

Qiansheng Fang
Quanmin Zhu
Feng Qiao *Editors*

Advancements in Smart City and Intelligent Building

Proceedings of the International
Conference on Smart City and
Intelligent Building (ICSCIB 2018)



Advances in Intelligent Systems and Computing

Volume 890

Series Editor

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

Advisory Editors

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India

Rafael Bello Perez, Faculty of Mathematics, Physics and Computing, Universidad
Central de Las Villas, Santa Clara, Cuba

Emilio S. Corchado, University of Salamanca, Salamanca, Spain

Hani Hagras, Electronic Engineering, University of Essex, Colchester, UK

László T. Kóczy, Department of Automation, Széchenyi István University, Gyor,
Hungary

Vladik Kreinovich, Department of Computer Science, University of Texas
at El Paso, El Paso, TX, USA

Chin-Teng Lin, Department of Electrical Engineering, National Chiao Tung
University, Hsinchu, Taiwan

Jie Lu, Faculty of Engineering and Information Technology, University
of Technology Sydney, Sydney, NSW, Australia

Patricia Melin, Graduate Program of Computer Science, Tijuana Institute
of Technology, Tijuana, Mexico

Nadia Nedjah, Department of Electronics Engineering, University of Rio de
Janeiro, Rio de Janeiro, Brazil

Ngoc Thanh Nguyen, Faculty of Computer Science and Management, Wrocław
University of Technology, Wrocław, Poland

Jun Wang, Department of Mechanical and Automation Engineering, The Chinese
University of Hong Kong, Shatin, Hong Kong

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

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

More information about this series at <http://www.springer.com/series/11156>

Qiansheng Fang · Quanmin Zhu ·
Feng Qiao
Editors

Advancements in Smart City and Intelligent Building

Proceedings of the International Conference
on Smart City and Intelligent Building
(ICSCIB 2018)



Springer

Editors

Qiansheng Fang
Anhui Jianzhu University
Hefei, China

Quanmin Zhu
University of the West of England
Bristol, UK

Feng Qiao
Shenyang Jianzhu University
Shenyang, China

ISSN 2194-5357

ISSN 2194-5365 (electronic)

Advances in Intelligent Systems and Computing

ISBN 978-981-13-6732-8

ISBN 978-981-13-6733-5 (eBook)

<https://doi.org/10.1007/978-981-13-6733-5>

Library of Congress Control Number: 2019932686

© Springer Nature Singapore Pte Ltd. 2019

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

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

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

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

The 2018 International Conference On Smart City and Intelligent Building

ICSCIB 2018

September 15–16, 2018, Hefei, China.



Organized by

Anhui Jianzhu University, Hefei, China

Sponsored by

University of the West of England, Bristol, UK
Shenyang Jianzhu University, Shenyang, China
Tsinghua University, Beijing, China
Dalian University of Technology, Dalian, China
Tongji University, Shanghai, China
Xi'an University of Architecture and Technology, Xi'an, China
Suzhou University of Science and Technology, Suzhou, China
Anhui Province Key Laboratory of Intelligent Building and Building Energy, Hefei, China

Organizer, Sponsors and Supporters

Organizer



Sponsors



Supporters



Conference Organizing Committee

General Chairs

Prof. Qiansheng Fang, Anhui Jianzhu University, China

Prof. Quanmin Zhu, University of the West of England, UK

Advisory Chair

Prof. Jiaping Liu, Xi'an University of Architecture and Technology, China

Members of Advisory Committee

Prof. Qianchuan Zhao, Tsinghua University, China

Dr. Zhaohui Liu, Chinese Society for Urban Studies, China

Prof. Runming Yao, University of Reading, UK

Organizing Chair

Prof. Yalong Yang, Anhui Jianzhu University, China

Program Chairs

Prof. Feng Qiao, Shenyang Jianzhu University, China

Dr. Yinong Chen, Arizona State University, USA

Publicity Chair

Prof. Zhenya Zhang, Anhui Jianzhu University, China

Publication Chair

Prof. Jing Na, Kunming University of Science and Technology, China

Special and Invited Session Chairs

Prof. Junqi Yu, Xi'an University of Architecture and Technology, China

Prof. Jiugen Zhang, Nanjing Tech University, China

Prof. Baochuan Fu, Suzhou University of Science and Technology, China

Student Activity Chair

Mr. Changcheng Sun, Anhui Jianzhu University, China

Registration Chair

Mr. Chenlei Xie, Anhui Jianzhu University, China

General Secretary

Ms. Mingyue Wang, Anhui Jianzhu University, China

International Program Committee

Chairs of IPC

Prof. Feng Qiao, Shenyang Jianzhu University, China

Dr. Yinong Chen, Arizona State University, USA

Members of IPC

Ahmad Taher Azar, Benha University, Egypt
C. Benim, Duesseldorf University of Applied Sciences, Germany
Yinong Chen, Arizona State University, USA
Zhixin Chen, Beijing University of Civil Engineering and Architecture, China
Lei Cheng, Wuhan University of Science and Technology, China
Ramon Costa-Castelló, Universitat Politècnica de Catalunya, Spain
Zhihua Cui, Taiyuan University of Science and Technology, China
Mingcong Deng, Tokyo University of Agriculture and Technology, Japan
Mark Dooner, University of Warwick, UK
Peiyong Duan, Shandong Normal University, China
Fuyan Guo, Tianjin Chengjian University, China
Lingzhong Guo, University of Sheffield, UK
Guowen Hu, Yancheng Institute of Technology, China
Chengdong Li, Shandong Jianzhu University, China
Jiejia Li, Shenyang Jianzhu University, China
Kang Li, Queen's University Belfast, UK
Jian Liu, Shenyang Jianzhu University, China
Shaowen Lu, Northeastern University, China
Hongbin Ma, Beijing Institute of Technology, China
Yingying Ma, South China University of Technology, China
Yue Ma, Beijing Institute of Technology, China
Jing Na, Kunming University of Science and Technology, China
Mokhtar Nibouche, University of the West of England, UK
Qingchang Ren, Xi'an University of Architecture and Technology, China
Ahmed Rhif, University of Carthage, Tunisia
Liangliang Sun, Shenyang Jianzhu University, China
Georgios Theodoropoulos, Southern University of Science and Technology, China
Hesheng Wang, Shanghai Jiaotong University, China
Jihong Wang, University of Warwick, UK
John Wang, Montclair State University, USA
Yahui Wang, Beijing University of Civil Engineering and Architecture, China
Liming Wei, Jilin Jianzhu University, China
Sundarapandian Vaidyanathan, Vel Tech University, India
Hui Xiao, Tongji University, China
Lai Xu, Bournemouth University, UK
Hongji Yang, Bath Spa University, UK
Xuelin Yang, Shanghai Jiaotong University, China
Jing Yong, Chongqing University, China
Junqi Yu, Xi'an University of Architecture and Technology, China
Shuai Yuan, Shenyang Jianzhu University, China
Yong Yue, University of Bedfordshire, UK
Jiugen Zhang, Nanjing Tech University, China

Yijun Zhang, Nanjing University of Science and Technology, China

Xiaofang Zheng, East China Jiaotong University, China

Yuguo Zhou, Qingdao University of Technology, China

Invited Speakers

Prof. Yi Jiang, Department of the Energy Conservation Research Center on Building, Tsinghua University, Beijing, China

Prof. Antonio Puliafito, Department of Computer Engineering, The University of Messina, Italy

Prof. Abbas Elmualim, School of Natural Built Environments, University of South Australia, Adelaide, Australia

Prof. Qianchuan Zhao, Department of Automation, Tsinghua University, Beijing, China

Preface

Urbanization is one of the major trends nowadays around the world, more and more people rush into the urban area to seek better and more opportunities for career and wealth, and the number of population in cities has increased exponentially. It is predicted that the urban population will be more than two-thirds of the world's population by 2050 with more megacities which inhabit more than 10 million residents. The economic and social development in a country or a region is an image of the development of its cities. Well-planned, implemented, and managed cities provide good chance to access close proximity, cultural diversity, and commercial competition. The rapid population growth in cities challenges urban living condition and environment. The future cities will need to become smarter to offer more efficient, sustainable, and livable environments for inhabitants by managing existing and imminent infrastructures and resources.

