station in down direction. The horizontal axis is time and the vertical axis is the speed of train in Fig. 56.6b. When the train No. 213 in down direction approaches Qin huangdao station, it decelerates at time 154 and slows down to $v = 2$ to pass through it at low speed safely, and accelerates to leave at the highest speed $v = 7$

as it reaches the maximum limiting speed in the interval section. About 10,000 tons unit train No. 114 and 20,000 tons combined train No. 112 both run at its maximum limiting speed $v = 7$, $v = 4$; train No. 114 catches the front train No. 112 at $t = 271$, slows down and follows it at $v = 4$ with a safe distance. Train No. 112 stops at Qin huangdao station waiting for No. 114; train No. 114 stops 10 s at the station and overtakes No. 112, and then accelerates to leave the station before No. 112 at the maximum speed $v = 7$; train No. 112 continues to stop 20 s at the station, then accelerates to the maximum speed $v = 4$. Train No. 114 and No. 112 complete this overtaking.

56.5.2 Analysis on the Braking Distance of the 20,000 tons Combined Train in the Straight and Ramp Line

The braking distance is a significant parameter in station designing and train operation. By analyzing the braking distance under different braking modes and the initial velocity in the model, the simulation results are compared with the calculated theoretical values to verify the validity and availability of the model.

Twenty thousand tons combined train in Da-Qin line is composed of four SS_{4G}locomotives and 210 C76 or C80-type coal cars. According to regulations, the pipe pressure of locomotive is 600 kPa. But the pipe pressure of it sets to 500 kPa when it is used to calculate the emergency braking distance for reasons brake safety. The calculated theoretical value is based on the reference [16] and the calculation process does not include the driver reaction time and equipment response time.

Under different initial velocity, the braking distance of simulation and the theoretical value comparison are shown in Fig. 56.7. The solid line is the theoretical value; dashed line is the simulated value; red line for the pipe pressure of brake system is 600 kPa; the blue line for the pipe pressure is 500 kPa. As shown, the braking distance of train increases nonlinearly by the increase of initial speed, and its rate increases accordingly. When the pipe pressure is 500 kPa and the initial speed is 120 km/ h, the braking distance in straight line is 1,315 m, which is less than the 1,400 m requirement; when the pipe pressure is set to 500 kPa and the initial speed is 90 km/h, the braking distance is 783.3 m, which is less than the requirement of 800 m. The emergency braking distance will increase about 15 % when the braking efficiency is at its 0.8 times. Shown in Fig. 56.8, the mean square percentage error (MSPE) of the braking distance theoretical values and simulation values is at 2–4 %, which indicate that the model is valid and available.

Because the heavy haul train is overweight and super long, its braking force is complex in different slope grades. The train must brake many times in long-big ramp line, to ensure the safety. In the simulation, the pipe pressure is assumed to 500 kPa and the initial speed is set 120 km/h. As shown in Fig. 56.9, the train

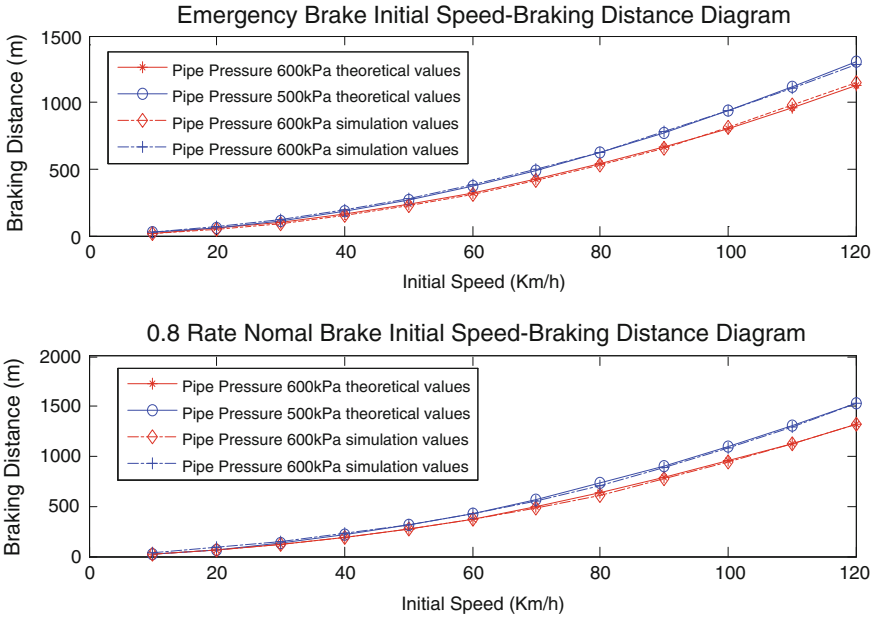


Fig. 56.7 Initial braking speed and the braking distance relationship diagram of train

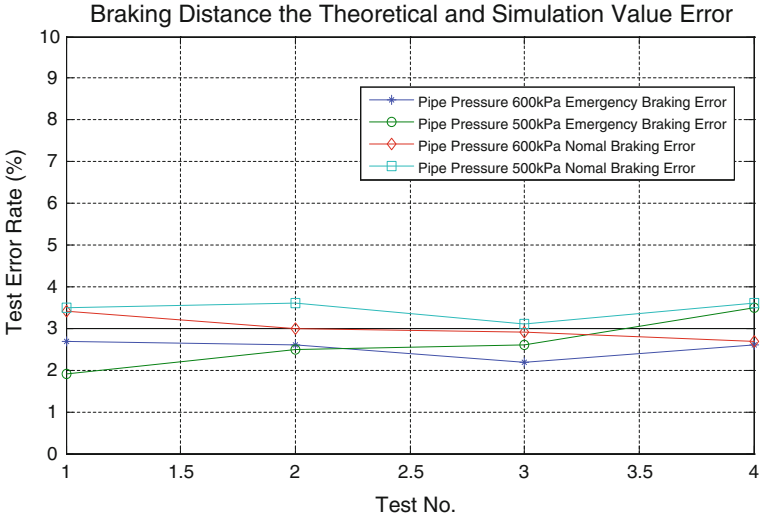


Fig. 56.8 Mean square percentage errors of the braking distance theoretical values and simulation values

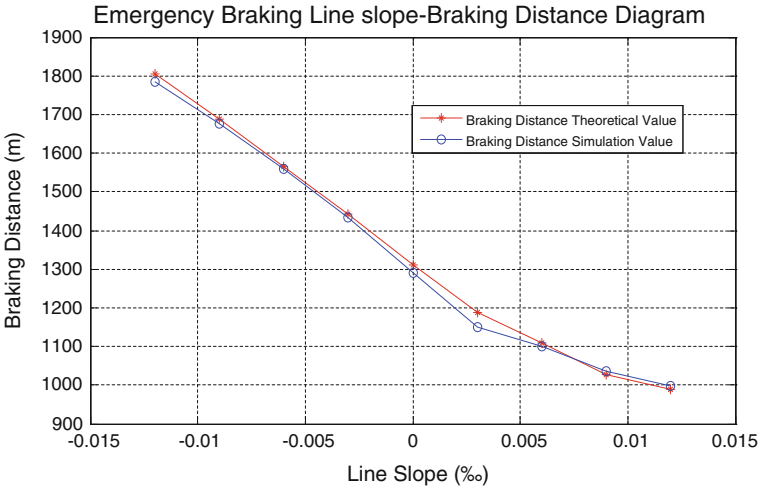


Fig. 56.9 Line slope—braking distance relationship diagram

braking distance increases linearly by the slope of the ramp. When the slope is $-12‰$, the braking distance reaches at 1,785 m. As a result, the theoretical and simulation value error is less than 3 %, which verifies the validity of the model.

56.6 Conclusion

According to the characteristics of heavy haul transport, this paper proposes a hybrid cellular automaton model with attributes based on the NS model, to simulate the meeting, tracking, overtaking of train, and other traffic flow phenomena. By analyzing the space–time diagram, the simulation results demonstrate the relationship of train's time–speed-distance. It is also analyses the relationship of the 20,000 tons combined train's initial speed—the braking distance and the slope-the braking distance in the straight and ramp lines. The results show that when the pipe pressure is 500 kPa and the initial speed is 120 km/h, the braking distance in straight line is 1,315 m, which is less than the 1,400 m requirement; The emergency braking distance will increase about 15 % when the braking efficiency is 0.8 times. Finally, the theoretical values and simulation values are compared to verify the validity and availability of the model. As it shows that he model has significant theoretical and practical value in the studies of the properties of heavy haul railway.

Acknowledgments This work was financially supported by The High-tech R&D project of The Ministry of Railways of the People's Republic of China (Grant No. 2011X021-c). The authors acknowledge this support.

References

1. Hailong Wang, Yongsheng Qian (2008) The simulation Of PDL passing capacity based on cellular automaton. Railway Transp Econ 30(7):82–84 (in Chinese)
2. Neumann JV (1966) The theory of self-reproducing automata. University of Illinois Press, Urbana
3. Wolfram S (1986) Theory and applications of cellular automata. World Scientific, Singapore
4. Nagel K, Sehreekenberg M (1992) A cellular automaton model for freeway traffic. J Phys I 2(12):2221–2222
5. Keping Li, Ziyou Gao, Bin Ning (2005) Modeling the railway traffic using cellular automata model. Int J Mod Phys C 16:921–932 (in Chinese)
6. Keping Li, Ziyou Gao, Bin Ning (2005) Cellular automata model for railway traffic. J Comput Phys 209:179–192 (in Chinese)
7. Bin Ning, Keping Li, Ziyou Gao (2005) Modeling fixed-block railway signaling system using cellular automata model. Int J Mod Phys C 16:1793–1801 (in Chinese)
8. Bin Ning (2005) The research of train following operation models and traffic flow features in rail transportation systems. Beijing Jiaotong University, Beijing (in Chinese)
9. Hualiang Zhou, Ziyou Gao, Keping Li (2006) Cellular automaton model for moving-like block system and study of train's delay propagation. Acta Physica Sinica 55(4):1706–1710 (in Chinese)
10. Feng Li, Ziyou Gao, Keping Li (2007) Analysis of the property of train flow in the fixed autoblock system. Acta Physica Sinica 56(6):3158–3165 (in Chinese)
11. Yinping Fu, Ziyou Gao, Keping Li (2007) The characteristic analysis of the traffic flow of trains in speed-limited section for fixed-block system. Acta Physica Sinica 56(9):5165–5171 (in Chinese)
12. Jing Xun, Bin Ning, Keping Li (2007) Network-based train-following model and study of train's delay propagation. Acta Physica Sinica 56(9):5158–5164 (in Chinese)
13. LiXin Qian (2010) The world latest progress of heavy railway transportation technology. Electr Driv Locomotives 1:3–7 (in Chinese)
14. Shengli Xia, Hao Yang (2011) Research of the model of heavy Haul railway in China. Railw Transp Econ 33(9):9–13 (in Chinese)
15. Pries W, Thanailakis A, Card HC (1986) Group properties of cellular automata and VLSI applications. IEEE Trans Com 35:1013–1024 (in Chinese)
16. Junyong Wang, Meichuang Yang (2008) Calculation and analysis of braking of 20,000t heavy haul combined train in Datong-Qinhuangdao line. Diesel Locomotives 4:5–7 (in Chinese)

Chapter 57
Research on AC Drive Test System Based on Vector Control

Zhe Zhang, Zhi-gang Liu, Lei-ting Zhao and Li-jun Diao

Abstract In this paper, we give a further analysis of mechanical model for metro vehicle and propose a "twin inverter-motor" reciprocal power-fed AC drive test system to simulate the various driving characteristics of metro vehicle. High performance of vector control as well as torque-closed loop control is applied to this test system and the results verify its good steady and dynamic performance.

Keywords AC drive · Vector control · Mechanical model · Twin "inverter-motor" AC drive test system

57.1 Introduction

Nowadays, AC drive technology is widely used in urban rail transit worldwide. As for the research, development and production of high-power AC drive system, it requires a series of power tests on converter, traction motor, control unit as well as whole vehicle, which all depend on the operation of the AC drive test system. According to the configuration of the main circuit, there are three types of AC drive test systems, namely power-consumed, power-fed, and reciprocal power-fed ones. At present, reciprocal power-fed AC drive test system, with slip frequency control strategy adopted, is mostly in use due to its good steady performance. However, it can hardly achieve good dynamic performance.

In order to overcome the disadvantage mentioned above, this paper proposes a "twin inverter-motor" [1] reciprocal power-fed AC drive test system and high performance of vector control is applied to it. This AC drive test system does not have to enlarge its power supply capacity because of its reciprocal power-fed

Z. Zhang (✉) · Z. Liu · L. Zhao · L. Diao
School of Electrical Engineering, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian, Beijing, China
e-mail: 12121585@bjtu.edu.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_57,
© Springer-Verlag Berlin Heidelberg 2014

working mode. Thus, its energy efficiency is enhanced. In addition, when using high performance vector control to operate twin motors independently, this test system is able to simulate various driving characteristics of metro vehicle.

57.2 System Description

Figure 57.1 shows the structure figure of reciprocal power-fed AC drive test system. It consists of rectifier, inverter 1-motor M_1, inverter 2-motor M_2, and control unit.

The rectifier supplies inverter 1 and inverter 2 with DC power. Motor M_1 and M_2 are coupled with each other coaxially. Through its control unit, one motor is used to achieve the role of traction motor of metro vehicle and the other is used to simulate various driving characteristics of metro vehicle.

57.3 System Control Principle

57.3.1 Model for Coaxial Twin Motors

Figure 57.2 is the diagram of the coaxial twin motors. The torque balance between twin motors is the necessary condition for the stable operation of this test system [2].

The mechanical balance can be described as the following equations:

$$T_1 + T_2 = (B_1 + B_2)\omega/p_n + (J_1 + J_2)p\omega/p_n \tag{57.1}$$

$$T_m = T_2 - B_2\omega/p_n - J_2p\omega/p_n \tag{57.2}$$

T_1, T_2 is the electromagnetic torque of motor M_1, M_2, B_1, B_2 is the friction factor of motor M_1, M_2, J_1, J_2 is the moment of inertia of motor M_1, M_2, p_n is the polar pair, ω is the mechanical velocity of motor, p is differential operator, and T_m is the actual load torque.

From Eqs. (57.1) and (57.2), we could easily have Eq. (57.3).

$$T_1 + T_m = B_1\omega/p_n + J_1p\omega/p_n \tag{57.3}$$

It is required that T_m sticks to the load torque, namely $T_m = -T_L$. Then we would have Eq. (57.4).

$$T_2 = B_2\omega/p_n + J_2p\omega/p_n - T_L \tag{57.4}$$

As we are going to simulate driving characteristics of metro vehicle, we should set up a specific mechanical model for metro vehicle to obtain T_L.

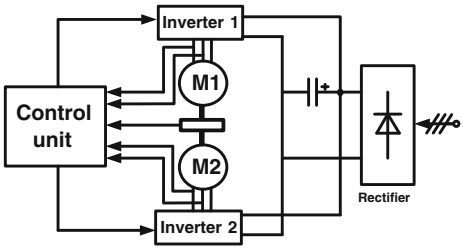


Fig. 57.1 Structure figure of reciprocal power-fed AC drive test system

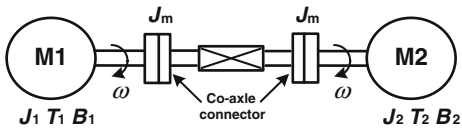


Fig. 57.2 Diagram of the coaxial twin motors

57.3.2 Mechanical Model for Metro Vehicle

For the purpose of analyzing simply and clearly, we model the metro vehicle with four wheels into single axle driven by one motor as shown in Fig. 57.3 [3].

The motion of metro vehicle is governed by the following equations:

$$F_{ad} = \mu Mg \tag{57.5}$$

$$T_L = F_{ad} \times r/G \tag{57.6}$$

$$v_d = r \times \omega_m \tag{57.7}$$

$$M\frac{dv_t}{dt} = F_{ad} - F_{\text{wind}} \tag{57.8}$$

$$J_m\frac{d\omega_m}{dt} = T_m - T_L - B\omega_m \tag{57.9}$$

v_{d} is the wheel velocity, v_{t} is the vehicle velocity, ω_{m} is the angular velocity of motor, F_{ad} is the adhesive force between rail and wheel, M is the total vehicle mass, J_{m} is the equivalent moment of inertia, T_{m} is the electromagnetic torque of motor, T_{L} is the load torque, B is he friction factor of motor, F_{wind} is the running resistance, r is the wheel radius, G is the gear ratio, and μ is the adhesion coefficient.

57.3.3 Adhesion Coefficient

How to attain adhesion coefficient is key to the mechanical model discussed above. Figure 57.4 is an illustration of adhesion coefficient under various road conditions based on slip velocity [4].

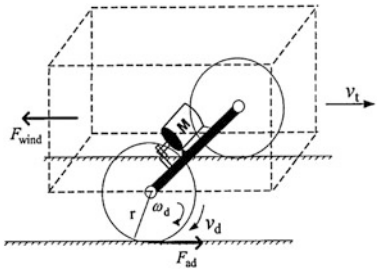


Fig. 57.3 Single axle model for metro vehicle

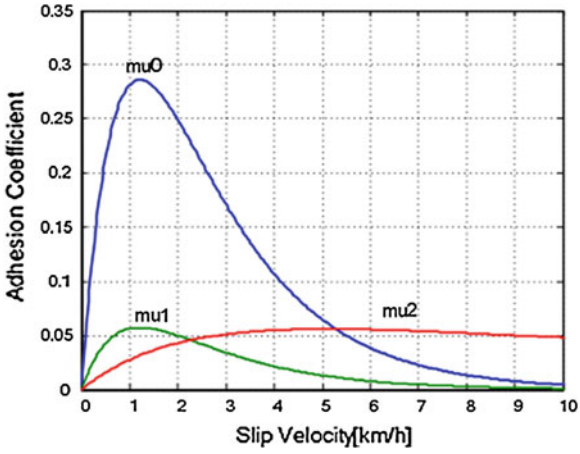


Fig. 57.4 Adhesion coefficient characteristics

Adhesion coefficient is modeled as follows:

$$v_{\text{slip}} = v_d - v_t \tag{57.10}$$

$$\mu(v_{\text{slip}}) = a \times \exp^{-b \times v_{\text{slip}}} - c \times \exp^{-d \times v_{\text{slip}}} \tag{57.11}$$

v_{slip} is the slip velocity, the parameters a, b, c, d can be designed depending on road conditions. The parameters are given in Table 57.1.

Once the control unit works out the slip velocity of metro vehicle and its corresponding adhesion coefficient, we might get the present load torque T_L through Eqs. (57.5) and (57.6).

57.3.4 Control Strategy for AC Drive Test System

In this paper, the induction motor for the AC drive test system is operated with indirect vector control [5]. Figure 57.5 shows the control diagram of twin motors based on vector control [6].

Table 57.1 The parameters for adhesion coefficient in Eq. (57.11)

	a	b	c	d
Dry (mu0)	1.0	0.54	1.0	1.2
Wet1 (mu1)	0.2	0.54	0.2	1.2
Wet2 (mu2)	0.08	0.05	0.08	0.5

Traction motor and load motor are operated with torque-closed loop as outer loop control. The reference value of T^* is given by the control unit according to the actual operation of metro vehicle and set to zero when the speed of traction motor exceeds the limit set-value ω^*. Inner loop control adopts double current closed-loop control with feed-forward voltage decoupling components added to its output. i_q* is the current component of torque and i_d* is the current component of field excitation.

As the given input of PWM control unit, stator voltage component of axle-d and axle-q should be able to control the corresponding stator current independently. Stator voltage component of axle-d and axle-q, in field-oriented control, can be described as follows:

$$u_{sd} = R_s i_{sd} + \sigma L_s p i_{sd} - \sigma L_s \omega_s i_{sq} + p\psi_r L_m / L_r \tag{57.12}$$

$$u_{sq} = R_s i_{sq} + \sigma L_s p i_{sq} + \sigma L_s \omega_s i_{sd} + \omega_s \psi_r L_m / L_r \tag{57.13}$$

$\sigma = 1 - L_m^2 / L_s L_r$, p is differential operator. It is obvious that these two stator voltage components are in a way coupled to each other. In order to achieve decoupling of these two stator voltage components, feedforward voltage decoupling section is introduced. The voltage decoupling components can be given as Eqs. (57.14)–(57.15) according to the reference value of stator currents i_{sd}* and i_{sq}*.

$$u_{sd} = -\sigma L_s \omega_s i_{sq}^* \tag{57.14}$$

$$u_{sq} = L_s \omega_s i_{sd}^* \tag{57.15}$$

Introduction of feedforward voltage decoupling section not only improves the response of current but also enhances the stability of the system [7].