This book is the collection of papers from the 2018 International Conference on Smart City and Intelligent Building (ICSCIB 2018), held in Hefei, Anhui, China, on September 15–16, 2018. ICSCIB 2018 was organized by the Anhui Jianzhu University (AHJZU), China, and sponsored by the University of the West of England, UK, and Shenyang JianZhu University, Shenyang, China; it is supported by the International Journal of Simulation and Process Modelling (IJSPM) from Inderscience Publishers, as well as, Tsinghua University, Beijing, China, Dalian University of Technology, Dalian, China, Tongji University, Shanghai, China, Xi'an University of Architecture and Technology, Xi'an, China, Suzhou University of Science and Technology, Suzhou, China, Anhui Province Key Laboratory of Intelligent Building and Building Energy, Hefei, China.

ICSCIB 2018 provided an international forum for professionals, academics, and researchers to present the latest developments from interdisciplinary theoretical studies, computational algorithm developments, and engineering applications in smart cities and smart buildings. This academic event featured many opportunities to network with colleagues from around the world in a wonderful environment. Its program covered invitation and presentations from scientists, researchers, and practitioners, who were working in similar areas to establish platforms for collaborative research projects in these fields. The conference invited leaders from

industry and academia to exchange and share their experiences, present research results, explore collaborations, and to spark new ideas, with the aim of developing new projects and exploiting new technology in these fields, and bridge theoretical studies and practical applications in all science and engineering branches.

At the conference, we invited four distinguished scholars, in Smart City and Intelligent Building, to make plenary speeches. They were Prof. Yi Jiang, Academician of the Academy of Engineering of China, from Tsinghua University, Beijing, China; Prof. Antonio Puliafito, from the University of Messina, Italy; Prof. Abbas Elmualim, from University of South Australia, Adelaide, Australia; Prof. Qianchuan Zhao, from Tsinghua University, Beijing, China. They gave speeches on “Insect Intelligence in Buildings & City”, “Software Defined Cities: Integrating the Cyber World with the Internet of Things”, “Smart and Livable Cities: the case of Adelaide, Australia”, “A new architecture of building control systems based on Internet of Things (IoT)”. Their speeches addressed the state-of-art development and the cutting-edge research topics in both theory and practical application.

This edition of ICSCIB 2018 covers a wide range of research areas in smart city and intelligent building. The topics of the papers in the volume covered the fields of smart city and intelligent building included smart traffic and transportation, smart energy storage and usage, emergency management systems, healthcare/service/administrative/public sector services, smart communities and campuses, big data and cloud computing in smart cities, smart utility management systems, ICT-based smart services, underground pipeline, and space-integrated management and smart infrastructures, it also dealt with the issues in intelligent buildings of building comfort and energy saving, assembly of buildings, building information modeling, building management systems, building automation systems and HVAC, construction robot and automation, IoT in smart home, lighting, fire and safety control systems, intelligent and smart home devices. The papers in this volume are categorized into 8 tracks in smart city and intelligent building as follows:

1. Building Energy Efficiency,
2. Construction Robot and Automation,
3. Intelligent Community and Urban Safety,
4. Intelligentization of Heating Ventilation Air Conditioning System,
5. Information Technology and Intelligent Transportation Systems,
6. New Generation Intelligent Building Platform Techniques,
7. Smart Home and Utility, and
8. Smart Underground Space.

Feng Qiao

Program Chair of ICSCIB 2018

Conference Website: <http://ICSCIB2018.ahjzu.edu.cn>

Acknowledgements

This edition of ICSCIB 2018 gathered 58 papers from about 100 submissions to the 2018 International Conference on Smart City and Intelligent Building (ICSCIB 2018), held in Hefei, Anhui, China, on September 15–16, 2018. ICSCIB 2018 was attended by more than 130 participants of scholars, researchers, and engineers in the areas of smart cities and intelligent buildings. On behalf of the Organizing Committee and the Program Committee of the conference, we are sincerely thankful to the authors and contributors to the conference who gave their full support and encouragement to this event.

We are thankful to the Editor-in-Chief of the Springer series on Advances in Intelligent Systems and Computing (AISC) and the staff from Springer, who have supported and helped to bring out the Proceedings of 2018 International Conference on Smart City and Intelligent Building.

We are very grateful to the famous scholars in smart city and intelligent building, who gave excellent keynote speeches on the conference, they are Prof. Yi Jiang, Academician of the Academy of Engineering of China, from Tsinghua University, Beijing, China; Prof. Antonio Puliafito, from the University of Messina, Italy; Prof. Abbas Elmualim, from University of South Australia, Adelaide, Australia; Prof. Qianchuan Zhao, from Tsinghua University, Beijing, China.

We would like to honor our members of the Organizing Committee to have done their best for their invaluable contribution in arranging the diverse offerings at the conference.

We are very grateful of the members of Program Committee and all of the anonymous reviewers worldwide, who helped to review and comment on the submitted papers. And we are thankful to the volunteers who kindly supported the conference and assisted the conference to move smoothly.

As the editors of this edition of ICSCIB 2018, we hope that the papers in this volume will stimulate further research and development in smart city and intelligent building, and the readers will get great help from this volume of AISC series on Advancements in Smart City and Intelligent Building.

Hefei, China
Bristol, UK
Shenyang, China

Qiansheng Fang
Quanmin Zhu
Feng Qiao

About This Book

During the rapid expansion of the urbanized area around the world, smart city and intelligent building has been one of the hottest research topics in academic research and engineering practice in the past decade.

This book gathers the Proceedings of the International Conference on Smart City and Intelligent Building (ICSCIB 2018), held in Hefei, Anhui, China, on September 15–16, 2018. It includes 58 papers from about 100 submissions for 8 tracks in the topics on “Building Energy Efficiency”, “Construction Robot and Automation”, “Intelligent Community and Urban Safety”, “Intelligentization of Heating Ventilation Air Conditioning System”, “Information Technology and Intelligent Transportation Systems”, “New Generation Intelligent Building Platform Techniques”, “Smart Home and Utility”, and “Smart Underground Space”, which cover a wide range in the area of smart cities and intelligent buildings.

This edition of ICSCIB 2018 covers the wide range in smart cities of smart traffic and transportation, smart energy storage and usage, emergency management systems, smart communities and campuses, big data and cloud computing in smart cities, smart utility management systems, ICT-based smart services, underground pipeline and space-integrated management and smart infrastructures, it also deals with the issues in intelligent buildings of building comfort and energy saving, assembly of buildings, building information modeling, building management systems, building automation systems and HVAC, IoT in smart home, lighting, fire and safety control systems, intelligent and smart home devices.

Contents

Part I Building Energy Efficiency	
Study on the Control Method of Temperature and Humidity Environment in Building Intelligent System	3
Kuan Huang, Haolin Song and Hongrui Fu	
Research on Human Thermal Comfort Model Based on Multiple Physiological Parameters	11
Yalong Yang, Wenmiao Wu, Qiansheng Fang, Xulai Zhu, Rui Zhang and Mingyue Wang	
Study on Thermal Comfort Model Based on Genetic Algorithm with Backpropagation Neural Network	23
Yalong Yang, Dejian Hong, Rui Zhang, Qiansheng Fang, Xulai Zhu and Wenmiao Wu	
An Optimization Model for PV and CCHP-Supplied Power System in Buildings	35
Jin Zhao and Jing Yong	
Control and Optimization of Indoor Environmental Quality Based on Model Prediction in Building	45
Anjun Zhao, Meng Zhou, Junqi Yu, Junlin Zhang and Xiong Yang	
Identifying Abnormal Energy Consumption Data of Lighting and Socket Based on Energy Consumption Characteristics	59
Liangdong Ma, Yiying Xu, Yugen Qin and Jili Zhang	
Part II Construction Robot and Automation	
An Improved Weight Control System for Slender Cigarette Production	75
Zhonghua Han, Xu Yang, Kaiyuan Bi, Xiaoting Dong and Xixian Sun	

Design of Building Environment Mobile Monitoring and Safety Early Warning Robot	85
Guoqing Yang, Yuhao Wang and Bing Chen	
Application of Probabilistic Reasoning Algorithm in Indoor Positioning Based on WLAN	95
Meng Li and Honglin Wang	
Multiple Rotorcrafts Environment Map Fusion for Atmosphere Monitoring	101
Pengxiang Bao, Lei Cheng, Xin Wang, Qin Liu and Qiuyue Yu	
Simulated Tests of Feedforward Active Noise Control (ANC) for Building Noise Cancellation	115
Tongrui Peng, Quanmin Zhu, M. Osman Tokhi and Yufeng Yao	
Research on Production Layout Design of Concrete Prefabricated Units Based on SLP	125
Yang Liu and Ding Li	
 Part III Intelligent Community and Urban Safety	
Blockchain in Smart City Development—The Knowledge Governance Framework in Dynamic Alliance	137
Yi Zhang, Wei Sun and Chenlei Xie	
Attendance and Security System Based on Building Video Surveillance	153
Kailai Sun, Qianchuan Zhao, Jianhong Zou and Xiaoteng Ma	
Simulation Study on Collaborative Evacuation Among Stairs and Elevators in High-Rise Building	163
Xiaodong Liu, Hui Zhang and Ping Zhang	
Dynamic Emergency Evacuation System for Large Public Building	173
Yongming Zhang, Zhe Yan, Xueli Zhu and Wenjie Piao	
Status of Intelligent Building Development of China—Questionnaire Analysis	183
Huai Li, Zhen Yu and Wei Liu	
 Part IV Intelligentization of Heating Ventilation Air Conditioning System	
Regression Model of Wet-Bulb Temperature in an HVAC System	197
Luping Zhuang, Xi Chen and Xiaohong Guan	

Research on Optimal Control Algorithm of Ice Thermal-Storage Air-Conditioning System	207
Junqi Yu, Xiong Yang, Anjun Zhao, Meng Zhou and Yanhuan Ren	
Decentralized Optimization Algorithm for Parallel Pumps in HVAC Based on Log-Linear Model	219
Junqi Yu, Xuegen Qian, Anjun Zhao, Shiqiang Wang and Qite Liu	
Partial Fault Detection of Cooling Tower in Building HVAC System	231
Liangliang Sun, Haiqi Jia, Hang Jin, Ye Li, Junning Hu and Congxin Li	
Application of Information Network and Control Network Integration Technology in Central Air Conditioning Data Management System	241
Yan Bai, Zhengmin Liu and Qingchang Ren	
MFAC and Parameter Optimization for a Class of Models in HVAC	249
Zengxi Feng, Junqi Yu, Zhongtian Rao and Anjun Zhao	
A Controller Algorithm (ILC) for the Variable Differential Pressure Control of Freezing Water in a Central Air Conditioning System	261
Qingchang Ren and Hongmei Jiang	
The Online Evaluation System of Chiller Plant in HVAC System	271
Jiaming Wang, Tianyi Zhao and Wei Li	
The Power Consumption Model of Chiller with Elman Neural Networks for On-line Prediction and Control	281
Zhiyang Jia and Tianyi Zhao	
 Part V Information Technology and Intelligent Transportation Systems	
Research on Driving Decisions in Winter and Summer Based on Survey Date	293
Mingxia Huang, Haiqiang Zhang and Zhu Bai	
The Performance Evaluation and Improvement of Urban Taxi Firms Using Data Envelopment Analysis and Benchmarking Approach	301
Zhu Bai and Shuai Bian	
Research on the DV-Hop Location Algorithm Based on the Particle Swarm Optimization for the Automatic Driving Vehicle	309
Pei Huang, Xinjian Xiang and Bingqiang Huang	

Multi-objective Optimization Coordination for Urban Arterial Roadway Based on Operational-Features 321
Feng Qiao, Haochen Sun, Lingzhong Guo, Haolin Song
and Zhaoyan Wang

Applicability Analytic of Closed Intersection Along Tramway Based on Simulation 333
Weidong Liu, Xingquan Guan, Yixiu Wu and Shanshan Fan

Vehicle Scheduling Optimization of Urban Distribution Considering Traffic Control 343
Qianqian Shao, Huifei Li, Yunfeng Zhang and Feng Guan

The Warrant of Slip Lane at Single-Lane Roundabout 353
Yuzhou Duan, Hui Li and Yabo Song

A New Method for the Minimum Concave Cost Transportation Problem in Smart Transportation 363
Chuan Li, Zhengtian Wu, Baochuan Fu, Chuangyin Dang and Jinjin Zheng

Study on the Characteristics of Vehicle Lane-Changing in the Intersection 371
Yan Xing, Jinling Wang, Weidong Liu, Lishuang Sun and Faguang Chong

Part VI New Generation Intelligent Building Platform Techniques

A P2P Algorithm for Energy Saving of a Parallel-Connected Pumps System 385
Qianchuan Zhao, Xuetao Wang, Yifan Wang, Ziyang Jiang
and Yunchuang Dai

A Distributed Algorithm for Building Space Topology Matching 397
Yifan Wang and Qianchuan Zhao

Decentralized Differential Evolutionary Algorithm for Large-Scale Networked Systems 407
Guanghong Han, Xi Chen and Qianchuan Zhao

Intelligent Building Fault Diagnosis Based on Wavelet Transform and Bayesian Network 419
Jundong Fu, Luming Huang, Li Chen and Yunxia Qiu

Fault Location of Distribution Network for Wavelet Packet Energy Moment of Dragonfly Algorithm 433
Jundong Fu, Jinglin Yue, Li Chen and Tianhang Leng

Graphical Programming Language Design for Decentralized Building Intelligent System 447
Shuo Zhao, Qiliang Yang, Jianchun Xing and Guangtong Xue

Insect Intelligent Building (I²B): A New Architecture of Building Control Systems Based on Internet of Things (IoT)	457
Qianchuan Zhao and Ziyang Jiang	

Part VII Smart Home and Smart Utility

The Fault Diagnosis Model Established Based on RVM	469
Yahui Wang, An Yun, Qinghong Ye and Yunfeng Zhao	

PLC-Based Intelligent Home Control System	477
Liming Wei, Yangyun Wu and Xiujuan Guo	

A Classification-Based Occupant Detection Method for Smart Home Using Multiple-WiFi Sniffers	485
Ping Wang, Huaqian Cao, Si Chen, Jiake Li, Chang Tu and Zhenya Zhang	