57.4 Experimental Results

According to the control diagram in Fig. 57.5, we set up a 380 KW test system using SIEMENS induction motor 1TB2010-0GA02. Its rated power is 190 KW, rated voltage is 1,050 V, and rated frequency is 60.8 Hz with two polar pairs.

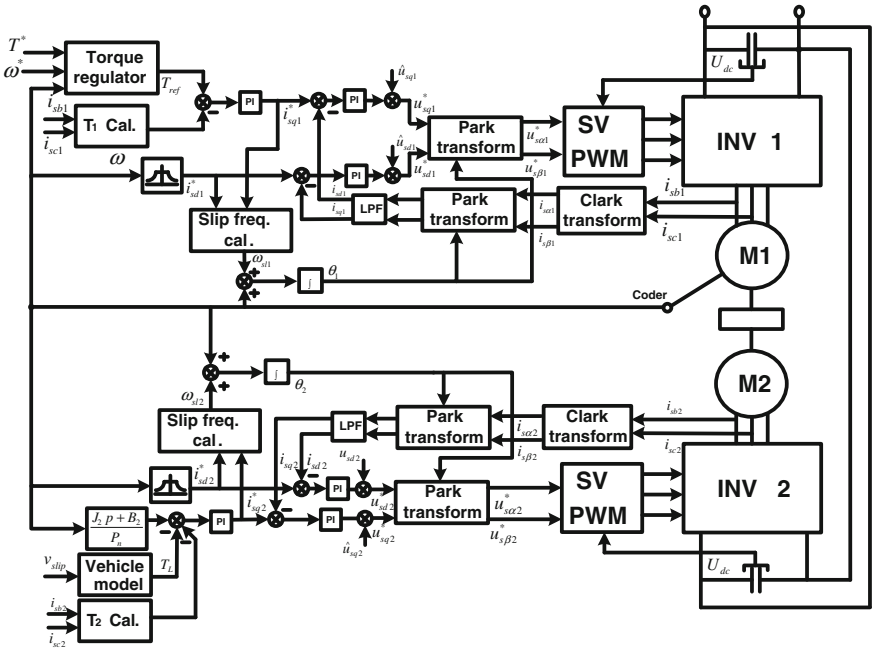


Fig. 57.5 Control diagram of twin motors based on vector control

Torque meter JN338 as well as speed sensor LENORD + BAUER GEL293 is also applied. Figure 57.6 shows the structure of this test system.

The key components of the control unit are DSP and FPGA. DSP (TMS320LF2812) serves to accomplish current and voltage sampling, SVPWM control strategy, FOC control strategy, and speed regulation. FPGA (Altera Cyclone II EP2C256C6) serves to accomplish data uploading, reverse of driving pulse and dead-time setting, motor speed detection, and fault protection.

In order to test steady and dynamic performance of the AC drive test system, we carry out a series of experiments.

Figure 57.7 shows the waveforms of steady operation. Figure 57.7a is waveform of stator current of M1, Fig. 57.7b is waveform of speed of axle, Fig. 57.7c is waveform of electromagnetic torque of M_1, Fig. 57.7d is waveform of torque of axle, and Fig. 57.7e is waveform of electromagnetic torque of M_2 when the test system was started. The reference torque for the load motor M_2 was set to zero and the traction motor operated at constant torque. When the speed of traction motor reaches the given speed 1,450 r/min, the reference torque for the load motor M_2 is set to −1,600 N.m. Finally, the test system is operating stably.

According to the load torque worked out by its control unit, this AC drive test system can adjust reference torque for load motor in real time so as to simulate various driving characteristics of metro vehicle.

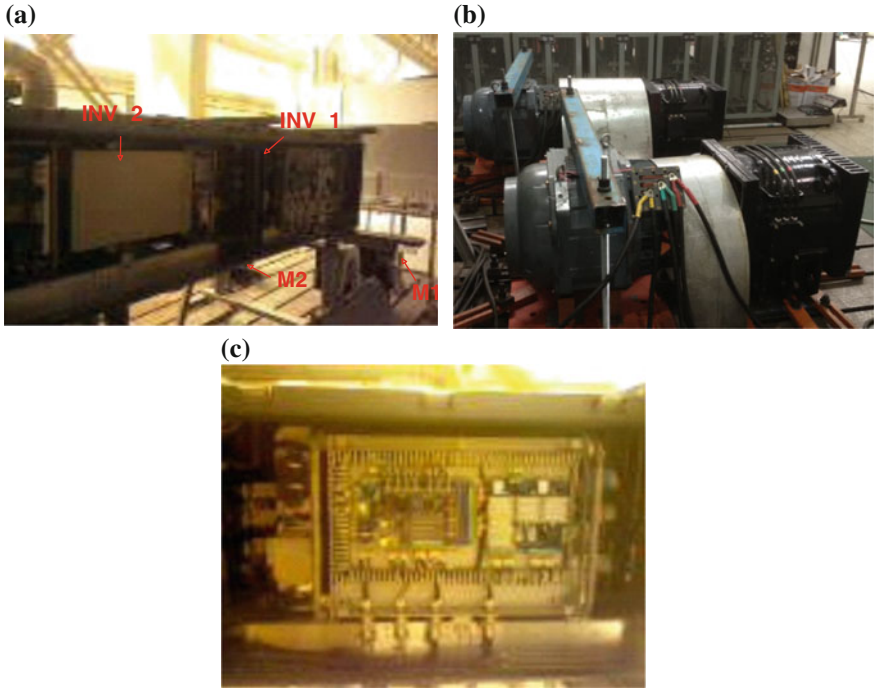


Fig. 57.6 Structure of AC drive test system. **a** Inverters. **b** Induction motors. **c** Control unit

Figure 57.8 shows the waveforms of driving characteristics while the test system is simulating that metro vehicle runs on a road in good condition. Figure 57.8a shows the waveforms of load torque and axle torque. Figure 8b shows the waveform of slip velocity of metro vehicle simulated by control unit. Figure 57.8c shows the waveform of adhesion coefficient in actual use.

Figure 57.9 shows the waveforms of driving characteristics while the test system is simulating that a wheel slip occurs when the condition of the road where the metro vehicle runs get worse. Figure 57.9a is the waveforms of load torque and axle torque. Figure 57.9b shows the waveform of slip velocity of metro vehicle simulated by control unit. Figure 57.9c shows the waveform of adhesion coefficient in actual use. It can be seen that the axle torque sticks exactly to load torque which it means. It can be seen clearly from Figs. 57.8 and 57.9 that the axle torque sticks exactly to the load torque. Figure 57.10 shows the waveforms of stator current and voltage of the traction motor.

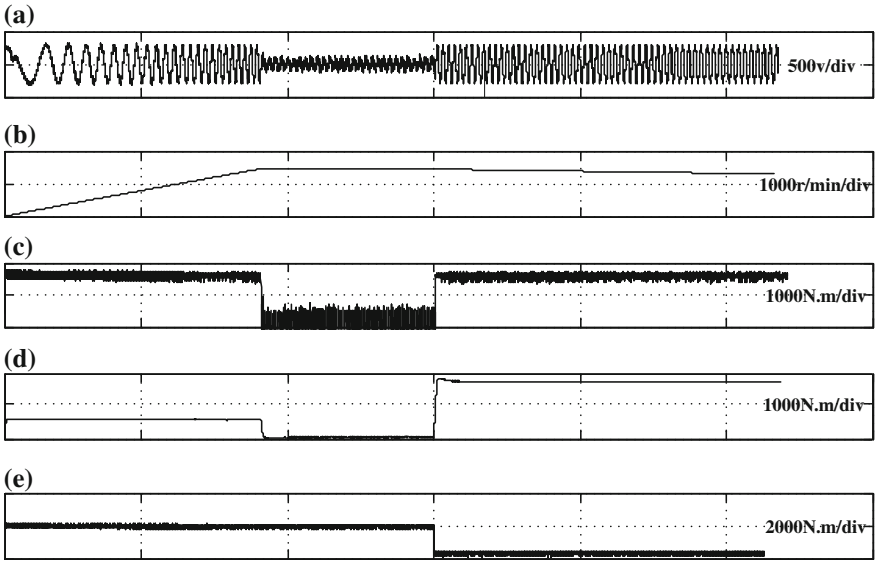


Fig. 57.7 Waveforms of steady operation

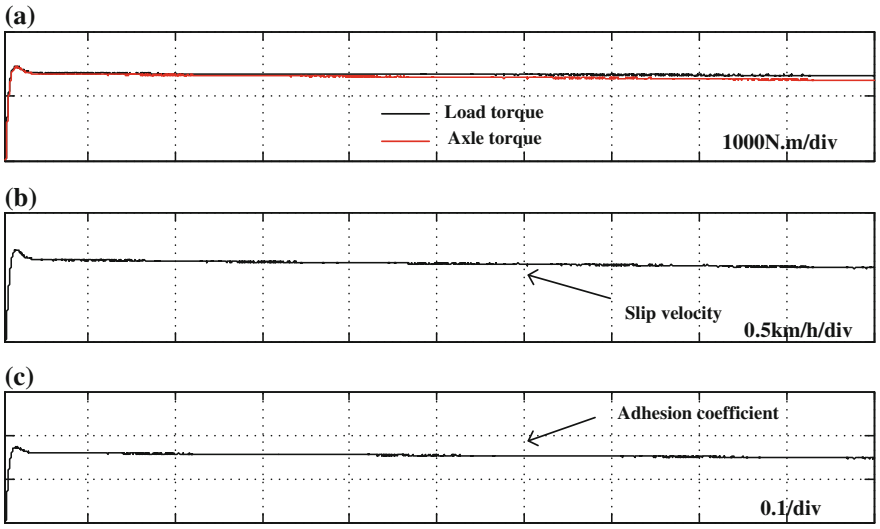


Fig. 57.8 Waveforms of driving characteristics

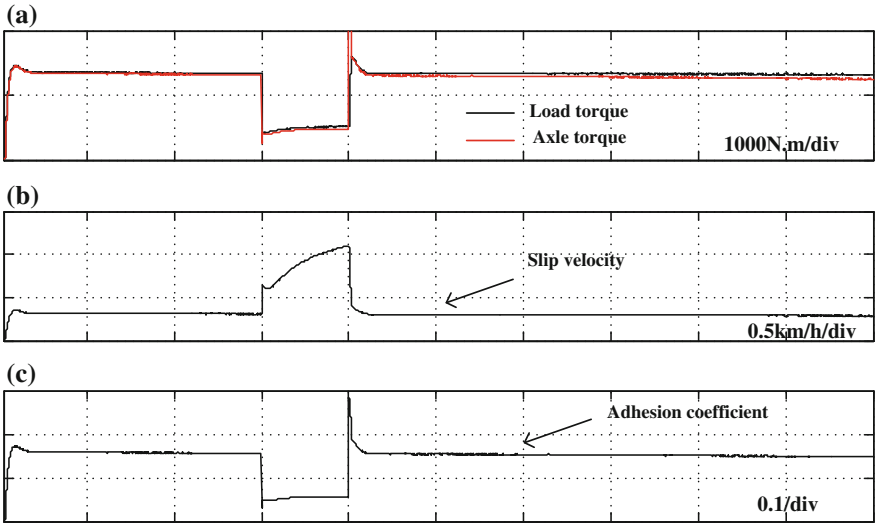


Fig. 57.9 Waveforms of driving characteristic

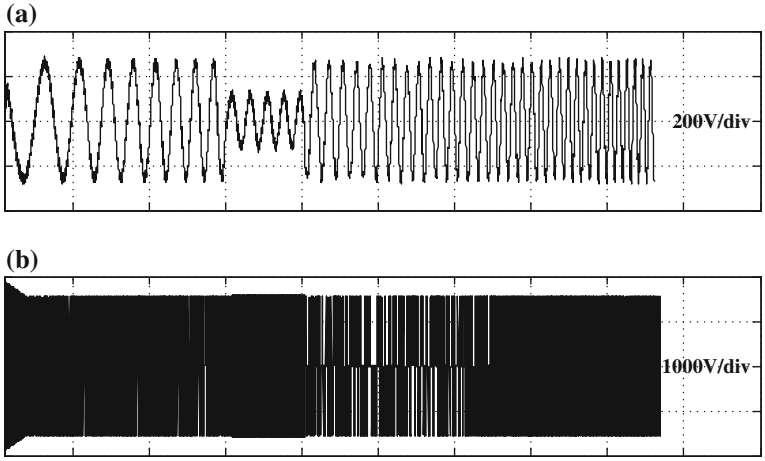


Fig. 57.10 Waveforms of stator current and voltage

57.5 Conclusion

In this paper, we have proposed a reciprocal power-fed AC drive test system. With this test system it is possible to simulate various driving characteristics of metro vehicle according to mechanical model for metro vehicle. The experimental results verify its excellent steady and dynamic performance.

Acknowledgments This work was supported in part by the China National Science and Technology Support Program under Grant 2013BAG24B01 and the National Natural Science Foundation of China under Grant U1134204.

References

1. Zheng Q, Lin F (2005) Modern Rai transportation and the reciprocal power-fed AC drive experiment system. Trans China Electrotech Soc 1:21–25 (in Chinese)
2. Sun D, Li Z, Lin W (2009) Double motors transmission experiment system based on vector control. Electric Drive 39(8):16–18 (in Chinese)
3. Lin W (2010) Analysis, modeling and optimization of metro traction system. Dissertation, Beijing Jiaotong University (in Chinese)
4. Byun Y-S, Kim M-S, Mok J-K, Kim Y-C (2007) Slip and slide simulator using induction motor. In: International conference on control, automation and systems, pp 1605–1608
5. Bauer F, Heining HD (1989) Quick response space vector control for a high power three-level-inverter drive system. EPE 1:417–421
6. Ma Z, Zheng T, Lin F (2005) Research on reciprocal power-fed AC drive test rig for electric traction applications. In: Conference on ICEMS, pp 1873–1876
7. Sun D (2012) Research on key control technologies of electric traction drive system for metro cars. Dissertation, Beijing Jiaotong University (in Chinese)

Chapter 58
Process Design and Analysis of Emergency Decision Support System for High-Speed Rail Transport Organization

Guoxing Han, Yong Qin, Tao Zhu, Jie Xu, Fei Dou and Kai Yan

Abstract The paper analyzes the emergencies which affect the rail transport organization and explores the causal relationships among the basic emergencies, abnormal phenomena, equipment fault, and dispatching orders. After that, it defines the causality of these concepts. Second, based on the principle of "safety-oriented," we establish a Causal Mapping Relationship Network (CMRN) model to describe the expression. The relationships based on CMRN model lay a solid foundation for the emergency warning and disposal theoretical system of rail transport organization. Third, designs the progress of emergency decision support system (EDSS) based on the theoretical system to reflect the dynamic evolution of emergencies during different period. Finally, it models the system process with petri net, and analyses its performance to prove the rationality of the process.

Keywords Transport organization · Emergency decision support · Causal mapping · Safety-oriented principle · Petri net

58.1 Introduction

Transport dispatch is the core of transport organization and its informatization ensures the safety in production. Nowadays, the construction of rail transport informatization in China has made a great breakthrough. However, affected by emergencies, the progress on Emergency Decision Support during abnormal

G. Han (✉) · Y. Qin · J. Xu · F. Dou · K. Yan
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian, Beijing, China
e-mail: han-guoxing@163.com

T. Zhu
Information Technology Center MOR, No. 1 Building, No. 5 Yard, Huichengmen Rd, Hai Dian, Beijing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_58,
© Springer-Verlag Berlin Heidelberg 2014

operation conditions seems to be slowing down the full implementation of rail informatization and intellectualization [1].

Models and algorithms are seldom seen on emergency intelligent decision support about rail transport organization. Li and Wang [2] proposed the method to generate the emergency disposal schemes and make emergency decision by case-based reasoning on digital emergency plan and historical cases relied on the railway emergency platform. Zhang [3] investigated and researched the crisis symptom of factors of staffs, equipment, circumstance and management, also summarized the reason rule and risk of disaster. Then he built a forewarning management system of rail transportation. All of them considered from a global perspective of the railway emergency warning, and were weak of solid theoretical foundation and not able to discover the mechanism of emergency occurring, evolution and affecting train operation. These shortcomings are not conducive to the development of rail transport dispatch informatization in emergency.

In Sect. 58.2 of this paper, we delve into the effect mechanism of railway emergency on train operation safety. Section 58.3 is aimed to create a model for emergency decision support theoretical system, so it summaries the causal relationships among the basic event, phenomenon, fault and dispatching order, also defines the causal mapping relationship and expression, at last presents the causal mapping relationship network (CMRN) model. Based on the theoretical work above and the principle of "safety-oriented," Sect. 58.4 proposes an emergency decision support system (EDSS) process of rail transport dispatch to reflect the emergency evolution of each stage, in addition, petri net is used to model the process and performance of the model is analysed to verify the accessibility, boundedness, and activity of the process.

58.2 Effect Mechanism of Railway Emergency on Train Operation Safety

The current operation dispatch and command mode have not considered the relationship between emergencies and orders or handlings. Consequently, this mode cannot recognize the effect mechanism of railway emergency on train operation safety very well, not provide early warning and prevent the deterioration and upgrade of the emergency either.

We, try to regard the "emergency and dispatching order" as dominant points, while those points between the two kinds dominant points are called implicit connection points, which is ignored. For the sake of convenient to describe, we name them: "abnormal phenomenon and equipment & device fault." To be sure, the sequence of these four points is "emergency, abnormal phenomenon, equipment & device fault and dispatching order." The points can be teased out form emergency plan, historical emergency disposal cases, and expert experience and so on, then the theoretical system of EDSS can be formed.

58.3 Causal Mapping Relationship and Network Construction of Railway Emergency

In order to lay a solid theoretical foundation for EDSS, the causal mapping relationships between the points of every two adjacent classes are defined. Then a model named CMRN is constructed, both of them are based on the knowledge system above.

58.3.1 Causal Mapping Relationship of Railway Emergency

With casual map [4], this section defines the mapping relationships of the points in railway emergency mechanism. To begin with, we give some new definitions:

Let e_i be a kind of railway emergencies that affect the rail transport operation, p_j be an abnormal phenomenon that different from the normal train operation, f_m be a kind of faults on railway dynamic equipment & static device and o_n be a dispatching order or disposal to prevent the train from risk or accident. Suppose *E, P, F, O* are four sets such that $e_i \in E$, $p_j \in P$, $f_m \in F$, $o_n \in O$ for all $i, j, m, n \in N^+$, and $E \cap P = \emptyset$, $P \cap F = \emptyset$, $F \cap O = \emptyset$.

Definition 1: If e_i, p_j are two given elements, and $e_i \in E$, $p_j \in P$. If there is a causal relationship between e_i and p_j, i.e. p_j is caused by e_i, then a mapping is defined between emergency e_i and abnormal phenomenon p_j, and is denoted by $e_i \mapsto p_j$.

Definition 2: If p_j, f_m are two given elements, and $p_j \in P$, $f_m \in F$. If there is a causal relationship between p_j and f_m, i.e. p_j leads to f_m, then a mapping is defined between abnormal phenomenon p_j and fault f_m, and is denoted by $p_j \mapsto f_m$.

Definition 3: If f_m, o_n are two given elements, and $f_m \in F$, $o_n \in O$. If there is a causal relationship between f_m and o_n, i.e. o_n is the necessary measure to be taken to ensure safety when f_m has occurred, then a mapping is defined between fault f_m and dispatching order o_n, and is denoted by $f_m \mapsto o_n$.

58.3.2 Causal Mapping Relationship Network Model of Railway Emergency

The railway emergency plan, historical disposal cases, and expert experience contain lots of information which we need to form the theoretical points. However, it cannot express the entire theoretical system to support the designing and developing of rail transport intelligent system. So, it is time to present the concept of Casual Mapping Relationship Network (CMRN):

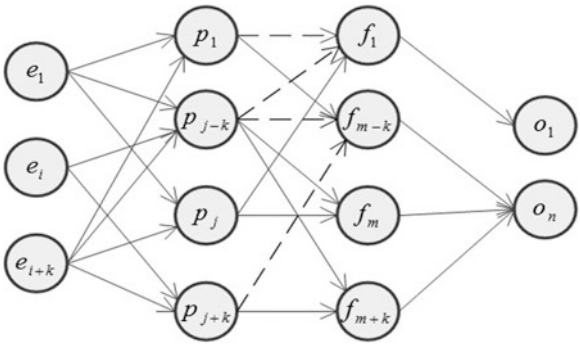


Fig. 58.1 The CMRN model

Definition 4: Given the elements e_i, p_j, f_m, o_n, if there are one-to-one mapping relationships existing between every two adjacent kind elements, i.e., $e_i \mapsto p_j$, $p_j \mapsto f_m$, $f_m \mapsto o_n$, then e_i, p_j, f_m, o_n create a chain of casual mapping relationship; A number of these chains form a network, which is the so-called *Casual Mapping Relationship Network* (CMRN) in this paper.