A p-Persistent Frequent Itemsets with 1-RHS Based Correction Algorithm for Improving the Performance of WiFi-Based Occupant Detection Method	497
Ping Wang, Huaqian Cao, Si Chen, Jiake Li, Chang Tu and Zhenya Zhang	

Day-Ahead Short-Term Optimization of Renewable Energy of Microgrid in Multiple Timescales	505
Xiaohui Wang and Shiqi Zong	

Modeling of Multiple Heating Substations Based on Long Short-Term Memory Networks	515
Qi Li, Bingcheng Han, Mingwei Yu and Jianglan Shang	

The Elman Network of Heat Load Forecast Based on the Temperature and Sunlight Factor	525
Qi Li, Shiqi Jiang and Xudan Wu	

Theoretical Study on Even Heating of Single Pipe Heating System	537
Xiaoli Yin, Mingsheng Liu, Zhixian Ma and Jili Zhang	

Illumination Variation Similarity Based Fault Diagnosis for HV-LED Lamp Driven by Segmented Linear Driver	547
Fukang Sun, Shaofeng Zhu and Ye Wang	

Point Illumination Calculation Method in Special-Shaped Space	557
Jundong Fu, Qing Chen, Yunxia Qiu and Li Chen	

Part VIII Smart Underground Space

Device-Free Activity Recognition for Underground Spaces Based on Convolutional Neural Network	573
Qizhen Zhou, Jianchun Xing, Xuwei Zhang and Wei Chen	

A Decentralized Parallel Kalman Filter in Multi-sensor System for Data Verification	583
Guoping Li and Shiqiang Wang	
DXF File Topological Information Extraction and Storage for Decentralized Distribution Network	593
Yuhan Zhang and Shiqiang Wang	
Research on Underground Device Operation and Maintenance Management System Based on BIMserver	601
Mengli Ding, Qiliang Yang, Jianchun Xing and Liqiang Xie	
A Fully Distributed Genetic Algorithm for Global Optimization of HVAC Systems	609
Shiqiang Wang, Jianchun Xing and Juelong Li	
Open-Neutral Fault Detection in Underground Space Based on Genetic Support Vector Machine	621
Xuechen Zhao, Ping Wang and Jianchun Xing	
Author Index	631

Part I
Building Energy Efficiency

Study on the Control Method of Temperature and Humidity Environment in Building Intelligent System



Kuan Huang, Haolin Song and Hongrui Fu

Abstract Due to the difficulty of establishing the accurate control model for building an intelligent system, a neural network predictive control method is proposed, in this paper, based on a weed optimization algorithm. Through considering indoor temperature and relative humidity environment factors, a control model of temperature and humidity environment is first established in an intelligent building. Then, the hidden layer nodes center of the RBF neural network is optimized by using the weed optimization algorithm. The above mentioned work focuses on improving the shortcomings of Orthogonal Least Squares (OLS) algorithm, and simultaneously simplifies the network architecture. The simulation results show that the RBF neural network predictive control method based on the weed optimization algorithm has better approximation ability and generalization ability contrasting with the OLS algorithm.

Keywords Radial basis function · Weed optimization algorithm · Neural network node centers · Building intelligent system

1 Introduction

Building an intelligent system is a nonlinear complex large system with multivariable and strong coupling. The control system has many uncertainties in the operation process, thus it is difficult to establish an accurate control model. Therefore, it is absolutely essential to study a control method that can overcome the uncertainty of the control object model; meanwhile, there is a good approximation ability and generalization ability. In view of the simple architecture of the RBF neural network, it has good nonlinear function approximation property and higher training rate, which can be adapted to the dynamic characteristics of an uncertain system. Therefore, we can

K. Huang (✉) · H. Song · H. Fu
Faculty of Information and Control Engineering, Shenyang Jianzhu
University, Shenyang, China
e-mail: huangkuan@sjzu.edu.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_1

combine the RBF neural network predictive control method with the actual application background of temperature and humidity environment of an intelligent building to carry on the preliminary research, and then to better solve the abovementioned problems.

In RBF neural network training algorithm, Yu et al. [1] proposed the Orthogonal Least Squares Algorithm (OLS algorithm), and applied it in the prediction process of the indoor temperature model of a greenhouse. OLS algorithm is easily achieved and the number of hidden layer nodes can be determined in the process of the weights training. However, Sherstinsky and Picard [2] carried out a systematic analysis for algorithm from the perspective of energy compression and they found they OLS algorithm is not necessarily can be achieved the most simplified network architecture. To solve this problem, a radial basis function (RBF) neural network predictive control method based on weed optimization algorithm is proposed in this paper.

2 Radial Basis Function (RBF) Neural Network and Orthogonal Least Squares Algorithm (OLS)

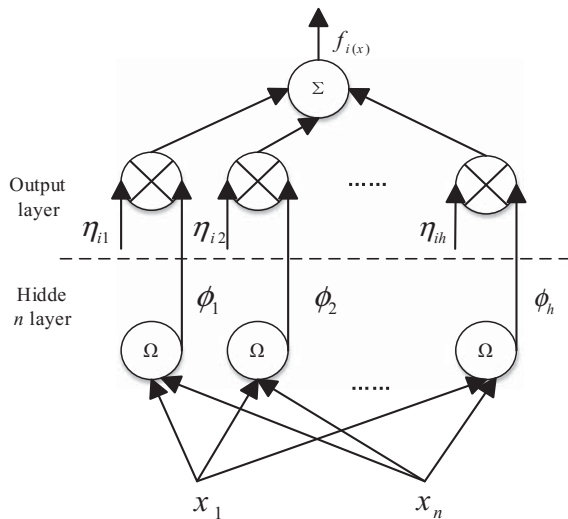
At the end of the 1980s, the Radial Basis Function Neural Network (RBF-NN) was proposed by J. Moondy and C. Darken, et al. It is a forward architecture neural network which is based on the localized reactions of human brain neuron cells to the external stimuli. The RBF network architecture is shown in Fig. 1. RBF network architecture can be divided into two levels, respectively, the hidden layer and the output layer. Hidden layer contains a number of cell nodes, and each cell node is corresponding to a parameter vector, which called the node center [3–5]. Through a certain nonlinear function mapping relationship, the Euclidean metric between the network input vector and the center parameter vector can be calculated by each node, which is sent to the output layer. Then, the value of the output layer node can be calculated in the linear combination of the output function value of each hidden layer node by the output layer [6].

In the implementation process of the RBF network, the input and output mapping $f : R^n \rightarrow R^m$, can be expressed as follows:

$$f_{i(x)} = \sum_{j=1}^h \eta_{ij} \phi_j = \sum_{j=1}^h \eta_{ij} \phi(\|x - c_j\|, \rho_j) \quad (1)$$

where h is the number of nodes in the hidden layer; $i = 1, 2, \dots, m$; m is the number of nodes in the output layer; η_{ij} is the transmission weight from node J in hidden layer to node I in the output layer; Input vector $X = [x_1 \ x_2 \ \dots \ x_n]^T$; c_j is the center value of the hidden layer node, $\|\dots\|$ is the Euclidean norm, representing the Euclidean metric between input vector X and hidden layer node J ; ρ_j is the

Fig. 1 RBF network architecture



normalized parameter value; $\phi(\cdot, \rho)$ is an unknown nonlinear function, using the Gaussian distribution function here:

$$\phi(Z, \rho) = \exp(-Z^2/\rho^2) \quad (\rho > 0) \quad (2)$$

The column vectors of the output matrices of the hidden layer are not orthogonal, And it has a mutual influence for the contribution value of output. So OLS algorithm can achieve a smaller network architecture, however, it cannot guarantee that the network architecture is the most simplified [7]. If the sequence of orthogonalization is changed, then the contribution value of the node centers will also be changed. The approximation error of RBF neural network is decided by the selection of the node centers in the hidden layer, and has nothing to do with the selection order of that. Therefore, the selection of the node centers of the RBF neural network can be considered as a route selection problem, and the optimal solution can be found by a suitable optimization method.