Sometimes the dispatching orders should be delivered before the equipment and device fault occurring. Though it has not caused any fault, it reaches the critical threshold of causing failure. In order to make the theory more fruitful and convincing, we give another concept—*the pre-treating abnormal phenomenon*:

Definition 5: If the abnormal phenomena caused by railway emergency have a threat on the dynamic equipment and static devices of rail transport operation, some abnormal disposal measures need to be taken to ensure the operation safety in order to avoid the any serious consequences. Then these abnormal phenomena are defined as the *pre-treating abnormal phenomenon*. Such as the strong winds, hard rain, heavy snow, etc. The casual mapping relationship chain denoted as $e_i \rightarrow p_j \cdots (f_m) \rightarrow o_n$, where the dotted line between p_j and (f_m) indicates that after p_j happened, o_n is taken to avoid happening of the potential risk (f_m) so as to ensure the reliability of railway equipment and device.

Now we can create the CMRN model with the definition above as shown in Fig. 58.1.

58.4 Process Design of Emergency Decision Support System

Rail transport operation system is a wide open dynamic system both in time and in space distribution. Many factors endanger the operation safety extremely. Thus, we design the process of EDSS for railway transportation.

Based on the model of CMRN and the principle of "Safety-Oriented," this section designs the process of EDSS. Once a sprout, the emergency is monitored by the system, and with the development, a series of constant early-warning

information is sent to dispatchers, urging them to prepare and prevent the potential hazards that may emerge, until it fades.

58.4.1 "Safety-Oriented" Principle

"*Safety-Oriented*," first proposed in "railway signal fault-safety principle" [5] by MOR, is a technical term of railway signal professional [6]. It refers that once the equipment or system is failure, this technology can guide the statement of equipment toward safety.

We definite three types of early-warning on the principle of "safety-oriented" to describe the impact of emergency on rail operation: Early-warning of emergency, abnormal phenomenon, and equipment fault. Early-warning of emergency aims to forewarn the abnormal phenomena; Early-warning of abnormal phenomenon may leads to equipment fault or potential risk, the phenomena may have led to the failure of devices and do harm to train operation. Early-warning of emergency and abnormal phenomenon is purpose to urge relative personals to reduce the possibility of these events. When the phenomenon pretreating failed, early-warning of equipment fault gives the dispatching order for dispatcher as decision support.

58.4.2 Process Design of EDSS with Petri Net

Petri net is the most suitable model for the description of business process model [7]. Therefore, we model the EDSS process with petri net and analyze performance of the model later.

The basic theory of petri net is easily seen in many references [8, 9], so this paper gives the petri net model (Seen Fig. 58.2). And Table 58.1 illustrates the meaning of each transit in the model.

58.4.3 Analysis the Petri Net Model of EDSS

Qualitative analysis and quantitative analysis are two kinds of methods for model analysis of system process at present. Due to the lack of a number of data, this paper takes the former one. With the incidence matrix and S-invariant, we study the performance of the petri net model for EDSS.

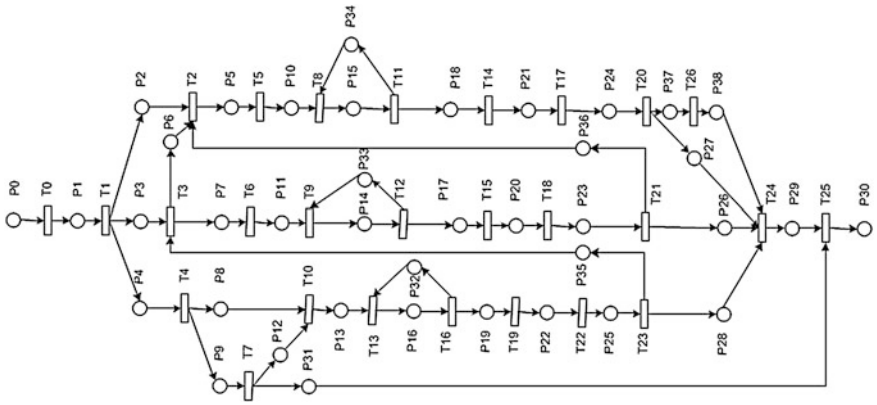


Fig. 58.2 The model of petri net for railway emergency decision support system

Table 58.1 Transit meaning of petri net model in Fig. 58.2

Transit	Meaning	Transit	Meaning
T_0	Collect the early-warning information	T_{14}	Deliver dispatching order
T_1	Classify the early-warning information	T_{15}	Deliver disposal scheme of abnormal phenomenon
T_2	Provide dispatching order reference	T_{16}	Approve disposal scheme of emergency
T_3	Judge pretreating abnormal phenomenon	T_{17}	Adjust stage operation schedule
T_4	Judge emergency quantifiably	T_{18}	Execute disposal scheme of abnormal phenomenon
T_5	Edit dispatching order	T_{19}	Deliver disposal scheme of emergency
T_6	Forewarning of abnormal phenomenon	T_{20}	Judge large-scale delay
T_7	Judge trigger threshold	T_{21}	Judge the pretreating result of abnormal phenomenon
T_8	Apply dispatching order	T_{22}	Execute disposal scheme of emergency
T_9	Edit disposal scheme of abnormal phenomenon	T_{23}	Judge the pretreating result of emergency
T_{10}	Forewarning of emergency	T_{24}	Generate historical case
T_{11}	Approve dispatching order	T_{25}	Finish the disposal process
T_{12}	Approve disposal scheme of abnormal phenomenon	T_{26}	Launch large-scale delay disposal
T_{13}	Edit disposal scheme of emergency		

According to the model in Fig. 58.2, we have the incidence matrix A.

$$A = \begin{pmatrix}
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1
\end{pmatrix}^T$$

Solving the equation:

$$A^T X = 0 \tag{58.1}$$

We can get the S-invariant of the petri net model in Fig. 58.2. As the petri net model has only input event: emergency (phenomenon or fault), so the token in the first place is 1. In each S-invariant, if the vector component value is 1, the token has reached the corresponding place and it does not while 0. Thus, some of the S-invariants can be translated in another pattern:

$$X_1 = (P_0, P_1, P_3, P_5, P_6, P_{10}, P_{15}, P_{18}, P_{21}, P_{24}, P_{27}, P_{29}, P_{30})^T;$$

$$X_2 = (P_0, P_1, P_4, P_9, P_{12}, P_{13}, P_{16}, P_{19}, P_{22}, P_{25}, P_{28}, P_{29}, P_{30})^T;$$

$$X_3 = (P_0, P_1, P_2, P_5, P_{10}, P_{15}, P_{18}, P_{21}, P_{24}, P_{27}, P_{29}, P_{30})^T;$$

$$X_4 = (P_0, P_1, P_4, P_8, P_{13}, P_{16}, P_{19}, P_{22}, P_{25}, P_{28}, P_{29}, P_{30})^T;$$

$$X_5 = (P_0, P_1, P_3, P_7, P_{11}, P_{14}, P_{16}, P_{17}, P_{20}, P_{23}, P_{26}, P_{29}, P_{30}, P_{32})^T;$$

$$X_6 = (P_0, P_1, P_4, P_7, P_8, P_{11}, P_{13}, P_{14}, P_{16}, P_{17}, P_{19}, P_{20}, P_{22}, P_{23}, P_{25}, P_{26}, P_{29}, P_{30}, P_{35})^T;$$

$$X_7 = (P_0, P_1, P_3, P_5, P_7, P_{10}, P_{11}, P_{14}, P_{15}, P_{17}, P_{18}, P_{20}, P_{21}, P_{23}, P_{24}, P_{27}, P_{29}, P_{30}, P_{36})^T;$$

$$X_8 = (P_0, P_1, P_3, P_5, P_7, P_{10}, P_{11}, P_{14}, P_{15}, P_{17}, P_{18}, P_{20}, P_{21}, P_{23}, P_{24}, P_{29}, P_{30}, P_{36}, P_{37}, P_{38})^T.$$

Therefore, we get the paths that token reaches. According to the basic theories of accessibility, activity and boundedness of petri net [9], the model of petri net and the process of EDSS is proved to be reasonable.

58.5 Conclusion

In this paper, we analyses the emergencies and make a further study in the effect mechanism. Casual mapping relationship is defined between the five classes of rail emergency and their direct or indirect impact. Then the CMRN model is proposed to rich the theoretical system of EDSS.

A new system process of emergency decision support based on the CMPN model is designed. Finally, petri net theory is used to check the reasonability of the process. By solving the equation $A^T X = 0$, we get the entire S-invariant and the paths that token reaches, proving the fact that: the process we design is reasonable.

Acknowledgments This work has been supported by National Key Technology Research and Development Program (2009BAG12A10), National Natural Science Foundation of China (Grant: 61074151), Research Fund of the State Key Laboratory of Rail Traffic Control and Safety (Grant: RCS2008ZZ003, RCS2009ZT002) and State 863 Projects: 2012AA112001.

References

1. Master of Railway PRC (2005) The overall planning of railway informationization. Railway Information, vol 4, pp 2–8 (in Chinese)
2. Li L, Wang F (2012) Railway incident and emergency decision-making research. J Inst Disaster Prev Sci Technol 14(3):58–60 (in Chinese)
3. Zhiqiao Zhang (2003) The studies on the risk of railway transportation disaster and forewarning management system[D]. WuHan University of Technology, Wuhan, pp 38–51 (in Chinese)
4. Wei Zheng, Xiaofeng Meng, Songjie Cao (2008) Fault diagnosis method based on causality mapping net for multi-sensor fusion [J]. J Syst Simul 21(20):5944–5946, 5952 (in Chinese)
5. China Railway Signal & Communication Corp. Standards Metrology Institute of MOR (1994) The People's Republic of China railway industry standard "railway signal fault-safety principle" TB/T. 2615–1994 (in Chinese)
6. Yujie Zhang (2008) Modeling and optimal methods research of the railway container based on the Petri nets[D]. Beijing Jiaotong University, Beijing, pp 35–45 (in Chinese)
7. Zhang L (2009) Workflow model based on stochastic petri nets and performance evalumion. In: IEEE international symposium on IT in medicine & education, vol 1. pp 245–249
8. Annappa B, Jiju P, Chandrasekaran K (2009) Petri net based verification of a cooperative workflow model. In: First international conference on networked digital technologies, pp 82–87
9. Chuang Lin (2004) Stochastic Petri net and performance evaluation of system, 2nd edn [M]. Tsinghua University Press, Beijing, pp 293–301

Chapter 59
Subway Vehicle Bearing Fault Diagnosis Methodology Research Based on PNN Neural Network and Wavelet Package

Xi Li, Yuan Zhang and Limin Jia

Abstract In order to diagnose different kinds of subway vehicle rolling bearing faults, a new method of fault diagnosis methodology based on improved wavelet packet and PNN (Probabilistic Neural Network) was put forward. Vibration signal of subway vehicle rolling bearing was collected by Piezoelectric Accelerometer. The collected signal was denoised by wavelet, and then decomposed by the improved wavelet packet, constructing the eigenvector. The signal was taken as fault samples to train the improved PNN neural network. The whole process finally recognizes fault types and realizes intelligent fault diagnosis. Test results show that the application of fault diagnosis method can effectively diagnose rolling bearing faults such as fatigue, peeling, and crack which occurred in inner ring, outer ring, and the rolling body surface during subway vehicle operation. The fault diagnosis method has high application value in the subway operation process.

Keywords Fault diagnosis · Wavelet analysis · PNN neural network · Subway bearing

59.1 Introduction

With the rapid development of subways throughout the world, safety becomes the most important problem in subway operation [1, 2]. Vehicle safety and early warning are gaining more and more attention for all kinds of researchers. In the

X. Li (✉) · L. Jia
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
e-mail: bjtulx@163.com

X. Li · Y. Zhang
Subway Operation Technology Centre, Mass Transit Railway Operation Corporation LTD, Beijing, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_59,
© Springer-Verlag Berlin Heidelberg 2014

actual operation process of the subway, equipment failure caused by vibrations such as frame crack, equipment fatigue, and bearing failure are potential hazards in the operation process. It is urgent to find appropriate methods to solve these problems and ensure the operation safety. Signal processing is an effective method used in subway device safety evaluation and early warning. There are many methods in signal processing, such as traditional Fast Fourier Transform, Short Time Fourier Transform, and so on [3–5]. But all these methods have some disadvantages, e.g., spectrum leakage, and cannot carry out time and frequency part window analysis at the same time. Over the years, the wavelet transform has attracted the increasing attention of electronic engineers because of its excellent achievement of time–frequency localization; so the harmonic detection with the wavelet transform becomes a very suitable method for the detection and time–frequency analysis of nonstationary signal. However, the segmentation of spectrum based on wavelet multiresolution analysis is asymmetric, causes inconformity of harmonic quantity in different frequency bands, and makes it different for harmonic measurement in the same frequency band. Therefore, we make use of wavelet packet and Probabilistic Neural Network (PNN) neural network to analyze vibration signals; the wavelet package transform is suitable if the information in both high and low frequency bands can be finely identified, and it will be an ideal processing tool to the nonstationary time-variable signal. The software of the system was developed on the MATLAB platform, which offers an effective way to apply wavelet transforms to measure instrument.

59.2 Research Methodology

59.2.1 Wavelet Package Analysis

Wavelet analysis includes continuous and discrete wavelet transform. It is a powerful tool for nonlinear and unsteady signal processing. Wavelet packet analysis is an extended method with more detailed analysis and reconstruction out of wavelet analysis for signal processing. It decomposes the frequency band into multiple layers, further decomposes the high frequency signal, is adaptive to choose corresponding frequency band, matches the signal spectrum according to the analysis of signal characteristics, and improves the time–frequency resolution [6, 7].

In this paper, we use the three-layer wavelet packet decomposition method, the structure is shown in Fig. 59.1. Wavelet packet decomposition deposes the signal into low frequency part a_1 and high frequency part d_1. In the next layer of decomposition, both the parts a_1 and d_1 are decomposed into low frequency and high frequency parts aa_2, da_2, ad_2, dd_2, and so on the signal can be further decomposed. Therefore, wavelet packet decomposition is a finer decomposition method.

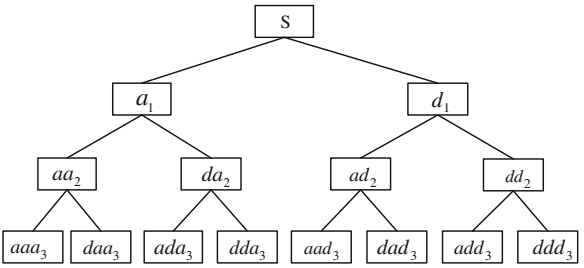


Fig. 59.1 Signal wavelet packet decomposition schemes

Define V_j Dimension subspace, W_j Wavelet subspace, U_j^n New subspace [8].

$$\begin{cases} U_j^0 = V_j \\ U_j^1 = W_j \quad j \in Z \end{cases}$$

And then, orthogonal decomposition of Hilbert space $V_{j+1} = V_j \oplus W_j$ can be unified by decomposition of U_j^n

$$U_{j+1}^0 = U_j^0 \oplus U_j^1 \quad j \in Z$$

Define subspace U_j^n is closure space of the function of, $u_n(t)$ space, and U_j^{2n} is the closure space of function $u_{2n}(t)$, $u_n(t)$ meet the dual dimension equation below:

$$\begin{cases} u_{2n}(t) \quad = \sqrt{2} \sum\limits_{k \in Z} h(k) u_n(2t-k) \\ u_{2n+1}(t) = \sqrt{2} \sum\limits_{k \in Z} g(k) u_n(2t-k) \end{cases}$$

In the formula, $g(k) = (-1)^k h(1-k)$, when $n = 0$:

$$\begin{cases} u_0(t) = \sum\limits_{k \in Z} h_k u_0(2t-k) \\ u_1(t) = \sum\limits_{k \in Z} g_k u_0(2t-k) \end{cases}$$

In the analysis, $\phi(t)$ and $\psi(t)$ meet the dual dimension equation

$$\begin{cases} \phi(t) = \sum\limits_{k \in Z} h_k \phi(2t-k) \cdots\cdots\cdots \{h_k\}_{k \in Z} \in l^2 \\ \psi(t) = \sum\limits_{k \in Z} g_k \phi(2t-k) \cdots\cdots\cdots \{g_k\}_{k \in Z} \in l^2 \end{cases}$$

$$U_{j+1}^n = U_j^n \oplus U_{j+1}^{2n+1} \cdots\cdots j \in Z \quad n \in Z_+$$

Wavelet packet decomposition and reconstruction algorithm: define $g_j^n(t) \in u_j^n$, $g_j^n(t)$ can be represented as

$$g_j^n(t) = \sum_l d_l^{j-n} u_n\left(2^j t - l\right)$$

Wavelet packet decomposition algorithm:

$$\begin{cases} d_l^{j,2n} = \sum_k a_{k-2l} d_k^{j+1,n} \\ d_l^{j,2n+1} = \sum_k b_{k-2l} d_k^{j+1,n} \end{cases}$$

Wavelet packet reconstruction algorithm:

$$d_l^{j+1,n} = \sum_k \left[h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+1} \right]$$

59.2.2 PNN Neural Network

The topological structure of the PNN is composed of input layer, model layer, accumulation layer, and output layer (decision layer), as shown in Fig. 59.2.

The input layer receives data from the sample value. The input pattern layer neurons in each component vectors are weighted using Gauss function operation. Accumulation layer node only connects corresponding category of sample nodes, calculates probability density function, so as to obtain maximum likelihood input sample belongs to a class. The output layer is model posterior probability estimation. Neuron of this layer is a competitive neuron.

Learning process of PNN is according to the distribution of input samples, determine center value, and width (smoothing factor) of model layer basis function, mapping complete sample to model categories to get the optimal estimate.

Gauss function center value choice mainly affects the performance of PNN network, if the center distance is too close it will cause the network to produce the approximate linear correlation and cannot get correct classification; if center distance is too far it will lead to the classification to be too rough. The common center value selection method includes random choice method, K means clustering method, mixed normal distribution model method, and orthogonal least square method.

Smoothing factor is another major factor influencing the performance of PNN network. If the value is too large, density estimation is smooth but details are lost; when the value is too small, density estimation will show more spikes protruding. Generally, smoothing factor method is to choose a fixed value in the data sample space.

59.3 Experiment Data

Through experiment analysis of variable fault bearing in subway operation and maintenance, part of experiment data is shown in Table 59.1.