3 Weed Optimization Algorithm Based on RBF Neural Network Node Centers

Weed optimization algorithm is a novel numerical optimization algorithm based on the population. Its operation process simulates the complete process of weed invasion, which has the characteristics of simple architecture, strong robustness, a few influencing factors, etc. In the operation process of weed optimization algorithm, the offspring will grow and spread around the parent individuals by the normal

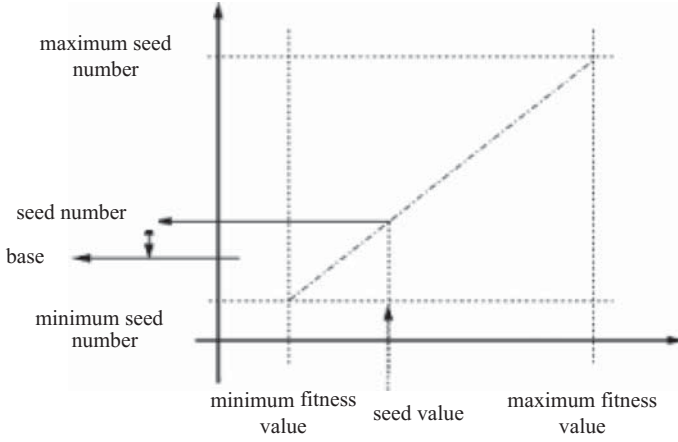


Fig. 2 Method for determining the number of new nodes

distribution. Meanwhile, the global search and the local search are carried out, and both weights can be adjusted according to the difference of iteration times. The global optimal solution is finally searched in the solution space of the node centers in a hidden layer. The weed optimization algorithm based on the RBF neural network node centers is introduced as follows:

- (1) Population initialization: The number of node centers (weeds) in the initial population was set to be N , and be randomly distributed in the D -dimensional sample space.
- (2) Growth and reproduction: Each node center will continue to evolve and generate new individuals (weed seeds) according to their fitness. A certain linear relationship between the number of new nodes (weed seeds) by parent nodes and the fitness of the parent is as follows:

$$N_s = \frac{f - f_{\min}}{f_{\max} - f_{\min}}(s_{\max} - s_{\min}) + s_{\min} \quad (3)$$

where N_s is the number of new nodes $n(n = 1, 2, \dots, N)$ generated by the node centers; $f_{\max}, f_{\min}$ is the best and the worst fitness of the current population; $s_{\max}$ and $s_{\min}$ are the set to be the maximum and minimum number of new nodes, respectively; f is the fitness of current node centers.

The method for determining the number of new nodes is shown in Fig. 2.

- (3) Spatial diffusion: The individual of the new node takes the individual of the node centers as the central axis, using the normal distribution to evolve and spread continuously in the D -dimensional space. In the specific iterative process, the standard deviation of each generation must conform to the following rules:

$$\sigma_{iter} = \frac{(iter_{\max} - iter)^n}{(iter_{\max})^n} (\sigma_{initial} - \sigma_{final}) + \sigma_{final} \quad (4)$$

where σ_{iter} is the standard deviation of the iteration, $\sigma_{initial}$ is the initial standard deviation, σ_{final} is the final standard deviation, $iter_{\max}$ is the maximum number of iterations, n is the nonlinear harmonic index.

- (4) Competitive exclusion: The fitness function is $f = 1/d + K_\varepsilon$, d is the number of node centers, ε is the approximation precision, K is real number, usually $0 \leq K_\varepsilon < 1$. In this process, when the number of node centers is larger, the generalization ability and approximation ability of the network is worse, and its fitness is lower.
- (5) Repeat steps 2–4 until the requirements of optimal solution are achieved or the number of iterations achieves the maximum, the whole operation process is stopped.

The training flow chart of RBF neural network based on weed optimization algorithm is shown in Fig. 3.

4 Simulation and Analysis

4.1 Data Acquisition

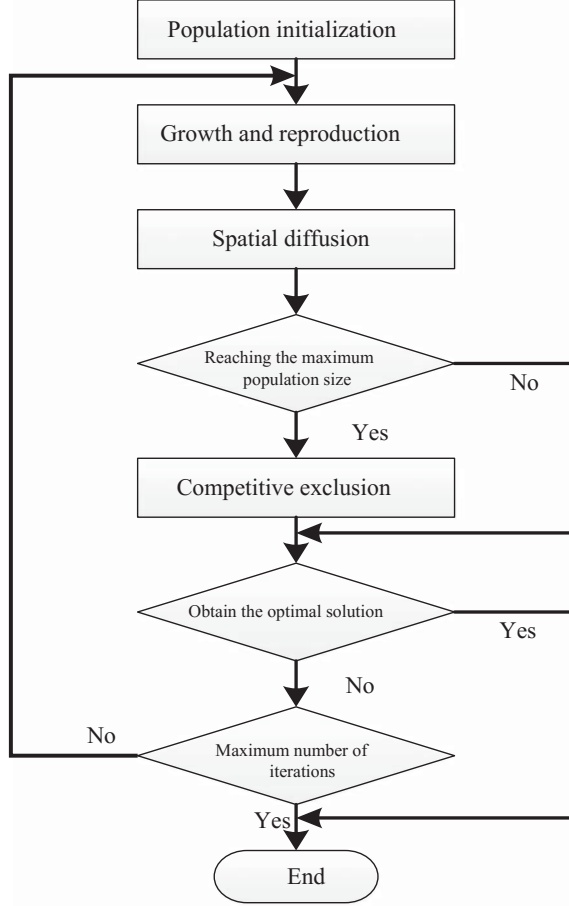
In order to verify the effectiveness of the proposed method, the indoor temperature and humidity environment control system of an intelligent building in our province is as an example. The data acquisition time is from 0: 00 on June 5, 2014 to 00: 00 on June 10 and a sampling time interval is 10 min. Taking 150 m² of residential as sampling objects¹, the outdoor temperature, wind speed, the degree of a window opening, solar irradiance, indoor relative humidity, indoor temperature and other 720 sets of data. Other data, such as air conditioning state data, is not considered due to no operation is performed. The obtained data will be uniformly normalized transformation according to the following equation:

$$y_i = \frac{1}{2} \frac{x_i - \bar{x}}{x_{\max} - x_{\min}} \quad (5)$$

where x_i is the measured value; $\bar{x}$ is the arithmetic mean; $x_{\max}$ is the maximum; $x_{\min}$ is the minimum; the obtained variation range for value y is $[-0.5, +0.5]$.

¹The simulation study of this research work has been conducted with approval obtained from the owner of the residential in Xiangfenghuayuan, Huanggu District, Shenyang, China.

Fig. 3 RBF neural network learning flow chart based on weed optimization algorithm



4.2 Simulation Results

The input parameters of building an indoor environmental model include the outdoor temperature, wind speed, the degree of the window opening, solar irradiance, etc. The output parameters include indoor temperature and humidity. The first 400 sets of collected data are taken as the training data sets, and the remaining data is used as the test data sets. Some of the main parameters data in the simulation study are as follows: $D = 2$, $N = 10$, $P_{max} = 100$, $N_{max} = 5$, $N_{min} = 0$, $\sigma_{initial} = 3$, $\sigma_{final} = 0.001$, $n = 3$, $iter_{max} = 25$. The actual curve is compared, respectively, with the simulation fitting curve of the OLS algorithm and the method based on the weed optimization algorithm. The details are as follows.