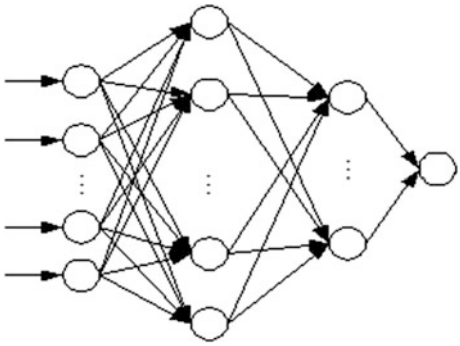


Fig. 59.2 Topology of PNN

Table 59.1 Part of experiment data

No fault	Inner ring fault	Outer ring fault	Roller fault
0.23406982421875	−0.02838134765625	−0.02044677734375	−0.0823974609375
0.1239013671875	−0.0177001953125	0.0152587890625	0.03204345703125
0.079345703125	−0.0201416015625	0.0042724609375	0.0390625
0.00823974609375	−0.02716064453125	−0.03753662109375	0.00244140625
0.015869140625	−0.0164794921875	−0.0347900390625	−0.029296875
−0.0103759765625	0.02197265625	−0.0042724609375	−0.030517578125
0.02166748046875	0.0244140625	−0.01983642578125	0.02105712890625
0.03662109375	0.01129150390625	0.00732421875	−0.009765625
0.17608642578125	0.00762939453125	0.03753662109375	0.00457763671875
0.29083251953125	0.03631591796875	0.05584716796875	0.00640869140625
0.39459228515625	0.0775146484375	−0.00091552734375	0.00335693359375
0.4217529296875	−0.04425048828125	−0.0091552734375	0.03082275390625
0.54412841796875	−0.00946044921875	−0.00335693359375	0.0250244140625
0.60211181640625	0.01678466796875	0.01983642578125	0.02655029296875
0.64483642578125	−0.0341796875	0.0311279296875	0.01861572265625
0.63934326171875	−0.01678466796875	−0.0091552734375	0.01220703125
0.64849853515625	0.02227783203125	0.00701904296875	0.0408935546875
0.526123046875	0.00701904296875	−0.0732421875	−0.0042724609375
0.41351318359375	−0.013427734375	−0.01312255859375	−0.068359375
0.31951904296875	0.02471923828125	−0.00732421875	0.03509521484375
0.3155517578125	−0.00396728515625	0.00579833984375	0.0872802734375
0.2886962890625	−0.02838134765625	−0.0152587890625	0.069580078125
0.23406982421875	−0.0177001953125	−0.02716064453125	−0.0347900390625
0.1239013671875	−0.0201416015625	−0.02044677734375	−0.0823974609375
0.079345703125	−0.02716064453125	0.0152587890625	0.03204345703125
0.00823974609375	−0.0164794921875	0.0042724609375	0.0390625
0.015869140625	0.02197265625	−0.03753662109375	0.00244140625
0.01007080078125	0.0244140625	−0.0347900390625	−0.030517578125
0.03570556640625	0.01129150390625	−0.0042724609375	0.02105712890625
0.02349853515625	0.00762939453125	0.0048828125	−0.009765625
0.0579833984375	0.00579833984375	0.00732421875	0.00457763671875
−0.0048828125	−0.01617431640625	−0.00274658203125	−0.001220703125

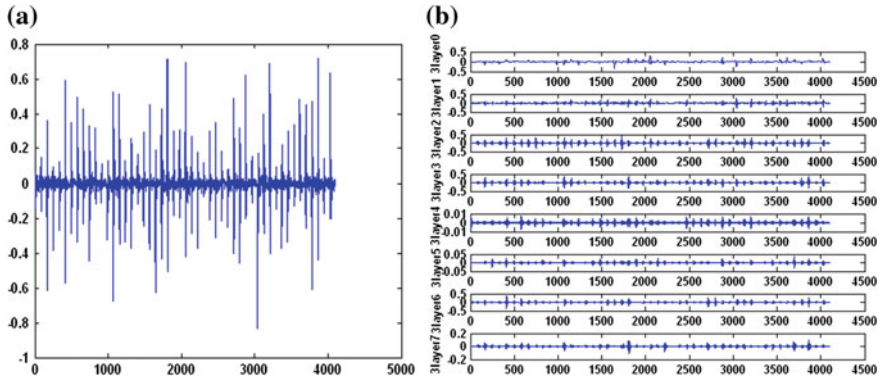


Fig. 59.3 **a** Roller fault time domain waveform. **b** Reconstruction waveform after roller fault three layers wavelet packet decomposition

59.4 Experiment Results Analysis

Time–domain waveform after FIR noise reduction to no fault [9], roller fault, outer ring fault is shown in Figs. 59.3a, 59.4a and 59.5a.

Figures 59.3b, 59.4b and 59.5b have three layers db1 orthogonal wavelet packet decomposition result.

59.4.1 Preparing Training Network

After three layers wavelet packet decomposed the data are collected through MATLAB; we reconstructed the third layer decomposed data, evaluated the frequency band energy according to the former formula, and normalized to the training samples as shown in Table 59.2.

59.4.2 Test Data

The test data was constructed with the same approach as collected test data, as shown in Table 59.3.

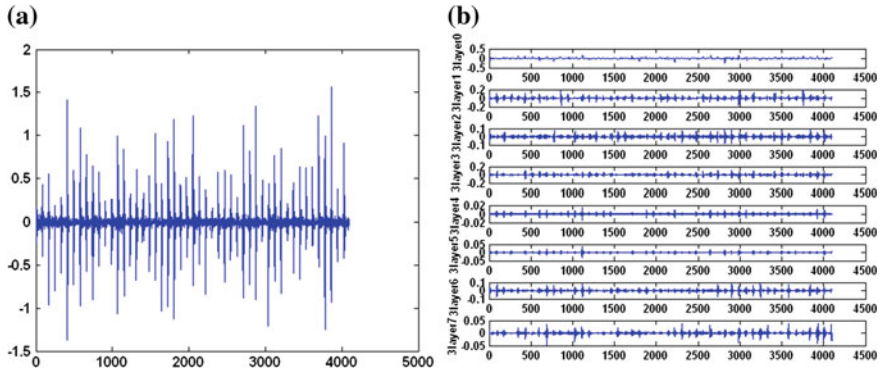


Fig. 59.4 **a** Inner ring fault time domain waveform. **b** Reconstruction waveform after Inner ring fault three layers wavelet packet decomposition

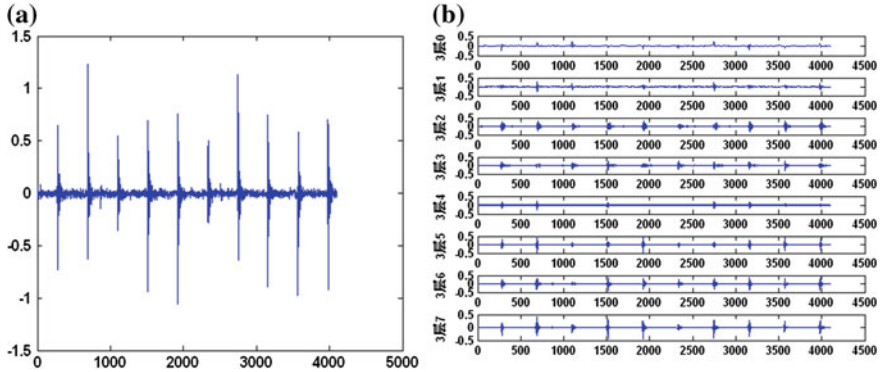


Fig. 59.5 **a** Outer ring fault time domain waveform. **b** Reconstruction waveform after outer ring fault three layers wavelet packet decomposition

59.4.3 Create PNN Network

Using newpnn () function in the Matlab neural network toolbox to design PNN network, net = newpnn (p, t, spread), p is input vector, t is target vector, spread is the density distribution of the radial basis function, here set as 0.1.

Neurons of input layer is 8, neurons of model layer is 4, neurons of accumulation layer, and output layer is 4, which corresponds to the four kinds of failure modes (1 for no fault). Using SIM () function simulation, the output vec2ind () function is converted into a vector form.

According to the training network, output result is yc_test = 1, 2, 3, 4. From the result, we can conclude that this method can effectively identify the fault type and characteristics. We can completely achieve the intended purpose of identifying problems through this methodology.

Table 59.2 Bearing operation sample data

Data Number	Sample								Fault state
1	0.9788	0.0078	0.0010	0.0087	0.0016	0.0006	0.0008	0.0007	No fault
2	0.0467	0.0997	0.0153	0.0655	0.6610	0.0051	0.0292	0.0775	Outer ring fault
3	0.0690	0.0677	0.1777	0.0112	0.0568	0.0701	0.4876	0.0599	Inner ring fault
4	0.0531	0.0717	0.8778	0.1259	0.0241	0.0363	0.2456	0.0655	Roller fault

Table 59.3 Test data

Data Number	Sample								Fault state
1	0.9722	0.0093	0.0042	0.0102	0.0004	0.0005	0.0024	0.0008	No fault
2	0.0937	0.0744	0.1225	0.0927	0.4234	0.0498	0.0914	0.0521	Outer ring fault
3	0.0226	0.0268	0.2782	0.0500	0.0450	0.0187	0.5361	0.0226	Inner ring fault
4	0.1841	0.1038	0.5236	0.0600	0.2669	0.0834	0.1039	0.0743	Roller fault

59.5 Conclusion

Aiming at the characteristics of subway vehicle running system, this paper proposes to acquire vibration signals of subway vehicles running gear bearing through the sensors, combines wavelet analysis with PNN neural network to analyze the vibration velocity or acceleration signals for feature extraction, identifies fault types, realizes the purpose of safety assessment of running gear bearing, and achieves remarkable results by actual tests. The method can be widely used in the safety assessment and fault diagnosis of rotating machinery such as bearing.

References

1. Li X, Zhu XN, Cai GQ (2010) Research on system integration alliance of urban rail transit safety monitoring. J Convergence Inf Technol 10:36–41
2. Li X, Zhu XN, Cai GQ (2010) A novel fault diagnosis expert system knowledge acquisition method of metro vehicle equipments. In: Proceeding of 2010 international conference on computer and automation engineering (in Chinese)
3. Nikolaou NG, Antoniadis IA (2002) Rolling element bearing fault diagnosis using wavelet packets. NDT and E Int 35(3):197–205
4. Guo QJ, Yu HB, Xu AD (2005) Modified moret wavelet neural networks for fault detection. In: International conference on control and automation, pp 1209–1214 (in Chinese)
5. Lin F-J, Shieh H-J, Huang P-K (2006) Adaptive wavelet neural network control with hysteresis estimation for piezo-positioning mechanism. IEEE Trans Neural Networks 17(2):432–444
6. Pang PL, Ding GB (2008) Wavelet-based diagnostic model for rotating machinery subject to vibration monitoring. In: Proceedings of the 27th Chinese control conference, 2008, pp 303–306

7. Liu HM, Wang SP, Ouyang PC (2009) Fault diagnosis based on improved elman neural network for a hydraulic servo system. Beijing Jiaotong University Press, Beijing (in Chinese)
8. Huang NE, Shen SP (eds) (2005) Hilbert-Huang transform and its applications. World Scientific, Singapore
9. Ramos G, López JJ, Pueo B (2009) Cascaded warped-FIR and FIR filter structure for loudspeaker equalization with low computational cost requirements. Digital Signal Process 19:393–409

Chapter 60
Study on the Best Transfer Station Selected in Beijing Subway Network

Qi Liu, Xingchen Zhang, Junhua Chen and Bin Xu

Abstract Transfer station is the most important node of urban rail network, affecting transfer efficiency and overall efficiency of the network as a whole. Based on the investigation and statistics of the transfer time and train running time between transfer stations, this paper first finds the shortest path between any transfer stations with Floyd algorithm. Then through analysis of transfer mode and passengers travel behavior, we can figure out the best transfer station for passengers, and take Beijing subway network as an example to validate the model and algorithm.

Keywords Transfer station · Transfer mode · Floyd algorithm · Shortest path · Traveler behavior

60.1 Introduction

With the increase in people's travel demands, the public transportation plays an increasingly important role in urban traffic. Especially, rail transit has developed rapidly in the recent years all over the country. As an "anchorage point" of the rail transit network, the transfer station plays an important role on the elaboration of the network running efficiency and the overall benefit. The transfer station has become the most prominent contradictions in network, from the actual operating status of Beijing, Shanghai, and other cities [1].

The research on the rail transit transfer has the following main aspects: (1) Research on the different transfer modes at the transfer station which analyzes their advantages and disadvantages [2]; (2) Research on the passenger flow behavior

Q. Liu (✉) · X. Zhang · J. Chen · B. Xu
Beijing Jiaotong University, Beijing, China
e-mail: 812079520@qq.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*,
Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_60,
© Springer-Verlag Berlin Heidelberg 2014

organization, which analyzes the characteristics of transfer station passenger flow [3, 4]; (3) Research on the connection between rail transit and other modes of travel [5]; (4) Summary of the law of walking time of passengers in the subway transfer channel, which analyzes the influencing factors of the passenger travel time [6].

Experts and scholars at home and abroad have performed a lot of comprehensive research on subway interchange and passenger travel choice. However, these aspects were rarely combined together for a systematic study which can achieve the simplification and fast turn for subway passengers. This paper intends to make the entire subway network and the surrounding environment facilities as a whole to establish the best transfer station mode. In the end the Beijing subway network is taken as an example to study the passengers' selection of the best travel route.

60.2 Analysis of Subway Transfer Mode and Passenger Travel Behavior Selection

60.2.1 Subway Transfer Mode

Because of the difference in the structure of transfer station, it has formatted different transfer modes. Generally, we divided it into several types such as one-platform-interchange, node transfer, channel transfer, composite transfer, and so on.

60.2.2 Passenger Travel Behavior Selection

Travel behavior selection is a traffic travelers' comprehensive reflection of the psychological and behavioral in particular traffic conditions. There are many complicated factors that can impact the traffic travelers' travel choice. Therefore, when choosing his travel routes, a traveler is bound to make decisions depending on his travel experience and the new acquired traffic information to choose the route whose travel time or cost is minimum. The effects on passenger travel behavior selection are: (1) The impact of age on the path selection; (2) The impact of profession on the path selection; (3) The impact of travel purpose on the path selection; (4) The impact of travel period on the path selection; and (5) Different travel distance has a certain influence on the path chosen by the passenger.

60.3 Model Constructions

60.3.1 The Shortest Path Model Between the Transfer Stations

We can put the transfer stations as nodes and the lines as arcs. Thus, we can transform subway network into network topology and we can use Floyd algorithm to solve the shortest path.

In the graph $G = (N, A)$, d_{rs} stands for the path between origin vertices r and destination s with the smallest resistance value. Matrix $D = \{d_{rs}\}$ consists of the smallest estimated on the resistance value of the path, while $D_n = \{d_{ns}\}$ the smallest resistance matrix in the n iteration. m_{rs} is the intermediate vertices from origin r to destination s in the shortest path between them. $D = \{m_{rs}\}$ is the path matrix.

The specific steps of Floyd algorithm are as follows:

Step 1: Initiate the distance matrix that shows the distance between vertices.

Step 2: Locate all the elements in the distance matrix in row n and column$n$. They are defined as the main row elements and the main column elements

Step 3: Compare the value of every element that is not the main row elements or the main column elements to the sum value of the main elements in the same row and column:

(1) If $d_m + d_{ns} \geq d_{rs}$, choose a new m_{rs}, go to Step 3 and start the next comparison process;

(2) If $d_m + d_{ns} < d_{rs}$, then $d_{rs} = d_m + d_{ns}$, replace the corresponding element in the shortest path matrix.

If all the intermediate vertices are checked, go to Step 2, $n = n + 1$ and continue to iterate until $n = N$ so that the smallest resistance matrix and the shortest path matrix can be finally settled.

60.3.2 The Shortest Path Model Between the Non-Transfer Stations

In fact, travel starting point or destination is not usually the transfer station. The choice of the shortest path between nontransfer stations is based on the shortest path model between the transfer stations, considering the influence factors of passenger transfer behavior. Then we can get the best transfer station for our travels. For example, older passengers in the selection of transfer stations will

consider less transfer times and the shortest transfer time. But the young will focus more on the shortest total travel time, and not care about the transfer times. So the travelers choose their travel paths according to their own situation.

60.4 Computational Results

In this section, we present computational results on the Beijing metro network. We implemented our algorithms in Mat lab Programming and tested them on the data by survey on Beijing metro network.

60.4.1 Classification of Transfer Station

According to the characteristics of each of the transfer stations, all the transfer stations can be divided into the following categories:

(1) The transfer station connecting two lines has played the role of the initial station or terminal. As a result of not affecting the selection of transfer station for travelers, the kind of station can be regarded as non-transfer station. Such transfer stations include Gongyi West Bridge station, etc.
(2) The transfer station connecting two lines has played the role of the starting station (terminal) in one line and the role of the intermediate station in the other line as the starting station (terminus) had no effect on the selection of travelers. Therefore, in the process of solving, we only consider its impact on the other line. Such transfer stations include Xierqi station, etc.
(3) The transfer stations connecting two or three lines all play the role of the intermediate station. As a result of the kind of station having effect on the selection of travelers for any line, the kind of station needs to be taken into account in the process of solving. Such transfer stations include Xizhimen station, etc.

60.4.2 Survey Data

(1) The train running time is between transfer stations.
According to the above classification of the transfer station, we know the first kind of the transfer station does not affect the selection of travelers. So such transfer stations are not involved in this study. Through the investigation we can receive the train running time between transfer stations, as shown in Fig. 60.1.

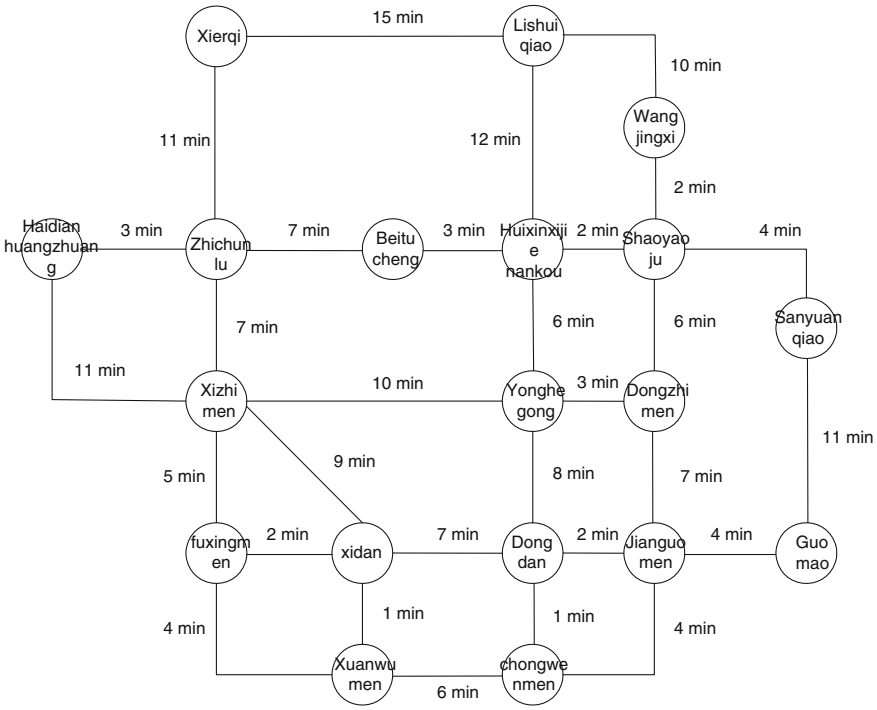


Fig. 60.1 Train running time between adjacent transfer stations

(2) The transfer time
According to the classification of the transfer station in 4.1, we know the second kind of transfer station affects only one line. So we will not consider transfer time of the kind of transfer station, and only consider transfer time of the third types of transfer station. The transfer time for every transfer station is shown in Table 60.1.

60.4.3 Analysis of the Results

(1) Solving the shortest path
We can solve the shortest path between transfer stations by using the method of MATLAB programming. The detailed results are shown in Table 60.2.
(2) The best choice of transfer station for travelers
For transferring, passengers not only consider the shortest transfer time, but also the transfer times and comfort. According to Tables 60.2 and 60.3, passengers can confirm to select the best transfer station.

Table 60.1 Transfer time in every transfer station (partial)

Transfer station	Transfer line	Transfer time (min)
Xizhimen	Line 4–Line 2	3
Xizhimen	Line 2–Line 4	1
Xizhimen	Line 13–Line 4	5
Yonghegong	Line 2–Line 5	2
Yonghegong	Line 5–Line 2	3
Dongzhimen	Line 2–Line 13	7
…	…	…

Table 60.2 Table of shortest route between transfer stations in Beijing subway network (partial)

The starting station	The terminal	Path
Dongzhimen (13)[a]	Jianguomen (2)	Dongzhimen (13) <Transfer> Dongzhimen (2) <Line 2> Jianguomen (2)
Dongzhimen (13)	Jianguomen (1)	Dongzhimen (13) <Transfer> Dongzhimen (2) <Line 2> Jianguomen (2) <Transfer> Jianguomen (1)
Dongzhimen (2)	Jianguomen (1)	Dongzhimen (2) <Line 2> Jianguomen (2) <Transfer> Jianguomen (1)
Dongzhimen (2)	Jianguomen (10)	Dongzhimen (2) <Line 2> Jianguomen (2) <Transfer> Jianguomen (1) <Line 1> Guomao (1) <Transfer> Guomao (10)
Dongzhimen (2)	Jianguomen (1)	Dongzhimen (2) <Line 2> Jianguomen (2) <Transfer> Jianguomen (1) <Line 1> Guomao (1)
Shaoyaoju (13)	Jianguomen (10)	Shaoyaoju (13) <Transfer> Shaoyaoju (10) <Line 10> Guomao (10)
Shaoyaoju (13)	Jianguomen (1)	Shaoyaoju (13) <Transfer> Shaoyaoju (10) <Line 10> Guomao (10) <Transfer> Guomao (1)
…	…	…

[a] Dongzhimen (13) indicates Dongzhimen station, Line 13

Table 60.3 Table of transfer mode and feature in Beijing subway network (partial)

Transfer station	Transfer path	Transfer mode	Crowdedness	Transfer feature
Dongzhimen	Line 2–Line 13	Tunnel	Low	Very long transfer distance, complicated transfer path, low efficiency
Shaoyaoju	Line 13–Line 10	Tunnel	Moderate	Moderate transfer distance, high crowdedness in peak hours
Guomao	Line 10–Line 1	Tunnel	High	Long transfer distance, extremely crowed
Jianguomen	Line 1–Line 2	Tunnel	Moderate	Short transfer path,moderately crowded in peak hours
…	…	…	…	…

For example, if one passenger wants to go to Yong'anli station (non-transfer station) in Line 1 from Liufang station (nontransfer station) in Line 13, there are two options:

Plan 1: Transfer at Donzhimen station and Jianguomen station;
Plan 2: Transfer at Shaoyaoju station and Guomao station

According to Table 60.2, we know that plan 1 is the best. But we find that the transfer time in Dongzhimen station from Line 13 to Line 2 is longer, and the transfer mode is more complex in Dongzhimen station. So the passenger maybe selects Plan 2.