Fitting curve of the intelligent building temperature and humidity environment based on the OLS algorithm and the weed optimization algorithm is shown in Fig. 4.

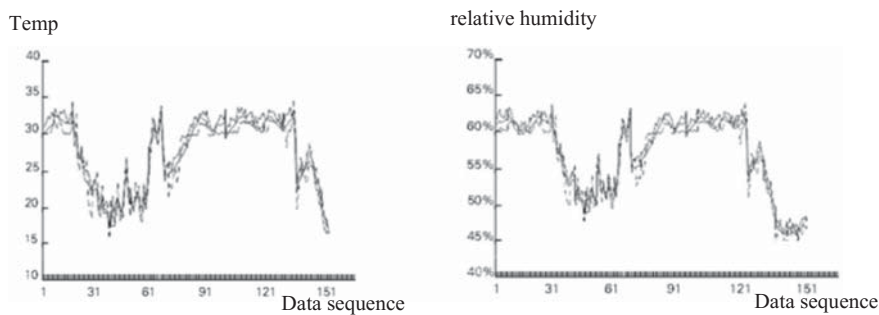


Fig. 4 Fitting curve of intelligent building temperature and humidity environment based on OLS algorithm and weed optimization algorithm

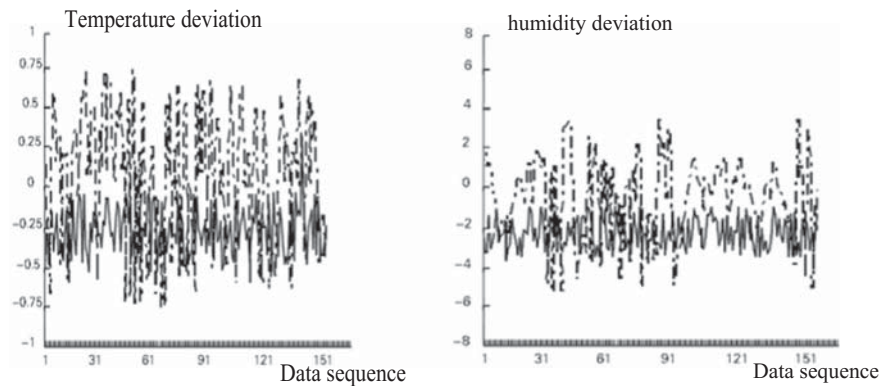


Fig. 5 Fitting deviation curve of intelligent building temperature and humidity environment based on OLS algorithm and weed optimization algorithm

The fitting deviation curves of the above two methods are shown in Fig. 5. Fitting deviation is the difference between the simulation fitting value and the actual value.

The solid line is the actual value curve; the dotted line is the fitting curve of the OLS algorithm; the dash-dotted line is the fitting curve of the weed optimization algorithm.

The dash-dotted line is the fitting deviation curve of the OLS algorithm; the solid line is the fitting deviation curve of the weed optimization algorithm.

According to the simulation results, the fitting deviation range based on the weed optimization algorithm of temperature and humidity environment is $[-0.5, -0.02]$ and $[-3.2, -1.2]$, the fitting deviation range based on OLS algorithm of temperature and humidity environment is $[-0.75, 0.75]$ and $[-5.1, 3.6]$. The number of nodes in the hidden layer of the weed optimization algorithm is 52. The number of nodes in the hidden layer of the OLS algorithm is 79. In the indoor temperature and humidity environment simulation test of intelligent building, the test result of the weed optimization algorithm is obviously better than the OLS algorithm. It can be seen that

the RBF neural network based on weed optimization algorithm is more compact in architecture, and has better approximation ability and generalization ability.

5 Conclusions

In the actual temperature and humidity environment of an intelligent building, it uses OLS algorithm to train the neural network mainly through the orthogonal regression matrix to determine the node centers of a hidden layer, and finally uses the pseudoinverse or LMS method to solve the weights of the output layer. This method has faster training speed, but, there is an error when the center of the hidden layer is determined by orthogonalization. If the entire training process is finished when the preset conditions are satisfied in the process of determining the node centers of the hidden layer by orthogonalization, which may result in the loss of some of the node centers and the network architecture cannot be guaranteed to be minimized. The RBF neural network predictive control method based on the weed optimization algorithm can guarantee the smaller network architecture and has better approximation ability, generalization ability, and fitting precision. It can solve the problem that the complex system is difficult to establish a precise control model, and it is of great significance to solve the related problems in other sub-control fields of an intelligent building system.

Acknowledgements This research work is partially supported by the National Natural Youth Science Foundation of China (Project Codes: 61305125), Shenyang Jianzhu University Discipline Content Education Project (Project Codes: XKHY2-66), the Natural Science Foundation of University (Project Codes: 2014068) and National Post Doctor Foundation (Project Codes: 2013M530955, 2014T70265).

References

1. Yu, C.G., Wang, J.P., Ying, Y.B.: Greenhouse temperature prediction model based radial basis function neural networks. *J. Biomath.* **21**(4), 549–553 (2006) (in Chinese)
2. Sherstinsky, A., Picard, R.W.: On the efficiency of the orthogonal least squares training method for radial basis function networks. *IEEE Trans. Neural Netw.* **7**(1), 195–200 (1996)
3. Zhang, Z.Z., Qiao, J.F.: Design RBF neural network architecture based on online subtractive clustering. *Control Decision* **27**(7), 997–1002 (2012) (in Chinese)
4. Ding, T., Zhou, H.C.: Prediction method research based on radial basis function neural network. *J. Harbin Inst. Technol.* **37**(2), 272–275 (2005) (in Chinese)
5. Wang, J.S., Gao, Z.N.: Traffic modeling and prediction based on RBF neural network. *Comput. Eng. Appl.* **44**(13), 6–11 (2008) (in Chinese)
6. He, F., Ma C.W.: Application of BP neural network based on genetic algorithm in predicting the air humidity of sunlight greenhouse. *Chinese Agric. Sci. Bulletin* **24**(1), 492–495 (2008) (in Chinese)
7. Lin, M.Q., Chen, Z.Q., Yuan, Z.Z.: Self-tuning controller for neural network predictive deviation compensation based on damped least square. *Inf. Control* **29**(1), 27–33 (2000) (in Chinese)

Research on Human Thermal Comfort Model Based on Multiple Physiological Parameters



Yalong Yang, Wenmiao Wu, Qiansheng Fang, Xulai Zhu, Rui Zhang and Mingyue Wang

Abstract Currently, the demand on thermal comfort of architectural environment is becoming higher and higher, the establishment of thermal comfort model based on the physiological parameters plays an important role in improving the indoor thermal comfort and building energy efficiency. The mean skin temperature, skin conductance, and heart rate are three main physiological parameters used to characterize the thermal comfort state of the human body, which will be of great significance to establish the thermal comfort model. In this paper, the physiological experimental program was designed in detail and the subjective questionnaires ruler was classified involving thermal sensation, thermal comfort, and sweat rate. Then, the human subjective thermal response distribution and the regular pattern were analyzed according to the vote on the state of thermal comfort and physiological parameters of subjects under five experimental conditions. As a result, the multiple physiological modeling of indoor human thermal comfort was performed by using partial least squares (PLS) method based on mean skin temperature, skin conductance, and heart rate. Furthermore, the experiments aimed at evaluating the accuracy of the established indoor human thermal comfort model were performed, the results indicate the accuracy of the established model is satisfactory.