60.5 Conclusions

This paper first analyzes the characteristics of subway transfer mode and the factors affecting passenger travel behavior. Then based on them, the paper builds the optimal transfer site selection model. Finally, Beijing subway network is taken as an example to validate the model and algorithm.

60.6 Acknowledgements

This work is partially supported by the National Basic Research Program of China (2012CB725406) and Beijing Municipal Education Commission Scientific Research and Graduate Education Projects (No. T11H100010).

References

1. Wang B, Li X, An Q (2010) Passenger flow characteristic and station design of metro transfer station. Urban Rapid Rail Transit 23(2):55–58 (in Chinese)
2. Wu M (2009) The transfer mode of subway station. Popular Sci (7):92–94 (in Chinese)
3. Song Y, Zheng X (2009) The research on passenger flow organization of metro transfer station. China New Technol Prod (12):87–88 (in Chinese)
4. Kang H, Zheng S, Wu Q (2009) The research on passenger flow organization of metro transfer station. Railway Transp Econ 31(8):89–91 (in Chinese)
5. Wang Y, Zhang X, Yuan Y (2008) Commuter transfer travel mode choice analysis of the city zone around metros terminal. Urban Railway 22(6):39–41 (in Chinese)
6. Du P, Liu C, Liu Z (2009) Walking time modeling on transfer pedestrians in subway passages. J Transp Syst Eng Inf Technol 9(4):103–109 (in Chinese)

Chapter 61
A New Approach to Diagnose Rolling Bearing Faults Based on AFD

Yu Liang, Li min Jia, Guo qiang Cai and Jin zhao Liu

Abstract A new fault diagnosis method of rolling bearing based on adaptive Fourier decomposition (AFD) is proposed. The new approach extracts the meaningful bearing vibration signal based on AFD algorithm instead of traditional band-pass filter; AFD decomposes the original bearing vibration signal into a series of mono-components, the kurtosis of each mono-component is calculated and clustered into two classes by fuzzy C-mean clustering (FCM). The mean of the two cluster centers is taken as threshold and the mono-components with large kurtosis is summed as bearing fault carrier signal because the bearing fault is sensitive to kurtosis; Then demodulated resonance technique is used to diagnose and locate the fault. The new approach can diagnose all kinds of rolling bearings' fault. Finally, the proposed approach is used to analysis the outer ring fault in case of N205EM type rolling bearing; the experiments indicates that the effectiveness and accuracy are significantly approved.

Keywords Fault diagnosis · Adaptive Fourier decomposition · Mono-component · Resonant demodulation

61.1 Introduction

Rolling bearings are essential parts of a rail transit train, and rolling element bearing failure is one foremost cause of rotating machinery breakdown. The importance of early bearing defects detection has led to continuous efforts to

Y. Liang · L. m. Jia (✉) · G. q. Cai
Beijing Jiaotong University, Beijing 100044, China
e-mail: jialm@vip.sina.com

J. z. Liu
China Academy of Railway Sciences, Beijing 100044, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_61,
© Springer-Verlag Berlin Heidelberg 2014

prevent the unpredictable disastrous failure. Now vibration-based condition monitoring and fault diagnosis has been the most widely utilized technique in the rolling bearing diagnosis [1].

Wavelet transform is an extensively used signal processing tool in bearing fault diagnosis [2]. Dilation and translation are basic operation of wavelet transform, which can effectively extract both time and frequency features of a signal. So, it is capable of analyzing nonstationary signals and it is suitable for vibration-based bearing fault diagnosis. But wavelet transform has many deficiencies in its application [3]. Huang et al. proposed a signal processing method, empirical mode decomposition (EMD) [4], with which the nonlinear or nonstationary signal could be decomposed into a small number of intrinsic mode functions (IMFs). IMFs determine all the instantaneous frequencies in the signal, and instantaneous frequency is an important characteristic during signal processing. So, EMD has been widely used in fault diagnosis [5]. The EMD also has deficiencies in its application. Time consuming is one of its disadvantages. Moreover, some IMFs have negative instantaneous frequency, which has no physical meaning. Demodulated resonance technique is the most commonly used methods on rolling bear fault diagnosis; it can diagnose the early tiny fault and determine the fault location through analyzing vibration signal. Wenyi wang used the resonant demodulation technology to diagnose the gear box fault, extracted fault feature, judged fault position, etc. [6]. Deyao Tang used resonant modulation in railway locomotive fault diagnosis; excited response impact fault by electronic resonator and extracts rolling bearing fault characteristics [7]. Although demodulated resonance technique has been widely used in bearing fault diagnosis, it has limitations. It needs to fix the filter bandwidth, the fault characteristic cannot be exacted if the inadequate filter bandwidth is fixed; different types of rolling bears have different bandwidths, and one filter bandwidth cannot be universal.

Tao Qian et al. suggested decomposing a nonlinear or nonstationary signal into mono-components with AFD [8–10]. With these mono-components, the global property of the signal can be revealed satisfactorily, i.e., a non-negative analytic instantaneous frequency function can be a global representative of conformal mappings. In this paper, an alternative approach of using AFD to the bearing fault diagnosis is proposed. With this approach, the bearing signal is decomposed into mono-components according its characteristic first, and then the mono-components are clustered into two classes according to their kurtosis. Those mono-components with kurtosis above the mean of the two class centers are extracted, and the sum of the extracted signals is preceded through envelope demodulation and transformed with FFT finally. Afterward, the bearing fault can be identified and located by using the spectrum diagram.

61.2 The Basics of AFD Theory

61.2.1 The Fault Characteristic Frequency of Rolling Bearing

The bearing failure manifests itself as pitting, spalling, flaking, or other forms. If a bearing has fault and keeps on running, the damage will spread since the localized stresses in the vicinity of the defect are increased. The motion of the rolling elements is disturbed by the damage, and it leads to the generation of short time impacts repeating at a corresponding frequency. The frequency is determined by the size of the bearing surface features, the sliding, and rolling velocities. The resulting vibration from the asperity interactions consists of random sequences of small impulses which excite all natural modes of the bearing and supporting structure. The natural frequencies which correlate with mean impulse rise time or mean interval between the impulses will be more strongly excited than the others. So bearing fault diagnosis is proceeded by taking envelop of the bearing vibration signal and calculating out the envelope frequency, which is called the bearing fault characteristic frequency. The formulas of the fault characteristic frequency are given as formula (61.1)–(61.3).

$$f_{\text{out}} = \frac{D-d}{2D} \times Z \times f_n \tag{61.1}$$

$$f_{\text{in}} = \frac{D+d}{2D} \times Z \times f_n \tag{61.2}$$

$$f_d = \frac{D^2 - d^2 \cos^2 A}{2D \times d} \times f_n \tag{61.3}$$

where, f_{out} is outer ring characteristic frequency; f_{in} is inner ring characteristic frequency; f_d is roller characteristic frequency; D is the pitch diameter; d is the roller diameter; A is the contact angle of roller; Z is the roller number; f_n is bearing inner & outer relative rotation speed frequency.

61.2.2 AFD Theory

AFD is a new method for analyzing nonlinear and nonstationary signal. AFD can decompose a complicated data into a finite number of mono-components which admit well-behaved Hilbert transforms. The decomposition method is adaptive and efficient since the decomposition is based on the local characters of the data. With the Hilbert transform, the mono-components yield instantaneous frequencies as functions of time that give sharp identifications of imbedded structures. In AFD, the main conceptual innovation is the introduction of “mono-component” based

on local properties of the signal, which makes the instantaneous frequency meaningful.

Instantaneous amplitude and instantaneous frequency are two important transient characteristics of a signal. If s is signal with finite-energy, then its corresponding analytic signal is $s_1 = s + iH(s) = \rho e^{i\phi}$, where H is Hilbert transform, ρ is the amplitude function and ϕ is the phase function of s_1, then the instantaneous frequency is the derive of ϕ. The instantaneous frequency ϕ' have clear physical meaning when it is a positive number, however by some other method, such as EMD, its instantaneous frequency may be negative number and has less physical sense.

A real-valued function $f = \rho\cos\theta$ is called the mentioned mono-component, if it satisfied formula (61.4).

$$H(\rho\cos\theta) = \rho\sin\theta \tag{61.4}$$

which, $\rho \geq 0$, $\theta' \geq 0$, $f \in L^p(R), 1 \leq p \leq \infty$.

H is Hilbert transform, ρ is the polar radius of real-valued function, θ is the polar angle of real-valued function, and $L^p(R)$ is integrable L^p functions space on the field of real number.

61.2.3 AFD Algorithm

The original signal is G, then its mono-components $f_1, f_2, \ldots, f_n$ has the general form

$$f_k = \left\langle G_k, e_{\{a_k\}} \right\rangle B_k(z), \quad k = 1, 2, \ldots, n \tag{61.5}$$

where, $G_k (k = 1, 2, \ldots, n)$ is a variable which appears in the process and $G_1 = G$;

$$e_{\{a_k\}} = \frac{\sqrt{1 - |a_k|^2}}{1 - \bar{a}_k z}$$

$$B_k(z) = B_{\{a_1,\ldots,a_k\}}(z) := \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1 - |a_k|^2}}{1 - \bar{a}_k(z)} \prod_{l=1}^{k-1} \frac{z - a_l}{1 - \bar{a}_l(z)} \tag{61.6}$$

where $a_l \in D$, $l = 1, 2 \ldots n, D = \{z \in C : |z| < 1\}$, C is complex plane.

Given $a_1 = 0$, then $e_{\{a_1\}} = 1$, $B_1(z) = \frac{1}{\sqrt{2\pi}}$, so

$$f_1 = \left\langle G_1, e_{\{a_1\}} \right\rangle B_1(z) \tag{61.7}$$

(1) Calculating f_2

$a_1 = 0$, $G_1 = G$, calculating G_2 first,

$$G_2 = \left(G_1 - b_1 * e_{\{a_1\}}\right) \frac{1 - \bar{a}_1(z)}{z - a_1} \tag{61.8}$$

where $b_1 = \langle G_1, e_{\{a_1\}} \rangle$.

Then a_2 is calculated by: G_2. The unit disk of complex plane is discrete into a series of vectors $D_n (n = 1, 2, \ldots 7705)$, and a_2 is one of D_n selected by maximal projection principle.

n_2 is the solution of formula (61.9),

$$\max_n \left\{ < \frac{\sqrt{1 - |D(n)|^2}}{1 - \bar{D}(n)z}, G_2 > \right\} \tag{61.9}$$

Then, $a_2 = D(n_2)$. When a_2 is known, $B_2(z)$ can be calculated by formula (61.6), then

$$f_2 = \langle G_2, e_{\{a_2\}} \rangle B_2(z) \tag{61.10}$$

(2) Calculating f_{k+1}

$G_1, G_2, \ldots, G_k, a_1, a_2, \ldots, a_k, , n_2, n_3, \ldots, n_k, b_1, b_2, \ldots, b_{k-1}$ are known, so G_{k+1} is calculating by formula: (61.9).

$$G_{k+1} = \left(G_k - b_k * e_{\{a_k\}}\right) \frac{1 - \bar{a}_k(z)}{z - a_k} \tag{61.11}$$

where, $b_k = D(n_k)$. Then a_{k+1} is calculated by: a_{k+1}. If the solution of formula (61.12) is $n_{k+1}, n_{k+1} \in \{1, 2, \ldots, 7705\}$,

$$\max_n \left\{ < \frac{\sqrt{1 - |D(n)|^2}}{1 - \bar{D}(n)z}, G_{k+1} > \right\} \tag{61.12}$$

Then $a_{k+1} = D(n_{k+1})$.

a_{k+1} is known, $B_{k+1}(z)$ can be calculated by formula (61.6), so,

$$f_{k+1} = \langle G_{k+1}, e_{\{a_{k+1}\}} \rangle B_{k+1}(z) \tag{61.13}$$

Repeat the above process and all the mono-components $f_1, f_2, \ldots, f_n$ of G are obtained.

61.3 Rolling Bearing Fault Diagnosis Algorithm Based on AFD

The meaningful signal representing the rolling bearing behavior is always covered by the detected signal, and the detected signal is usually modulated by some other signals which are un-relevant to rolling bearing. The algorithm of rolling bearing

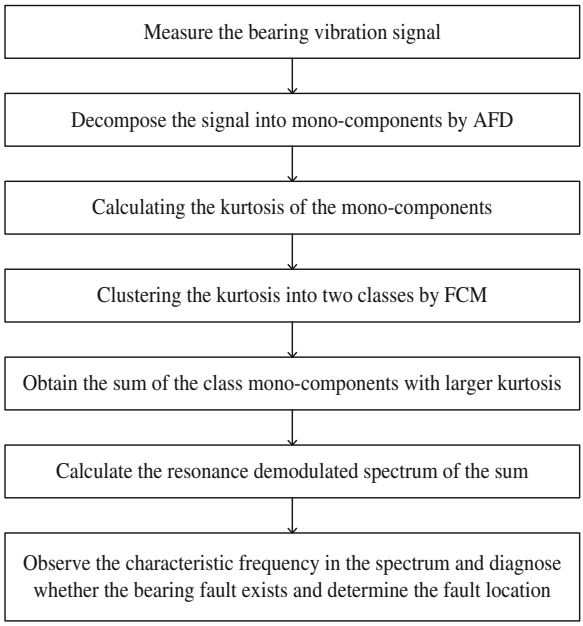


Fig. 61.1 The steps of proposed procedure for bearing fault diagnosis

fault diagnosis is to demodulate the detected signal and obtain periodical fault characteristic frequency. It will be not efficient or even it cannot obtain the meaningful information if it completely demodulates the whole detected signal, because the information of the modulated signal is contained in some frequency band [11]. So, it needs to search a frequency band containing the abundant modulated signal information according to a principle, and extract partial signal with the periodical fault component, which has the same period as original periodic fault impact stimulation.

The approach to diagnosing rolling bearing faults based on AFD does not use the conventional filter, and it directly decompose the signal into a series of mono-components $f_1, f_2, \ldots, f_n$. The BPFO harmonics of mono-components with low frequency are weak and mixed and covered by the noise, and kurtosis is sensitive to the impact component in the signal which is periodical fault impact stimulation that means if demodulated signal in a frequency band contains more information of the modulated signal, the kurtosis should be larger. So the kurtosis of mono-components are calculated and clustered into two classes, and the class of mono-components with larger kurtosis is taken as the carrier signal. Then it calculates the resonance demodulated spectrum of the carrier signal, and whether the bearing has fault or not is determined finally. The main steps of the proposed fault diagnostic process are shown in Fig. 61.1. The algorithm diagnosing bearing fault based on AFD is designed as the following four steps:

(1) Decomposing the original bearing signal into a series of mono-component;
(2) Calculating the kurtosis of mono-components, clustering the kurtosis into two classes, exacting the mono-components with larger kurtosis, then obtaining the sum of the selected mono-components as the modulated signal;
(3) Doing Hilbert transformation to the sum and getting the Hilbert envelope spectrum;
(4) Doing FFT to the Hilbert envelope spectrum and checking whether fault frequency peak exists in the frequency spectrum, and determining whether the bearing has fault or not and the fault location.

These steps show that bearing fault diagnosis does not need to fix filter parameters beforehand as traditional demodulated resonance technology, and this approach is adaptable to different types of bearing.

61.4 Fault Analysis and Experimental Verification of Outer Ring

The experiment uses N205EM type rolling bearing to verify the effectiveness of the AFD algorithm, and diagnoses the bearing outer ring faults. N205EM type rolling bearing is single row cylindrical roller bearing. The bore size is 25 mm, outside diameter is 52 mm, roller diameter is 7.5 mm, pitch diameter is 39 mm, and roller number is 12. The vibration signal of bearing outer ring and roller is obtained from a vibration experiment table and the relative parameters are as follows: bearing rotation speed is 1,500 rpm, signal acquisition frequency is 10,000 Hz. The signal measured time is 0.4096 s. According to these parameters and formula (61.1), the result of calculating fault characteristic frequency of outer ring is 117.2 Hz. Then calculating the resonance demodulated spectrum of the measured original vibration signal based on AFD, and comparing with the calculated theoretical characteristic frequency and judging whether the bearing has fault.

The measured original vibration signal of outer ring is shown in Fig. 61.2. The unit of the amplitude is volt.

It needs to extract the carrier signal from the new bearing vibration signal first. It decomposes the new signal by AFD and 500 mono-components are obtained. Then each kurtosis is calculated and clustered into two classes, which is shown in Fig. 61.3, the cluster centers are the asterisks and the mean of the centers is a circle. The two class centers is corresponding the 20th and 270th mono-component, so 145 mono-components is summed as the carrier signal. As shown in Fig. 61.3, the kurtosis of the 145th mono-component is almost the turning point.

It takes Hilbert transform of the outer ring carrier signal and does fast Fourier transform on the Hilbert envelop, the result is shown in Fig. 61.4a. The first peak is obviously appear at 117.19 Hz, and there is damping peak appears at 117.19 Hz,

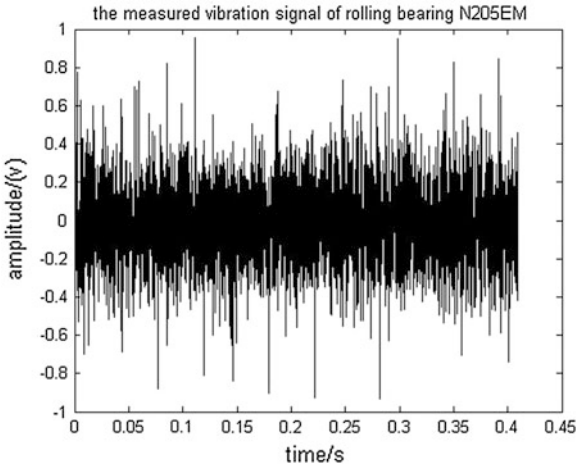


Fig. 61.2 The outer ring vibration signal measured from rolling bearing N205EM

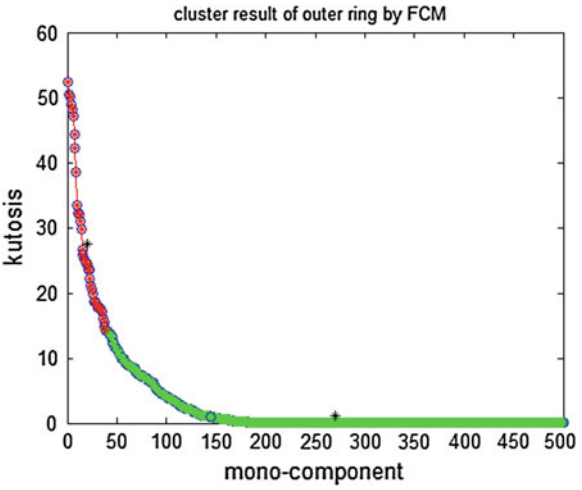


Fig. 61.3 Kurtosis clustering result of outer ring by FCM

which is close to the outer ring fault characteristic frequency, so it can be judged that there is fault on outer ring.

Meanwhile, the traditional demodulated resonance technology is used to analyze the new bearing vibration signal. It first filters the signal by Butterworth bandpass filter and the filter bandwidth is setting as 1,000–2,000 Hz, which is the general bandwidth of the carrier signal. Then it also takes Hilbert transform and does fast Fourier transform on the Hilbert envelop, and the result is shown in Fig. 61.4b.