Keywords Thermal comfort · Partial least squares method · Physiological parameter · Predicted mean vote

Y. Yang · W. Wu · Q. Fang (✉) · X. Zhu · R. Zhang · M. Wang
Anhui Province Key Laboratory of Intelligent Building and Building Energy Saving, Anhui
Jianzhu University, Hefei 230022, China
e-mail: qsfang@sohu.com

Y. Yang · W. Wu · Q. Fang · X. Zhu · R. Zhang · M. Wang
School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230022,
China

Y. Yang
Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_2

1 Introduction

With the development of the economics and the improvement of people's living standard, the dilemma between the increasing requirement of indoor thermal comfort and the supply of energy shortage becomes increasingly urgent [1, 2]. As a result, finding the most favorable alternatives is very necessary [3]. Generally, the air conditioner was used to improve the indoor thermal comfort by providing indoor air with constant temperature and humidity [4]. However, people's thermal adaptability will be weakened by using a conventional air conditioner [5]. To improve the human thermal comfort to a brand new height, it is necessary to perform research on the indoor thermal environment.

In recent years, physiology was used to investigate the thermal comfort by researchers, which aims at finding the thermal comfort mechanism from a physiology point of view. The effect of different thermal environment on the human physiological parameters was investigated based on physiological indexes [6]. However, the thermal comfort research involves physiology, psychology, ecology, etc., human thermal comfort evaluation plays an important role in human thermal comfort research, the accurate evaluation of human thermal comfort state is the foundation of performing the thermal comfort research deeply [7]. In this paper, the establishment of indoor human thermal comfort about multiple physiological models was performed by using the partial least squares method based on skin temperature, skin conductance, and heart rate. Further experiments were performed to verify the established model [8].

2 Methods

Taking the influence factor of human thermal comfort as an example, $\{X_1, X_2, X_3\}$ are regarded as subjects in the study, which represent skin temperature, skin conductance, and heart rate. The human thermal comfort is considered as dependent variable Y , and the three subjects are considered as independent variable X , which aim at obtaining n sample points. To establish the relationship between the human thermal comfort value and the three influence factors, the principal components of X and Y are extracted first.

t_1 is assumed as the first principal components of $\{X_1, X_2, X_3\}$, u_1 is the first principal component of Y , the requirements are as follows:

- (1) t_1 and u_1 should be carrying variation data at the most; and
- (2) The correlation between t_1 and u_1 should be much higher.

The above two requirements mean that t_1 and u_1 can represent X and Y to the greatest extent possible, and the principal components u_1 of the dependent variable can be explained greatly by the principal component t_1 of the independent variable.

The algorithm implementation is as follows:

- (1) After standardizing of the source data having been done, the independent variable X is changed into a matrix X_0 , and the measurement matrix Y_0 can be obtained after the dependent variable Y is standardized, as shown in Eq. (1),

$$X_0 = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}, Y_0 = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (1)$$

- (2) The first principal component t_1 can be obtained by calculating the X groups of variables, w_1 is the coefficient vector of first principal component and unit vector, the first principal component u_1 can be obtained by calculating the Y groups of variables, c_1 is the coefficient vector of first principal component and unit vector, as shown in Eq. (2),

$$t_1 = X_0 w_1, u_1 = Y_0 c_1 \quad (2)$$

According to $Var(t_1) = \max$, $Var(u_1) = \max$, and $\rho(t_1, u_1) = \max$, Eq. (3) can be obtained.

$$\begin{cases} \max_{w_1, c_1} (X_0 w_1, Y_0 c_1) \\ w_1^T w_1 = 1 \\ c_1^T c_1 = 1 \end{cases} \quad (3)$$

where w_1^T is the transpose of w_1 , c_1^T is the transpose of c_1 .

- (3) Establishment of the regression equation

The following two regression equations are about X_0 to t_1 and Y_0 to c_1 .

$$X_0 = t_1 \alpha_1^T + E_1, Y_0 = c_1 \beta_1^T + F_1 \quad (4)$$

where $\alpha_1^T = \frac{X_0^T t_1}{\|t_1\|^2}$, $\beta_1^T = \frac{Y_0^T u_1}{\|u_1\|^2}$, E_1 , F_1 is the residual matrix of X_0 , Y_0 . Furthermore, replace X_0 , Y_0 with a residual matrix E_1 , F_1 to find the second principal component t_2, u_2 .

- (4) Assuming the rank of $n \times 3$ matrix is $r = \min(n, 3)$, as a result, there are $t_1, t_2, \dots, t_r$ satisfying Eq. (5).

$$\begin{cases} X_0 = t_1 \alpha_1' + t_2 \alpha_2' + \dots + E_r \\ Y_0 = c_1 \beta_1' + c_2 \beta_2' + \dots + F_r \end{cases} \quad (5)$$

$$t_k = w_{k1} X_1 + w_{k2} X_2 + \dots + w_{kp} X_p, k = 1, 2, \dots, r \quad (6)$$

Substituting Eq. (5) into Eq. (6), then we can obtain a nonstandard partial least squares equation as shown in Eq. (7) after combining similar terms having been done.

$$\hat{Y}_j = a_{j0} + a_{j1}X_1 + a_{j2}X_2 + a_{j3}X_3 \quad (7)$$

3 Experiments of Human Thermal Comfort

In this paper, the difference of human skin temperature, skin conductance, and heart rate under different indoor temperature was studied, which aims at investigating the influence of thermal comfort on human physiology and searching the change rule between physiological parameters and temperature. In addition, to establish the multiple evaluating models of human thermal comfort based on multiple physiological parameters, the thermal sensation, thermal comfort, and sweating rate are collected.

3.1 Experimental Environment

The building of the laboratory for human thermal comfort was equipped with a precise HAIRF air conditioner numbered HADC0191. The refrigeration capacity of the air conditioner is 18.2 kW. The cooling and heating effect of the air conditioner can be achieved quickly by using underfloor air supply. In addition, the indoor light intensity is controllable and the wall is heat-insulating, theoretically, the indoor thermal environment can be considered as stable and homogenous. In the laboratory, there are 3 laboratory cabinets, 2 computers, 1 book cabinet, and 1 sofa, the human skin temperature, skin conductance, and heart rate can be obtained by using vital signs monitor named Biofeedback x-pert, the experimental conditions were as shown in Table 1. The experiments were oriented towards undergraduate students, the ages of the students are between 23 and 26 to establish the thermal comfort model of the human body in the classroom of a university. Furthermore, all the subjects have been in Hefei city (China) for more than 2 years, they have already gotten used to the climate in this city.