The advantage of using AFD to diagnose the outer ring fault is shown by the comparison Fig. 61.4a, b. The characteristic in Fig. 61.4a is obviously clearer than Fig. 61.4b. With the traditional demodulated resonance, it even cannot determine whether the outer ring has fault or not in Fig. 61.4b.

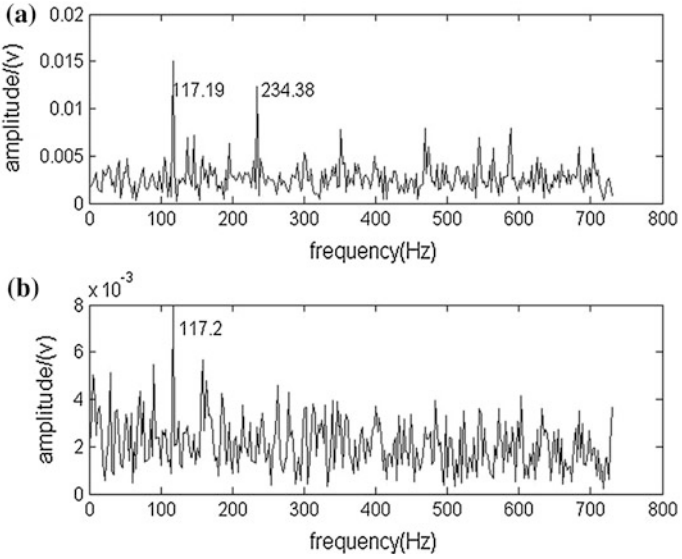


Fig. 61.4 **a** AFD demodulation spectrum, **b** traditional resonance demodulation spectrum

61.5 Conclusions

This paper proposes a new approach to diagnose rolling bearing fault by AFD. The new approach can adaptively diagnose the rolling bearing fault without preset parameters and it can adaptive to all kinds of rolling bearing. The new approach does not need to filter signal by band-pass filter and preset filter parameters as the traditional methods of bearing fault diagnosis. The experimental verification shows the effectiveness and accuracy of new approach.

Acknowledgments This work is supported by the National High Technology Research and Development Program of China (863 Program) (2011AA110501).

References

1. Howard I (1994) A review of rolling element bearing vibration "detection, diagnosis and prognosis". Melbourne, Victoria 3001 Australia. DSTO aeronautical and maritime research laboratory
2. Cheng JS, Yu DJ, Yang Y (2005) Time–energy density analysis based on wavelet transform. NDT E Int 38:569–572
3. Peng ZK, Peter TW, Chu FL (2005) An improved Hilbert–Huang transform and its application in vibration signal analysis. J Sound Vib 286:187–205

4. Huang NE, Shen Z, Long SR et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond 454:903–995
5. Gai GH (2006) The processing of rotor startup signals based on empirical mode decomposition. Mech Syst Sig Proc 20:222–235
6. Wang WY (2001) Early detection of gear tooth cracking using the resonance demodulation technique. Mech Syst Sig Process 15(5):887–903
7. Tang DY (2006) Generalized resonance. resonant demodulation fault diagnosis and safety engineering-for railway. China railway press
8. Qian T, Wang YB (2011) Adaptive decomposition into basic signals of non-negative instantaneous frequencies—a variation and realization of greedy algorithm. Adv Comput Math 34(3):279–293
9. Qian T, Zhang LM, Li ZX (2011) Algorithm of adaptive fourier decomposition. IEEE Trans Sig Process 59(12):5899–5906
10. Qian T, Wang YB, Dang P Adaptive decomposition into mono-components. Adv Adapt Data Anal. doi: 10.1142/S1793536909000278
11. Bozchalooi IS, Liang M (2008) A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals. Mech Syst Sig Process 22:915–933

Chapter 62
Train Control Management System Safety Assessment

Changyuan Liu, Xiaoming Li and Panpan Yang

Abstract Train control management system (TCMS) is responsible for the real-time transmission of monitoring the running state of train operation. The TCMS system safety assessment is an important means to ensure the quality of the products. This paper describes the various stages of the TCMS system in the whole life cycle assessment activities, and defines a complete audit process to prove that the TCMS has certain guidance and significance in the system safety assessment.

Keywords Software · Safety · Assessment · TCMS

62.1 Introduction

The Independent Safety Assessment is performed according to the requirements of the CENELEC standards. This will include an assessment of the life cycle processes and their output documents. The objective of assessment is to arrive at a judgement that TCMS possesses the defined safety integrity level based on credible evidence and is fit for its intended purpose.

62.2 Standard

The standards to be used are allocated as follows:

- The software according to EN 50128
- The hardware and safety case according to EN50129.
- The general requirements according to EN50126.

C. Liu (✉) · X. Li · P. Yang
Fundamental Engineering Department of R&D Center, Changchun Railway Vehicles Co Ltd, Changchun 130062, China
e-mail: liuchangyuan@cccar.com.cn

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_62,
© Springer-Verlag Berlin Heidelberg 2014

- The development and proof of hardware according to EN 50155.
- The software communication for safe data transmission according to EN 50159.

62.3 Program of Assessment

62.3.1 Planning of Assessment

The purpose of this Assessment Plan is to regulate the execution and quality management of all assessment steps of the Assessor regarding:

- the procedure,
- the performance,
- the prerequisites or the conditions to be considered, and
- the recording of the assessment.

The complete assessment activities were divided into two parts. One part will contain all assessment works concerning TCMS system and TCMS hardware, according to EN 50129. The other part will deal with the assessment of the TCMS software, according to EN 50128.

62.3.2 Safety Analysis

It is highly recommended to start and carry out the assessment concurrent to the development process and beyond system acceptance as appropriate as shown in Fig. 62.1. Safety-Related Application Conditions (SRAC) are generated by safety analysis at each level of the project whenever a related Safety Case has been identified or not [1].

62.3.3 Scope of Audit

An audit of the product quality and safety management system will be planned and executed regarding the development of the specific TCMS application.

Within the scope of the audit, the auditors will observe processes in the design department for the specific TCMS application in order to gain an understanding of the overall design processes. The auditors will verify the processes for conformity with the requirements of the respective standards and the descriptions in the relevant documents and quality procedures. Only persons who are directly involved in the development, verification, and validation will be audited. The client overall project organization and the related roles and responsibilities and the verification and validation process will be subject to auditing.

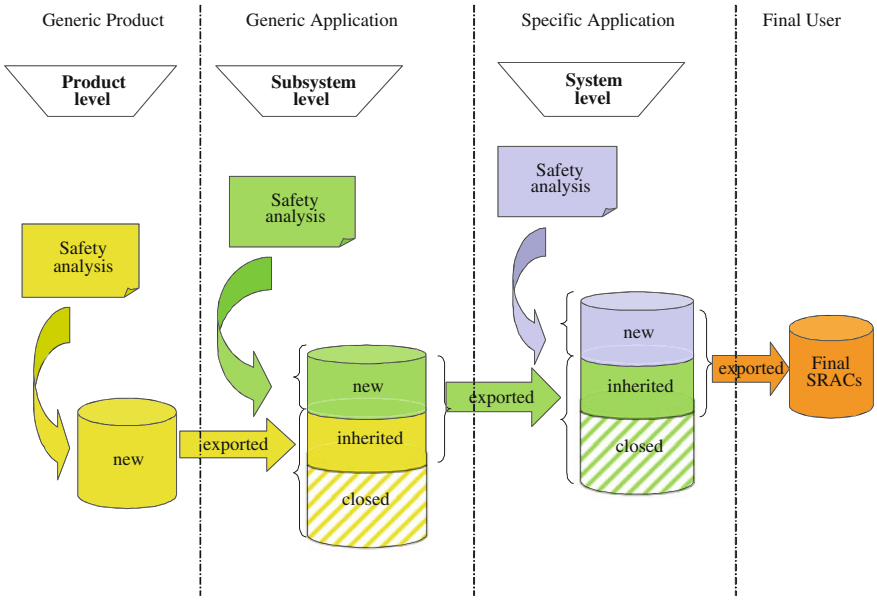


Fig. 62.1 Exported constraints and SRAC management

62.3.4 System and Hardware of Audit

The safety analyses, plans, technical specifications, descriptions, manuals, and test documentations will be inspected. These documents will be reviewed for compliance with the selected standards and consistency with the respective specification.

Specific focus will be given on evidence of compliance with the requirements for the specified Safety Integrity Level, in particular the evidence on the measures against the systematic failures, as well as the remaining failure probability (probability of undetected errors) for random hardware failures. Table 62.1 shows the typical assessor activity during the life cycle.

62.3.5 Software of Audit

Selected software documents will be assessed with respect to existing conformance with the CENELEC EN 50128 life cycle demands for SIL 2 (a typical process is shown in Fig. 62.2) [2].

Table 62.1 Typical assessor activity during the life cycle

Phase	Assessors activity
Concept	Review the RAMS principles and safety integrity requirements
System definition	– Assess the relevance and completeness of high-level safety requirements and preliminary hazard analysis – Assess the, safety plan and quality plan – Assess the system boundary, constituents, interfaces with other systems, and operational environment
Risk analysis	– Assess the risk analysis methods and process – Assess the risk tolerability criteria
System requirements	– Assess the preliminary hazard analysis is consistent with the hazards considered in the risk analysis – Assess the system RAM and safety requirements according to: – Testability – Completeness
Design and implementation	Assess that – Processes, methods and techniques of system safety plan are applied (also for subcontractors and suppliers) – all planned verification of that system design are done and robust Assess the – Hazard analysis including operational and interface hazards
System validation (acceptance and commissioning)	Assess that – every safety requirement is tested, verified and adequately fulfilled – all non-compliances against safety requirements are recorded, assessed and justified, and placed in the concluding clause of the safety case – every safety related fault is described and its risks are evaluated and controlled to an acceptable level
System acceptance	Assess that – the operation and maintenance manuals include all safety related application conditions and are considered adequate for safe operation and maintenance of the system – all constraints imposed on the system from the operational environment are included in the operation and maintenance manuals – all system risks including those relating to operation and maintenance have been adequately identified, controlled, and managed

(continued)

Table 62.1 (continued)

Phase	Assessors activity
Operation and maintenance	If an assessors role is continued after acceptance, then he/she should assess that – the essential safety documentation are handed over to the operators or the new project team – the operator/new project team continues to maintain a hazard log – new hazards are recorded, evaluated, and their risks are managed
Modification and retrofit	If an assessors role is continued after acceptance, then he/she should assess that – the Impact of any change is risk assessed and managed – there's a change and configuration management process in place and appropriately applied
Decommissioning and disposal	When applicable, assess that – the decommissioning is performed according to a plan – appropriate risk control measures are adopted and applied – relevant competencies are catered for during the decommissioning

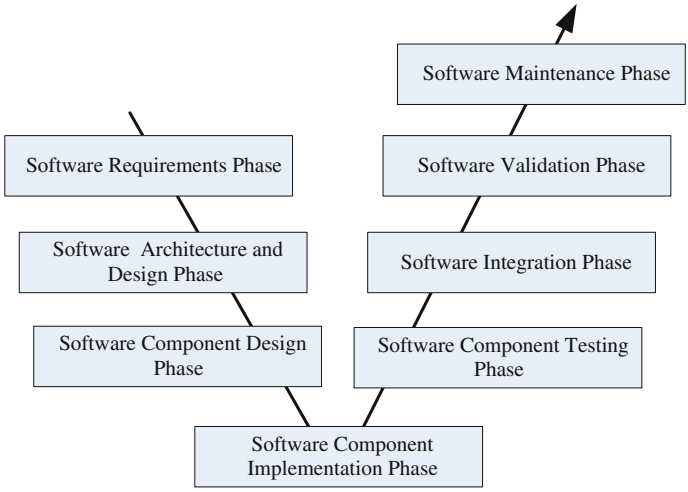


Fig. 62.2 Software life cycle

62.3.6 Test Witnessing

The conduction of audit and test witnessing by an independent assessor is demanded by the CENELEC standards.

A Test Witnessing and V&V (follow up) audit will be performed to gain evidence that the TCMS system has been implemented and tested correctly according to the specific TCMS application design documentation and rules and standards. During the test witnessing, tests of the specific TCMS design and integration tests will be observed and the results compared with the test specifications. The CLIENT V&V and test processes, design, manufacturing, train integration, test performance, and results will be subject to test witnessing.

The audits will address, among other aspects, the following subjects [3]:

- The supplier's organizations involved (with special attention to the Quality Department and the V&V Department).
- The Project Quality Organization.
- The Project Safety Organization.
- The Project V&V Organization.
- The development process (the life cycle of V&V).
- Inspection and testing.
- Environmental conditions.
- Tthe change control and configuration management.
- The non-conformity record (NCR) management.
- Demonstration of parts of system and subsystem functionality.
- Management of project documentation.

One followup audit will be performed, which will be combined with the Test Witness. The results will be summarized in an Audit and Test Witnessing Report.

Special emphasis will be given on the following process aspects:

- Application of quality assurance system and project specific safety plans.
- Project organization, roles and responsibilities, qualification, and experience for the specified Safety Integrity Levels.
- Safety target and safety level appointment and refinement process.
- Hazard and risk analysis process, hazard log handling, and followup process.
- Requirement specification and traceability process, requirements acceptance criteria.
- Operation and maintenance procedure, test acceptance criteria.
- Operation and maintenance procedure.

Table 62.1 shows assessor activities during the life cycle phases for a safety-related product or system development. Table 62.1 is a proposal of how to ensure the appropriate assessor activity. The assessor's output from each phase is generally a documented feedback or report distributed to the parties described in the safety plan. The assessor may additionally generate and deliver examination reports during certain phases covering activities, document structures, and analysis [4].

The Test Witnessing will be performed on a sampling basis, by interviews with the testers, review of the test procedures, by inspecting the test setups and attending manufacturing tests, and an on-site demonstration of the respective system. The Test Witnessing does not serve the purpose to prove that the specifications meet the said criteria. It must rather be evaluated that these criteria have sufficiently been observed by other persons (V&V responsible persons) during their inspection of the specifications.

In particular it will be checked whether tests plans and specifications are appropriate and whether they have been accurately implemented, whether the test environment is suitable for its purpose (e.g., that the operational scenarios under test have correctly been established/simulated, that the test results are clearly and unambiguously recorded), and whether the tests have shown the necessary coverage of safety aspects as well as the expected results and thus achieved compliance with the defined test acceptance criteria [5].

Special attention will be paid to the TCMS interfaces to various train subsystems with respect to TCMS functions. Test witnessing will concentrate on a typical train equipment installation.

The client role during the Test Witnessing on-site is to demonstrate the practical application of its documented procedures for testing. The auditors verify the effectiveness of the implemented test procedures in accordance with the requirements of the specified SIL and additional requirements agreed upon.

The test witnessing results and acceptance of corrections for deviations will be documented in the Test Witnessing Report. It contains the reason, scope, performance, results, and recommendations if any.

The auditors will document nonconformities which they observed in deviation reports and decide about the classification into major (critical) and minor (non-critical) deviations. For these deviations, corrections and corrective actions will be defined. Corrections shall be implemented and verified.

62.3.7 Management System of Audit

An audit of the client's Quality and Safety Management System will be prepared and executed regarding the development and production of the TCMS system. Only persons who are directly involved in the development, verification/validation, and production for the project will be audited. The audit will consider the requirements of the applicable standards regarding supplier organization and processes. An audit questionnaire/checklist will be prepared by the assessors in advance. Services will consist of assessment of organizations, roles and responsibilities, assessment of the adequacy of processes, tools, methods, criteria, tests, assessment of documentation process, spot checks of individual aspects, and spot checks of individual/selected documents.

The safety analyses, plans, technical specifications, descriptions, manuals, and test documentations will be inspected. These documents will be reviewed for compliance with the selected standards and consistency with the respective specification.

The client will submit a Safety Case for the TCMS for the application, following the structure and principle content according to EN 50129, namely the TCMS Specific Application Safety Case.

This safety case has to provide evidence of Quality Management, evidence of Safety Management, evidence of Technical and Functional Safety, demonstration of specified operation in the presence of specified, external influences, references to related Safety Cases, and the SRAC. The Safety Case document(s) itself will be analyzed for completeness, validity, unambiguity, comprehensibility, and consistency in reference to the CENELEC Standards [6].

62.4 Conclusion

Safety approval is a subsequent stage to assessment and considers the outcome of the judgement arrived at the assessment phase in order to allow the acceptance of implementation or deployment of the product or system. Safety approval is generally given by a safety authority that may employ the outcome of the assessment or additional evidence to arrive at their judgement.

References

1. EN50129 (2003) Railway applications—communications, signalling and processing systems—safety related electronic systems for signalling, pp 28–29
2. EN50128 (2011) Railway applications—communications, signalling and processing systems—software for railway control and protection systems, pp 22–23
3. EN 50126-1 (1999) Railway applications—the specification and demonstration of reliability, availability, maintainability and safety (RAMS), pp 25–30
4. TR 50506-2 (2008) Railway applications—communication, signalling and processing systems—application guide for EN 50129-Part 2:safety assurance, pp 66–69
5. Xiong Y (2009) Software testing technology. Beijing University of Aeronautics and Astronautics, Beijing, China (in press), pp 9–13
6. IEC61508 (2010) Functional safety of electrical/electronic/programmable electronic safety-related systems-Part 1:general requirements, pp 14–17

Chapter 63
Research on Disruption Management of Single Machine Scheduling

Lei Xiao

Abstract Based on common due date, this paper discusses the disruption management problem in the middle of JIT. If a hardware has been finished ahead of schedule, measures have to be adopted to arrange the hardware over again in order to reach the optimization. In addition, this paper discusses the paradox that we do not punish but encourage the hardware ahead of schedule. Corresponding to such disruption management problem, this paper gives two types of algorithm and an example.

Keywords Scheduling · Disruption management · JIT · Dynamic planning algorithm

63.1 Introduction

For a long time, the object function in the research about scheduling problem is canonical, and it is not descending function about the time [1], i.e., make span, average flow. With the appearance of Just In Time, the object function cannot reflect the relation of produce—storage—demand. The scheduling problem about JIT reflects the question that the hardware ahead of schedule do not be encouraged, therefore, the good scheduling ought to be adopted so that the whole hardware can be got through on time. The case roots in the fact that the hardware require to be got through on time every phase.

L. Xiao (✉)
National School of Development, Peking University, Beijing 100871, China
e-mail: xiaoleiyouhua@sina.com

School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 40073, China

Postdoctor Innovative Practice Base, Shenzhen Statistical Society, Shenzhen 518031, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_63,
© Springer-Verlag Berlin Heidelberg 2014

There have been a lot of good results about the JIT problem based on common due date since 1970s. The consequence can be assorted two classes. One class is to look for the optimal scheduling and the optimal due date, the summary can be found in [2]. Another is to look for the optimization scheduling on the condition of the appoint time, this question is more different than the above question commonly, the summary can be found in [3]. The JIT question about common due date is discussed at best in these questions.

Give the following sign about scheduling:

d : common due date

$\min_{\sigma,d} f(\sigma, d)$: look for the optimal scheduling and the optimal due date vector.

$\min_{\sigma} f(\sigma, d)$: look for the optimal scheduling if the due date is given.

$\min_{d} f(\sigma, d)$: look for the optimal due date if the scheduling is given.

The question about $\min_{\sigma} f(\sigma, d)$ is often divided into two classes: nonfinite and finite. Give the value of d, $\min_{d} f(\sigma, d) = f^*$, if $d_1 \geq d$, $\min_{d} f(\sigma, d) = f^*$, then the question is called nonfinite, otherwise it is called finite.

Now the research about this question is summarized below:

Model A:

$$\min_{\sigma} \sum_{j=1}^{n} \left(\alpha_j T_j + \beta_j E_j\right) \tag{63.1}$$

Equation (63.1) is NP—hard question, on the condition of consistency, [4] use taboo to give the approximate algorithm about the finite form (63.1), [5] give the dynamic planning algorithm about fake multinomial time. Sware [6] give the three form of the optimization scheduling, and.

Model B:

$$\min_{\sigma,d} \sum_{j=1}^{n} \left(\alpha_j T_j + \beta_j E_j\right) \tag{63.2}$$

After let

$d - C_j = E_j - T_j$,

Bector et al. [7] equivalent (63.2) to (LP):

$\min \sum_{j=1}^{n} \left(\alpha_j T_j + \beta_j E_j\right)$

S.T. $d + T_j - E_j = C_j$

$d \geq 0$, $E_j \geq 0$, $T_j \geq 0$

While give the following results: for the scheduling σ, exit the optimization $d^*(\sigma) = \mathrm{C}_b(\sigma)$, b satisfy the relation: $\sum_{j=1}^{b-1} \beta_j \leq \sum_{j=b}^{n} \alpha_j, \sum_{j=1}^{b} \beta_j \leq \sum_{j=b+1}^{n} \alpha_j$.

Based on some basic conception of sensitivity analysis and linear programming, [7] give some algorithm, and [8] give branch and band algorithm.

Model C:

$$\min_{\sigma,d} \sum_{j=1}^{n} \left(\alpha_j T_j + \beta_j E_j + \theta C_j\right) \tag{63.3}$$

Chen et al. [9] give O (nlogn) algorithm. for m machine, [10] prove if $m = 2$, the question is changed to NP—hard, and give a special example which is able to be settled by polynomial algorithm.

This paper discusses the disruption management problem in the middle of the JIT question: if a hardware has been finished ahead of schedule, measures have to be adopted—the hardware has to be arranged over again in order to reach the optimization. In addition, this paper discusses the question which we do not punish but encourage the hardware ahead of schedule. For example, in a cargo handling problem, trucks and warehouse need to satisfy the relation: n trucks reach the same harbor at zero time, and hope to finish tasks at the same time. If a truck cannot finish work at its due date, or finishes work ahead of schedule, warehouse need to consider the schedule over again in order to minimize the cost.

63.2 Problem Description

This paper adopts the method of the three parameters denotation [10], and adopts the following basic hypothesis:

1. If the machine begins to work, it cannot stop, unless all the jobs have been completed.
2. Once the hardware is being done, it cannot be interrupted.
3. A set of n hardware, are machined at zero time.

This paper also appears following the variables and parameters:

d : common due date, apparently $d_j = d, j \in N$
C_j : the time which the hardware j is completed
E_j : the advance of hardware j, $E_j = \left(d - C_j\right)^+ = \max\left\{0, d - C_j\right\}, j \in N$
T_j : the delaying of hardware j, $T_j = \left(C_j - d\right)^+ = \max\left\{C_j - d\right\}, j \in N$
α_j : the compensation gene of hardware j, $\alpha_j > 0, j \in N$
β_j : the ahead encouragement gene of hardware j, $\beta_j > 0, j \in N$
γ_j : the common cost gene of hardware j, $\gamma_j > 0, j \in N$

This paper considers the following problem: looking for a scheduling in order to minimize the function:

$$\sum_{j=1}^{n} \left(\alpha_j T_j - \beta_j E_j + \gamma_j C_j\right).$$

63.3 Problem Settlement

63.3.1 Neighborhood Search Algorithm

Neighborhood search algorithm begin with the initial solution, then use a generator to search better solution than the initial solution in neighborhood. If find a solution than the initial solution, the new solution will substitute it, then algorithm continue; otherwise the algorithm stop.

Neighborhood search algorithm technique:

Step1, according to one formula, fetch a scheduling for germ and calculate the value.

Step2, according to one formula, construct a neighborhood around the germ and calculate value. If the new value is not better than germ, stop. Or else go to Step3.

Step3, according to one formula, choose the better value for new germ, return to Step2.

example:

Step1, germ 1-2-3-4-5

neighborhood 2-1-3-4-5, g = 91

1-3-2-4-5 g =79 (choose)

1-2-4-3-5 g =95

1-2-3-5-4 g =79

Step2, new germ 1-3-2-4-5

neighborhood 3-1-2-4-5, g =83

1-3-4-2-5, g =97

1-3-2-5-4, g =75 (choose)

Step3, new germ 1-3-2-5-4 g = 75(choose)

neighborhood 3-1-2-5-4, g =79

1-2-3-5-4, g =79

1-3-5-2-4, g =79

1-3-2-4-5 g =79

Optimal value : 1-3-2-5-4, g = 75

The following measures can help us to achieve better solution:

(1) Choose some germs for the algorithm so that many solutions can be achieved, then we can choose the most optimal solution.
(2) Choose the scheduling which has been improved for the germ in the neighborhood, and construct a new neighborhood around the new germ.
(3) Construct a neighborhood as large as possible.

In fact, many complex questions is difficult to find the suitable initial solution, sometimes though the suitable initial solution is found, the neighborhood is often too complex to be constructed. We can apply dynamic planning algorithm to solve this problem.

63.3.2 Dynamic Planning Algorithm

Dynamic planning algorithm is a math means in operational research which can solve multistage optimization problem. It was produced in 1950s by R. Bellman, born in America. however, Not all NP-hard problems can be solved by dynamic planning algorithm,

When $Z = \sum_{j=1}^{n} g_j(C_j), j = 1, 2 \cdots n$ is canonical, the problem can be solved by dynamic planning algorithm.

$\sum_{j=1}^{n} (\alpha_j T_j - \beta_j E_j + \gamma_j C_j)$ is canonical, so it can be solved by dynamic planning algorithm.

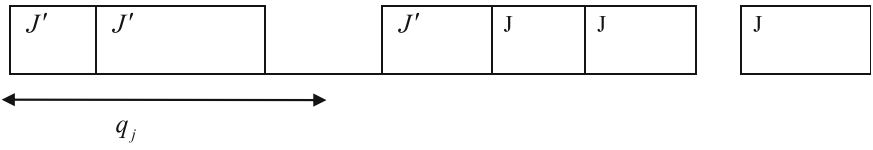
For the scheduling problem 63.1 $|\bullet|\sum_{j=1}^{n} (\alpha_j T_j - \beta_j E_j + \gamma_j C_j)$, let

$$\begin{aligned} g_j &= \alpha_j T_j - \beta_j E_j + \gamma_j C_j \\ &= \alpha_j \max\{0, C_j - d\} - \beta_j \{0, d - C_j\} + \gamma_j C_j \end{aligned}$$

$J \subset N, J' = N/J, N = \{1, 2 \cdots n\}$ when J denotes one subset of n hardware, J' is the complemented of J, $q_j = \sum_{j \in J'} t_j$ denotes all the time which complete the hardware in the set J'.

Suppose a schedule, when the hardware in J' is ahead of J', then bases on the optimization theory, if the schedule is the optimization, no matter what the schedule of the hardware of J', the schedule of hardware in J is also the optimal value.

Let G (J): the lowest cost which the hardware need not ahead of q_j in J



Construct the following recursion algorithm:
First, let

$$\mathrm{J} = \emptyset,\ \mathrm{G}(\emptyset) = 0 \tag{63.4}$$

Second, let $|J| = 1$, for every J:

$$G(J) = g_j\left(q_j + t_j\right) + G(J - \{j\}) \tag{63.5}$$

Then, let $|J| = 2$, for every J:

$$G(J) = \min_{j \in J}\ g_j\left(q_j + t_j\right) + G(J - \{j\}) \tag{63.6}$$

Commonly, for every $|J| = \mathrm{k}$, calculate 63.6,
When $|J| = n$, get the optimal value.
Example Initial time: $J = \emptyset$, G $(\emptyset) = 0$:

J	t_j	d_j	α_j	β_j	γ_j
1	1	7	1	1	1
2	2	7	1	1	1
3	3	7	1	1	1
4	4	7	1	1	1

Circle 1: $|J| = 1$

J	q_j	$G(J)$
{1}	9	13
{2}	8	13
{3}	7	13
{4}	6	13

Circle 2: $|J| = 2$

J	q_j	$G(J)$	Optimal value
{1, 2}	7	22	(1, 2)
{1, 3}	6	20	(1, 3)
{1, 4}	5	18	(1, 4)
{2, 3}	5	20	(2, 3)
{2, 4}	4	18	(2, 4)
{3, 4}	3	18	(3, 4)

Circle 3: $|J| = 3$

J	q_j	$G(J)$	Optimal value
{1, 2, 3}	4	23	(1, 2, 3)
{1, 2, 4}	3	19	(1, 2, 4)
{1, 3, 4}	2	17	(1, 3, 4)
{2, 3, 4}	1	17	(2, 3, 4)

Circle 4: $|J| = 4$

J	q_j	$G(J)$	Optimal value
{1, 2, 3, 4}	0	12	(1, 2, 3, 4)

63.4 Conclusion

This paper analyzes a disruption management problem in single machine scheduling according to neighborhood search algorithm and dynamic planning algorithm. Different from the general JIT problem, the paper studies the following instance: one hardware is completed ahead of schedule, and the case is encouraged.

Neighborhood search algorithm adapt to the question which the initial solution is acquired easily and the neighborhood can be found quickly, in addition, the algorithm is not adapt to the question which has too much variables. Though the construction of dynamic planning algorithm is more complex than neighborhood search algorithm, applying dynamic planning algorithm, the paper gets not only the global optimal value but also some local optimal values, which is in favor of the analysis. however, the algorithm has not standard model, and the variable has to satisfy the "uneffectivity" condition, which reduce the currency of dynamic planning algorithm. Certainly, we can adopt different heuristic algorithm to study the problem in view of disruption scheduling.

References

1. Baker KR (1974) Introduction to sequencing and scheduling. John Wiley and Sons, New York
2. Cheng TCE, Gupta MC (1989) Survey of scheduling research involving due date determination decisions. Euro J Oper Res 38:156–166
3. Baker KR, Scudder GD (1990) Sequencing with earliness and tardiness penalties: a review. Opns Res 38:22–36

4. Yan S, Tu Y (1997) Multifleet routing and multistop flight scheduling for schedule perturbation. Eur J Oper Res 103:155–169
5. Lee CY, Danusaputro S, Lin CS (1991) Minimizing weighted number of tardy jobs and weighted earliness-tardiness penalties about a common due date. Computers Opns Res 18:379–389
6. Sware W (1996) The weighted commom due date single machine scheduling problem revisited. Computers Opns Res 23:255–262
7. Bector CR, Gupta YP, Gupta MC (1988) Determination of an optimal common due date and optimal sequence in a single machine job shop. Int J Prod Res 26:613–628
8. Vickson RG (1980) Choosing the job sequence and processing times to minimize total processing plus flow cost on a single machine. Opns Res 28:1155–1167
9. Chen B, Potts CN, Woeginger GJ (1998) A review of machine scheduling: complexity, algorithm and approximability. In: Du, D–Z. et al. (eds) Handbook of combinatorial optimization, Kluwer Academic Publishers, Massachusetts, pp. 21–169

Chapter 64
A Study on Parking Problems and Countermeasures of Urban Central Commercial District

Xianyong Miao and Manzhen Duan

Abstract Parking facilities are important to the construction and development of urban central commercial district. With the increase of urban economy and people's living standards, more and more cars, followed by more and more city parking is not convenient. Especially in the city center business district, "parking difficult" problem is particularly serious. In order to solve the parking problem, Yuan Yang-Cheng in T city commercial district is taken as the object, according to the investigation on the parking behavior. The survey results show that the key factors affecting the behavior of parking is price and the parking is interested in the convenient parking of parking guidance system. Also, the reasonable use of price lever parking supply and demand has great room for adjustment.

Keywords Central commercial district · Parking behavior · Static traffic · Parking management

Project supported by: Specialized Fund for national Facilities and Information Infrastructure for Science and Technology (2013DDJ1ZY06); Special for science and technology evaluation and statistics (2013SE-0603).

X. Miao (✉)
School of Traffic and Transportation, Beijing Jiaotong University, No. 3 Shangyuancun, Haidian District, Beijing 100044, People's Republic of China
e-mail: hblgdxmxy2008@163.com

M. Duan
School of Civil Engineering and Architecture, Hebei United University, Tangshan 063009 Hebei, China

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_64,
© Springer-Verlag Berlin Heidelberg 2014

64.1 Introduction

Whether parking facilities are perfect or not, they not only affect the relative area of road traffic flow and the road traffic safety, but are also related to the rise or fall of the commercial district. In recent years, with the rapid increase of motor vehicles in our country, parking problems have emerged in many big cities, especially in the traditional commercial districts. T City Yuan Yang-Cheng and Ba Fang shopping areas appear as similar cases. At present, the parking management has become an important aspect of city traffic management. Therefore, to study the parking behavior analysis of citizens in the central business district of the parking lot to ease the city business district traffic demand has an important role in reducing road traffic [1].

The author based on a comprehensive survey to find the solution of parking problem in business district parking center in T city Yuan Yang-Cheng areas. The paper attempts to analyze the survey data in the region and provide the basis for a Parking Policy Formulation Department [2].

64.2 Investigation of Urban Central Commercial District Traffic Characteristics and Parking Behaviors

64.2.1 Traffic Characteristics of the Urban Central Commercial District

The traffic characteristics of a city center business district in China has the following characteristics: (1) A strong radiation force and trip attraction; (2) Large traffic flow, traffic load, heavy traffic delay; (3) The rate of bus' coverage is high, and the public transportation is highly developed; (4) The various travel mode interactions are serious; (5) The parking demand is high and the shortage of parking facilities.

The parking behaviors are closely related to the parkers' behavior characteristic. The commonly used indicators of parking are for purpose of parking, the walk distance from the park, and parking time [3].

64.2.2 Investigation of Parking Behavior

To survey the parker's parking behavior survey in the central business district, usually it needs to prepare some preliminary investigation before the implementation of traffic survey in an area. The questionnaire design is very important. This paper selects parking, their gender, age, occupation, parking time, amount of

consumption, whether the first hope parking, the parking lot reasons, often near Ocean City Commercial District, parking, car from the parking lot to the destination can tolerate walking distance, 15 indexes of parking, parking choice is considered the primary factor, etc. According to the reality of the traffic situation in our country, the investigation was based on the city of T, taken by field data collection and field interview relevant personnel in the survey on parking, the survey questionnaire for parking and face-to-face interview. The investigators interview the parking lot parking, on-site questionnaire at the specified time and place. In addition, it follows a concise and efficient principle. The project includes vehicle characteristics, parking behavior-related factors, and parking field characteristics of individual characteristics.

This research selected two underground parking lots of T City Yuan Yang-Cheng area (EM, YYC), 2 ground off-street parking (PS, YYCG), and three road parking lots (LF Road, CN Road, HY). Yuan Yang-Cheng area parking peak usually appears in the morning 10:00–11:30, and 15:30–17:30 in the afternoon according to the statistics of the relevant agencies. Therefore, it focuses on the time of the investigation in this time period. At the same time, taking into account the weekdays and holidays of the object area visitors will be different, so parking behavior is not the same. The research work was conducted in April 1, 2012 to May 10. The survey questionnaire is 362 copies. The questionnaire at the same time also specifies the investigation time, the parking lot of the instantaneous parking usage (instantaneous rate = instantaneous parking number/parking berth) conducted a survey, and statistical results as shown in Table 64.1.

Table 64.1 shows that there is a serious imbalance in the usage rates among different parking lots. The main characters are: (1) the usage rate of underground parking lots is far below that of the ground parking lots, whether on ordinary days or on festivals. On ordinary days, the usage rate of underground parking lots is 23.5 %, and the usage rate of ground parking lots is 74.8 %; on holidays, the usage of underground parking lots is 67 % and ground parking lots 96.8 %. (2) The average usage rates of ground parking lots are: 96.8 % in holidays, and over 100 % when traffic rush with many cars crowding on the passageways; 74.8 % on ordinary days. The average usage rates of underground parking lots are: 67 % on holidays and 23.5 % on ordinary days. It is obviously shown that the usage rate on holidays is much higher than on ordinary days.

The reasons are as follows: (1) the price charged by underground parking lots are generally much higher than that charged by ground parking lots; (2) the design of underground parking lots is not convenient; (3) it is convenient and fast to park in ground parking lots; (4) there is asymmetry in parking information.

There is always a phenomenon that the parking space has unbalanced use among different parking lots. Thus on the one hand some drivers cannot find the right parking position, and on the other hand some parking lots are left with many parking positions not used. These cars have produced a redundant traffic volume in order to find vacant parking positions. The car owners prefer the ground parking lots rather than the underground parking lots, which has intensified the unbalanced usage phenomenon in parking space utilization. Car owners attach much

Table 64.1 The information of the target parking lot

Parking lot	Type	Charge	Use rate (%)	
			Weekday	Weekend
PS	Outside of the road	¥ 2/h	81	100
YYCG	Outside of the road	¥ 2/h	86	96
LF Road	On the road	Free	89	95
CN Road	On the road	¥ 3/h	57	98
HY Road	On the road	¥ 2/h	61	95
EM	Underground	¥ 6/h	21	66
YYC	Underground	¥ 6/h	26	68

importance to whether the parking lot is near to the destination when choosing a parking lot. From this point of view, the reason they prefer the ground parking lots is that it is much cheaper and convenient. We suggest that the administrative department of parking lots make reasonable price standard for parking, decrease the parking price in underground parking lots, increase appropriately the parking prices in ground parking lots and street parking lots, encourage people to use the underground parking lots, and make the ground parking lots mainly used for short-time parking.

The advantage of ground parking lots is that it is cheap and convenient; the disadvantage of underground parking lots is that the parking price is much higher than that in ground parking lots. In order to use the underground parking recourse appropriately, decrease the unbalanced parking space using phenomenon, meets the requirements of the car owners, our strategy is to increase the parking price in ground parking lots, and decrease the parking price in underground parking lots.

From the view of the malls' profit, with the purpose of farthest attract the customers and increase the sale, the merchant can attract customers by making some favorable measures, such as offering a favorable parking price standard to the customers shopping in the mall, and decrease the customers parking cost. From the view of customers, the best choice is while enjoying shopping, they can also enjoy a convenient and cheap parking position.

64.3 The Main Findings

64.3.1 Distribution of Parking and Parking Time

The survey of parking and parking time results in Figs. 64.1 and 64.2 are shown.

Figure 64.1 shows that the parking proportion is very different on holidays and during working days. Parking for the purpose of work usually accounted for 31 %, while holidays accounted for 12 %. The parking ratio on holidays for the purpose of shopping accounted for 58 %, accounted for 72 %; for the purpose of entertainment parking usually accounted for 3 %, accounted for 12 % of the holiday.

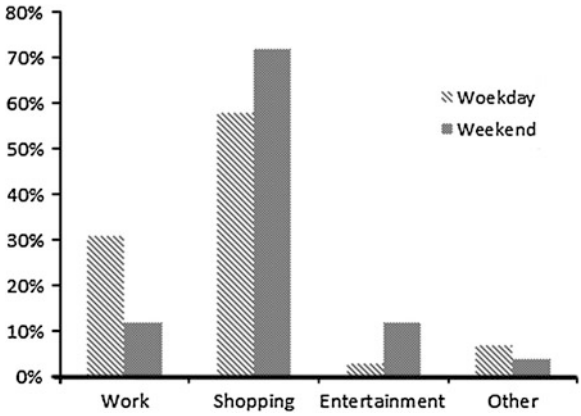


Fig. 64.1 The purpose of parking

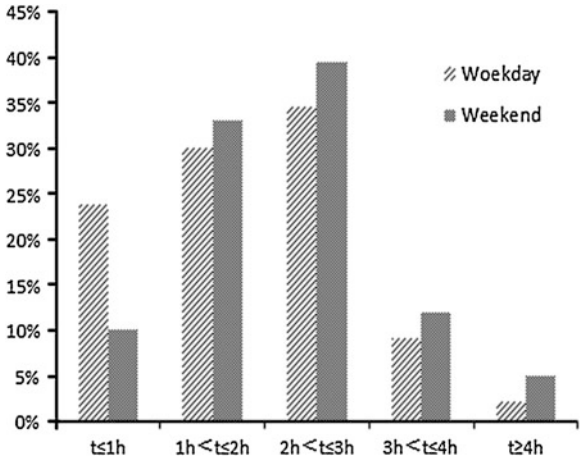


Fig. 64.2 The distribution of parking time

To work for the purpose of parking at 19 % points higher than the holidays, to shopping for the purpose of parking holidays is 14 % points higher than usual, for the purpose of entertainment parking holidays is 9 % points higher than usual. In addition, as can be seen from the graph, the region regardless of holidays or on weekdays to shopping for the purpose of parking are ranked first. This fully demonstrates the Yuan Yang-Cheng business district land properties. Therefore, we should rethink an approach to reduce stop time for those more than 3h parking to improve parking turnover when formulating strategy.

As can be seen from Fig. 64.2, parking times in 1–2 h weekday and holiday parking ratio are basically the same, but the parking time is less than 1 h weekday and holiday parking ratio is significantly different there. As can be seen from the chart, the parking time is less than the parking ratio of 1 h than usual holiday high 10 % points, to 19.3 %. In 2 h $< t \leq$ 3 h, 3 h $< t \leq$ 4H and $t >$ 4 h period, the holiday parking ratio is higher than usual, which fully shows that during the

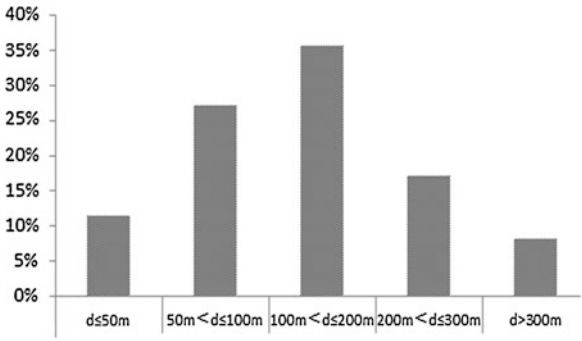


Fig. 64.3 The tolerable walking distance

holidays, people usually spend more time on shopping and entertainment. In addition, the proportion of the holiday parking above 3 h 37.1 %, normal 29.1 %, an average of 33.1 %. This shows that, in Yuan Yang-Cheng business district about 33 % parking spaces are stopping time 3 h above the parking occupancy in a long time, thus greatly reducing the parking turnover rate, reduce the use of parking facilities in the commercial area of the rate of. Therefore, it should rethink an approach that makes 3 h more parking to reduce stop time, and improve parking turnover when formulating strategy.

64.3.2 Considering the Factors of the Parking Choice

Figure 64.3 shows that parking can accept the longest walk distance is less than or equal to 50 m accounted for 11.5 %; in 50, 100 m accounted for 27.3 %; in 100, 200 m accounted for 35.7 %; in 200, 300 m accounted for 17.2 %; more than 300 m accounted for 8.3 %. Overall, that can accept the longest walk distance below 200 m accounted for 74.5 %; that can accept the longest walk distance below 300 m accounted for 91.7 %. Therefore, in the parking facilities planning, can be considered the service set parking radius of 200–300 m.

As can be seen from Fig. 64.4, the parker will be the primary factor to consider price as first, and safety is second. Therefore, in the parking facilities planning, construction, and management, the first thing to consider is prices and security in the parking lot. Parking choice parking price is low, mainly in order to reduce parking costs, and T, most of the parking spaces are free and the price is low, drive customers pay a larger amount of parking fees have not formed the habit of. Select the parking accounted for only 13 %, so you may not have too much to consider this factor in the development of parking policy. Choose the price as parker accounted for 52 %, accounted for the largest proportion, show driving consumer sensitivity to parking prices, so change the parking charges can be changed, the parking lot non-equilibrium parking, improve the utilization of underground parking rate.

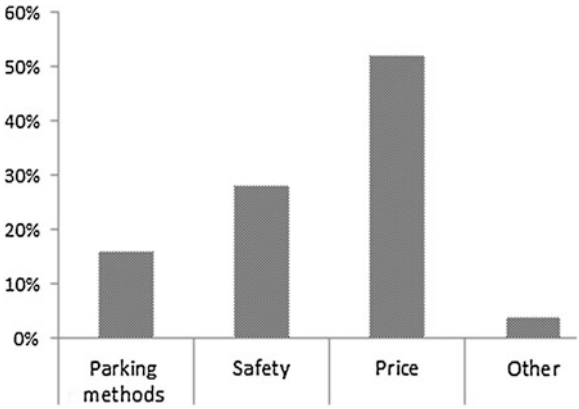


Fig. 64.4 The factors of the parking choice

Through the experts assessment and survey results, the effective shopping time in the parking is 2–3 h for consumers, more than 3 h will not bring more sales to the shopping malls, shopping time is invalid. In order to give the market to drive the consumer benefits, also in order to improve store sales and profits. First, through reasonable increase in garage parking prices, and then through the implementation of preferential policies to stop shopping, not in the shopping vehicle in other parking lot, as in the shopping car of consumer shopping for parking spaces, a small amount of who consumers choose not to drive, not only can reduce the occupation of parking spaces, but also reduce the city road traffic. On the other hand to reduce parking fees through the shopping amount, so as to prompt more consumers in the shopping malls. Through the implementation of shopping preferential parking policy, the car consumer sales can increase. The measures can not only significantly improve the mall parking turnover rate, but also increase the number of free parking spaces, and also reduce the number of roadside parking, improve road capacity, and reduce traffic congestion.

64.4 Conclusion

1. To raise the price and preferential policy. Through reasonable increase in garage parking prices, and then shopping preferential parking policy, make non-shopping vehicle in other parking lot, as in the shopping car of consumer for parking, the parking lot can provide maximum service for the shopping mall. This method is very effective, and not only solves the problem of difficult parking peak shopping district, but also improves the shopping malls and parking profit greatly, and is also convenient for driving customer parking demand who enjoy the preferential parking service.
2. Lifting parking management means and gradually realize the intelligent parking management. Parking guidance and information system is the technical means

to achieve intelligent parking management system, parking in the parking lot can reduce the peak stage to find parking roundabout traffic generation, alleviate traffic pressure, to provide parking scheme is optimized, parking will be efficient use. The parking management of parking to scientific, rational, visualization, humanized management, provides the high quality service for parking, play a role in order to improve the market competitiveness. Customers can also get more convenient parking service from the information management of parking.

3. Commercial and office land development should be matched. Commercial land in nonworking day to attract passengers and office space to attract the work day passenger flow, in order to balance work and non-workdays' parking demand. The commercial area of commercial land development and office land development should be matched so as to make the office of land for commercial use to provide potential flow, the parking lot and the construction during the working day and the nonworking day utilization balance.

This paper conducted in-depth studies on the parking status on T city's central business district and made a further analysis and development. To conduct a full investigation on the parking situation area, parking index and paying more attention to parking is carefully designed and analyzed the investigation statistics and summary. The most concerning factors are the parking behavior and characteristics of parking. The key factor that influences the parking behavior is the parking price. In order to solve the strange phenomenon of ground parking in T central business district, which is especially crowded with less underground parking, a better balance between the driving customers' convenience of parking is taken into account for the needs of improving market margins. The measures to effectively solve the problem of parking business district, reducing the city road traffic capacity, reducing traffic accidents, and improving the quality of the city environment also plays a significant role. The results are of universal significance. Some problems facing the central business district have been put forward some corresponding measures and management recommendations which have very high reference value for other city commercial districts in parking field management.

References

1. Cheng T-x, Wang X, Zhang Y (2012) Empirical study on parking behavior in central commercial district of Tianjian. Res Dev World Sci Technol 34(5):851–856 (in Chinese)
2. Guan H-z (2000) Some issues in the parking facility development in central commercial area. Urban Plan 12:55–57 (in Chinese)
3. Huang R (2003) A study on parking problems and ideas of urban central commercial district of China. Xi'an University of Architecture and Technology,Xi'an, 2003 (in Chinese)

Chapter 65
Research on Structural Modeling Technique of Vehicles Outward Based on AVI

Bing Wang, Dudu Guo, Wenlei Sun and Gulbahar Tohti

Abstract This paper realized the reconstruction of vehicle complex curved surface through the application of reverse engineering method. This technology can establish 3D entity mode of vehicles outward which was helpful to identity and judge vehicles for the automatic vehicles identification area in the intelligent transportation system.

Keywords The automatic vehicle identification · Vehicles outward structure · Reverse engineering · Reconstruction technique · 3D entity model

65.1 Introduction

The automatic vehicle identification (AVI) of intelligent transportation system (ITS) was the general designation of technology on vehicles automatically identified when it passed through a particular place [1]. Vehicle identity usually refers to the inherent nature of the vehicle itself (including appearance structure) and other properties. Technologies such as license plate recognition, bar code and electronic license plate, electronic identification card and smart card were often used in the applications of AVI [2]. These technologies realized vehicle identification indirectly. If there was no on-board equipment, the method would be failure. At present, because of the difficulty to form the social demand of the AVI on-board equipment oriented to traffic management, research on developing new technology was kept on at home and abroad [3].

This paper intended to establish 3D entity model vehicle outward through the application of reverse engineering technology. The technology can directly

B. Wang · D. Guo (✉) · W. Sun · G. Tohti
Traffic Engineering Department, Xinjiang University, Urumqi 830000, China
e-mail: guodudu1122@126.com

L. Jia et al. (eds.), *Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)-Volume II*, Lecture Notes in Electrical Engineering 288, DOI: 10.1007/978-3-642-53751-6_65,
© Springer-Verlag Berlin Heidelberg 2014

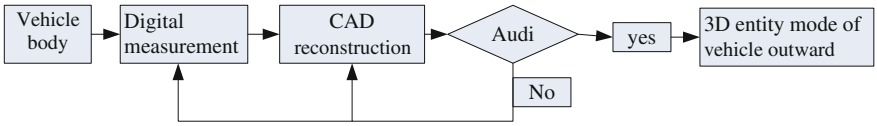


Fig. 65.1 The process of the design of vehicle structure outward modeling in reverse engineering

determine and identify the vehicles. It used the method of complex curved surface reconstruction in reverse engineering to realize the reconstruction of vehicle complex surface outward from point cloud data to the 3D solid model. This method combined the operation platform of reverse engineering with much software in the VC environment. The process of the design of vehicle structure outward modeling is shown in Fig. 65.1.

65.2 Technical Outline

The external structure of vehicle is complex. The reconstruction of complex curved surface still needs to use the new method to achieve the high quality of the vehicle outward of 3D entity model.

The method of complex curved surface reconstruction was based on Bezier of triangular domain and NURBS surface of rectangle domain [4]. Reconstructed results based on Bezier method of triangular domain cannot be accepted by most of the CAD software, which made this kind of reverse software to be equipped with module of the NURBS surface conversion [5]. During the process of conversion, it would inevitably lose information and produce error. What is more, the surface reconstruction condition of NURBS was high and it must deal with a large number of scattered data to fit, which made a large reconstruction error. Also, it was difficult to accurately reflect the original appearance of the vehicle.

The reconstruction method of this research was based on NURBS surface of triangular domain [6]. The method is a combination of the flexibility and accuracy of the Bezier surface reconstruction method based on triangular domain and the universality and the standard of the reconstruction results based on the rectangle domain. It can reconstruct the appearance of complex surface of vehicle high quality, be accepted by most of the CAD software, and solve the modeling problem of the fast surface of a large number of scattered data. Specific technical steps were as follows:

(1) This method first extracted the data information of the basic properties of the vehicle surface appearance using the raster scanning device, then preprocessed scattered data by measure and optimized data quality.
(2) It treated point cloud data with triangulation and reconstructed Bezier surface with G^1 continuous on the triangular mesh.

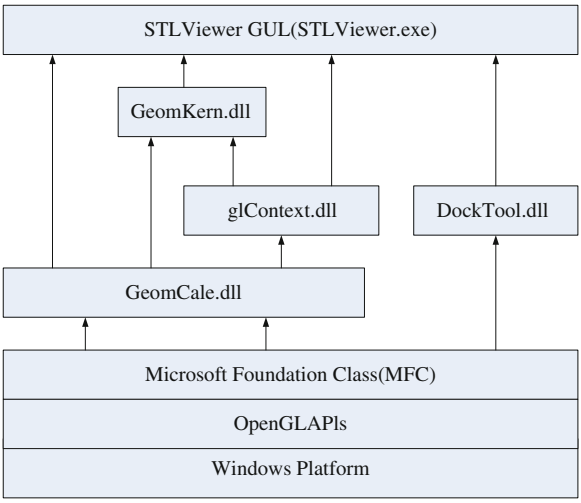


Fig. 65.2 Operation platform of the hierarchical structure of the system

(3) It extracted new characteristics of initial reconstructed surface and merged triangular domain, and used the transition surface Coon to reconstruct Surface on rectangle domain. A complete Coons surface was formed through surface smoothing and joining treatment.
(4) Coons surface was converted into NURBS surface using the CAGD relationship between Coon and NURBS surface.

65.3 Technical Methods

In this study the concrete method of the complicated surface reconstruction was as follows:

(1) An MFC single document/view project was created under the VC environment as the work platform of the reconstruction application system.
(2) In the platform, the reconstruction method was combined with the two reverse engineering software which was Imagware and RE-SOFT, and unified data in the two software interface through the exchange standard IGES format of the initial graphics.
(3) The file of "*. IMW" format of characteristics of curve and surface which were extracted by the two software was imported to the 3D modeling software UG NX to generate 3D entity.

The hierarchical structure of the operation platform of the application system is shown in Fig. 65.2. The technical route is shown in Fig. 65.3. The data process of the application system is shown in Fig. 65.4.

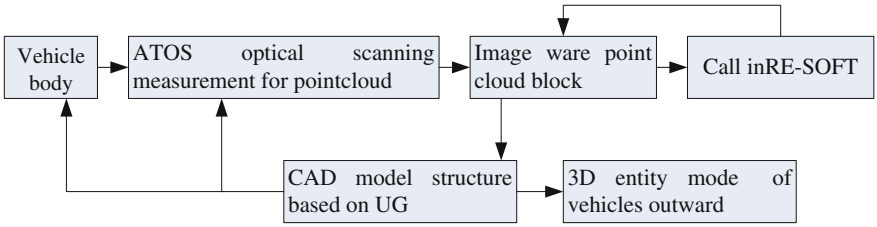


Fig. 65.3 The reconstruction of the application system

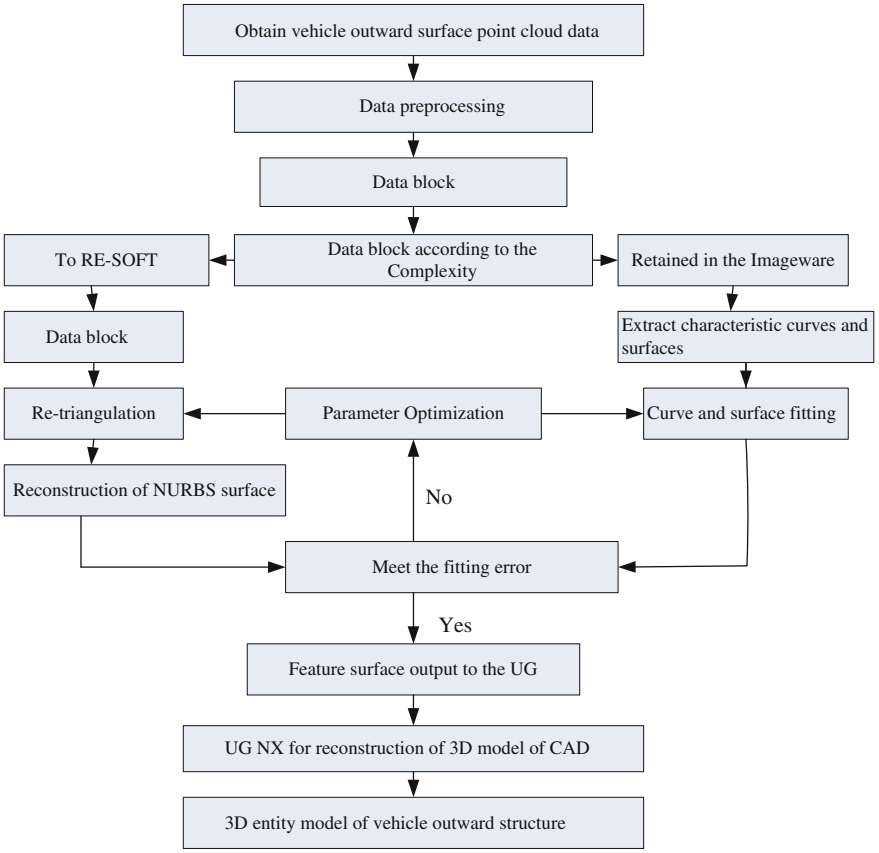


Fig. 65.4 Application data process between systems

65.4 Case

This paper described the full process of technology by setting the left anterior plate of the domestically produced vehicles as an example.

(1) The reconstruction of curve and surface of the point cloud data was realized by collecting surface data through optical scanning and preprocessing the data.
(2) The positioning and segmentation of point cloud data. The positioning of the point cloud data was to define a coordinate system on the point cloud data to make it coincide with the coordinate system in reverse engineering software. On the one hand, it provided size for the subsequent generated entity. On the other hand, the split point cloud data can be treated and does not need to care about the location them. The system itself will automatically splice them according to the coordinate system of the user. In addition, point cloud data of segmentation can be imported to other software of reverse engineering according to the positioning coordinate system. After processing, origin cloud model was reloaded and did not change the position. The segmentation of point cloud data mainly depended on the complexity of the point cloud data. The point cloud data of simple rules were retained in the reverse engineering software Imagware to reconstruct and the complex point cloud data was imported to RE—SOFT of the algorithm as a standard format of IGES to reconstruct. This method was convenient and flexible and can reconstruct a good NURBS surface.
(3) The NURBS surface constructed by step(2) in a unified IGES format was imported to Imagware and was merged with the curve and curved surface simple rules reconstructed in Imagware.
(4) The curve and surface extracted by the above steps in a "*. IMW" format file was imported UG which was the 3D modeling software to finally realize the 3D entity mode of high quality of complete left anterior plate of vehicle.

The process of the key technology is shown in Fig. 65.5.

65.5 Conclusion

In conclusion, this technology can establish 3D entity model with high quality of vehicle outside. The research was helpful to determine the identity of the vehicle in the field of AVI in ITS. It was the only way for ITS development by using multidisciplinary technology to realize various ITS projects [7]. With the rapid development of reverse engineering and deepening of the key technologies [8], it will be a great opportunity of reverse engineering for the ITS domain.

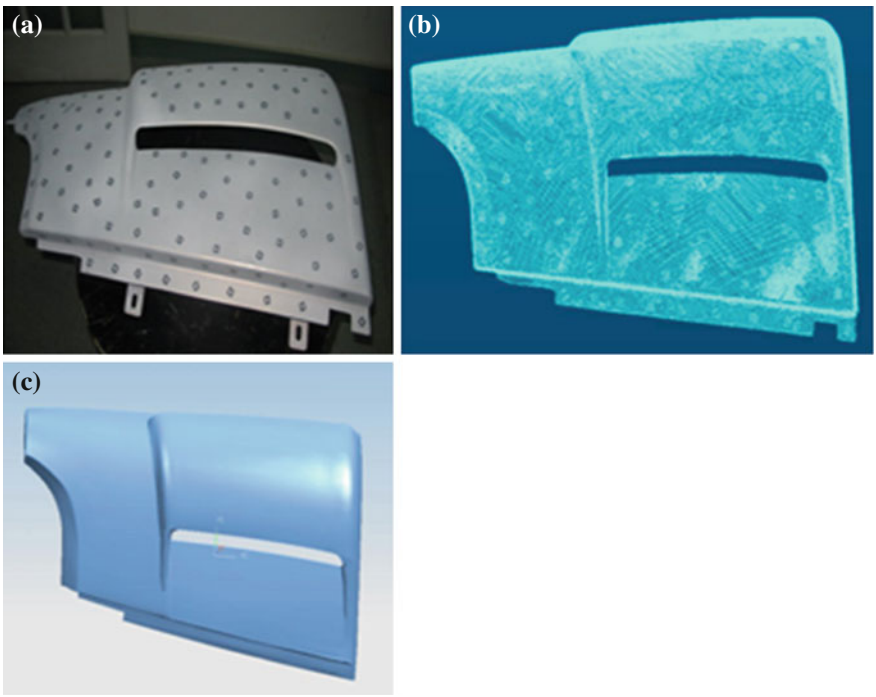


Fig. 65.5 The process of the key technology. **a** Sticker point cloud **b** Original point cloud **c** The 3D entity model of reconstruction

Acknowledgments This paper is supported by Key Project of Xinjiang Colleges and Universities Scientific Research (No.: XJEDU 2011I10).

References

1. "Research on intelligent transport system architecture in China" task force (2003). Research on intelligent transport system architecture in China. China Communications Press, Beijing (in Chinese)
2. Chen X (2007) Intelligent transportation system. China Railway Publishing House, Beijing (in Chinese)
3. Wang X (2004) Reverse engineering technology and application. Chemical Industry Press, Beijing (in Chinese)
4. Tao J, Dong S (2003) Reverse engineering technology. Machinery Industry Press, Beijing (in Chinese)
5. Ma J (2005) Reconstruction of the complex nonlinear system reconstruction technology. Tianjin University Press, Tianjin (in Chinese)
6. Liu H (2004) Reverse engineering technology and application. Chemical Industry Press, Beijing (in Chinese)

7. Xinpin Y (2007) Intelligent transport system-principle, method and application. Technology Press of Wuhan University, Wuhan (in Chinese)
8. Zhou J (2006) Research status and development trends of reverse engineering technology. Mod Manufact Technol Equip 2006(3):23–26 (in Chinese)