3.2 Design of Questionnaires

The subjective parameters used to reflect the human thermal comfort include thermal sensation, thermal comfort, and sweat rate, because the subjective values cannot be measured by instruments, the subjective questionnaire survey was used to obtain thermal comfort values scientifically and reasonably. The designed voting scales of the questionnaires are as shown in Tables 2, 3, 4 and 5. In Table 2, ASHRAE

Table 1 Experimental conditions

Items	Parameters
Dimensions of laboratory	Laboratory: 8.7 m × 6.4 m × 2.6 m (length × width × height), working area: 8.4 m × 6.1 m × 2.4 m (length × width × height)
Materials of laboratory	The inner and outer walls: 240 mm in width, Roof: waterproofing and thermal insulation, Windows: double glass and steel frame
Parameters of constant temperature and humidity air conditioner	Supply air rate: $G_s = 5200 \text{ m}^3/\text{h}$, Air changes: $N = G_s/LWH = 40 \text{ time/h}$

Table 2 Voting scale of thermal sensation (TSV)

TSV	Cold	Cool	Slightly cool	Neutral	Slightly warm	Warm	Hot
Scale	−3	−2	−1	0	1	2	3

Table 3 Voting scale of thermal comfort (TCV)

Slightly warm	Warm	Hot	Slightly warm	Warm	Hot
Scale	0	1	2	3	4

Table 4 Voting scale of sweat rate

Degree of perspiration	Dry skin no sweat	No sweat stick	Skin moist perspiration	Little sweat	More sweat
Scale	0	1	2	3	4

(American Society of Heating, Refrigerating and Air-Conditioning Engineers) seven-level scale was chosen as the subjective judgment scale of a thermal sensation of human to the surrounding environment [9]. In Table 3, five-level scale was chosen as the voting scale of human thermal comfort [10]. In Table 4, five-level scale was chosen as the voting scale of human sweat rate. In Table 5, three-level scale was chosen as the voting scale of the expectations of the human body to the thermal environment.

Table 5 Voting scale of thermal environment expectations

Expectation temperature	Reduce	Constant	Rise
Expectation humidity	Reduce	Constant	Rise
Expectation wind	Decrease	Constant	Enlarge
Indicators	−1	0	1

Table 6 Experimental conditions

Condition	Temperature (°C)	Relative humidity (%)	Wind (m/s)	Activity (W/m ²)	Thermal resistance (clo)
1	24	60%	≤0.3	69.78	0.3
2	26				
3	28				
4	30				
5	32				

3.3 Test Conditions and Procedures

According to the monitoring results of indoor thermal environment parameters of typical summer weekday and experimental conditions, the setting conditions are as shown in Table 6. The dresses of the experimental subjects were controlled, and the subjects were asked to sit or move moderately.

The experiment selected 20 students aged between 22 and 30, 14 boys and 8 girls. Before the experiment, the experimental subjects were asked to move moderately to adjust their psychological state. After half an hour, the experimental subjects were asked to wear equipment. In the experiment, the thermal environment parameters were recorded, and the human physiological parameters were gathered. Then questionnaire of subjects on thermal sensation, thermal comfort and sweat rate was performed, the data were extracted, recorded and saved.

4 Establishment and Evaluation of the Human Thermal Comfort Model

4.1 Experimental Data

In this part, the experimental data obtained from the experiment are analyzed and summarized, and the regularity of skin temperature, skin conductance and heart rate to the change of ambient temperature is mainly studied. The effects of indoor temperature on skin temperature, heart rate, and skin conductance were investigated.

4.2 Establishment of the Human Thermal Comfort Model

In this research, the human thermal comfort model was established based on a partial least squares regression method, the independent variable *X* represents the six

points skin temperature, the dependent variable Y represents the human thermal comfort, a number of newly integrated variables with the best explanatory ability were extracted from independent variables set and dependent variable set, the cumulative explanatory ability relationship was obtained according to the extracted factor of cross-validation.

t_1/u_1 is the rate of the linear combination of the independent variable set $X = (x_1, \dots, x_p)$ and the dependent variables set $Y = (y_1, \dots, y_p)$, when the rate is close to 1, which indicates the decomposing variable is obviously linear correlation, accordingly, the correlation of input variables and output variables is good, which indicates the fitting effect of the established model is good. In addition, the temperature of head, leg, thigh, and arm was positively correlated with a voting scale of thermal comfort (TCV), and the conclusion indicates that the temperature of chest and back was negatively correlated with TCV.

After the standardized variables having been restored, the contribution value of each parameter to TCV was obtained, which results in the physiological evaluating model of the relationship between the skin temperature and thermal comfort as expressed by Eq. (8). Where, A~F represent the six points skin temperature of forehead, chest, calves, thighs, arm, and back. What is more, the range of TCV is within 0–4.

$$TCV = 0.54A - 0.07B + 0.16C + 0.45D + 0.26E - 0.05F - 42.14 \quad (8)$$

Based on the established human thermal comfort model of the average skin temperature, skin conductance and heart rate are regarded as input, and the values of thermal comfort are regarded as output. After the standardized variables having been restored, which results in the multiple physiological evaluating models as expressed by Eq. (9). Where, A~C represent the values of average skin temperature, skin conductance, and heart rate. The range of TCV is also within 0–4.

$$TCV = 0.51A + 0.013B + 0.046C - 20.435 \quad (9)$$

4.3 Evaluation of the Human Thermal Comfort Model

As shown in Fig. 1, the experimental data were simulated by regression equations, the accuracy of the model prediction is verified by the subjective thermal sensation of the subject represented by the vote value, the values of Vote/Pred reflect the ratio of predictive value and the real value, the slope of Vote/Pred is approximately 1, which indicates the predictive value of the multiple physical models is reasonable. The simulating diagram is as shown in Fig. 2, the black curve represents the vote value of thermal comfort, the red curve represents the predictive value of the established model, the maximum deviation, minimum deviation, and average deviation between real value and predictive value are 0.32, 0.03, and 0.16, respectively, which indicates

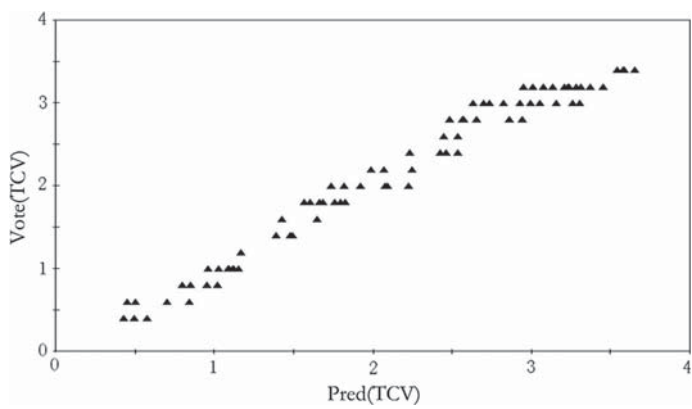


Fig. 1 The ratio of prediction values and factual values

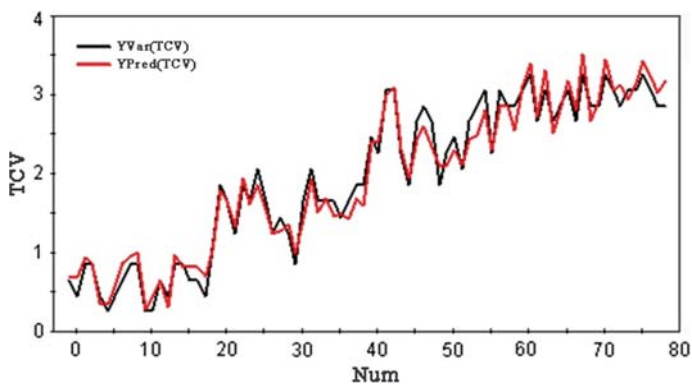


Fig. 2 Traceability of the multivariate physiological evaluation model

the variation trend of human thermal comfort can be tracked by the established model perfectly, most of the predictive data are accurate, and the performance of the established model is satisfactory.

5 Results and Discussion

Figures 3 and 4 are prediction diagrams of the thermal comfort model based on the partial least squares method and classical Predicted Mean Vote (PMV) model. For the thermal comfort model based on the partial least squares method, only a small number of points are far from the diagonal line, they represent that the difference between the predicted value of the model and the subjective true value are too large, and the predicted value cannot effectively predict the true thermal sensation of the

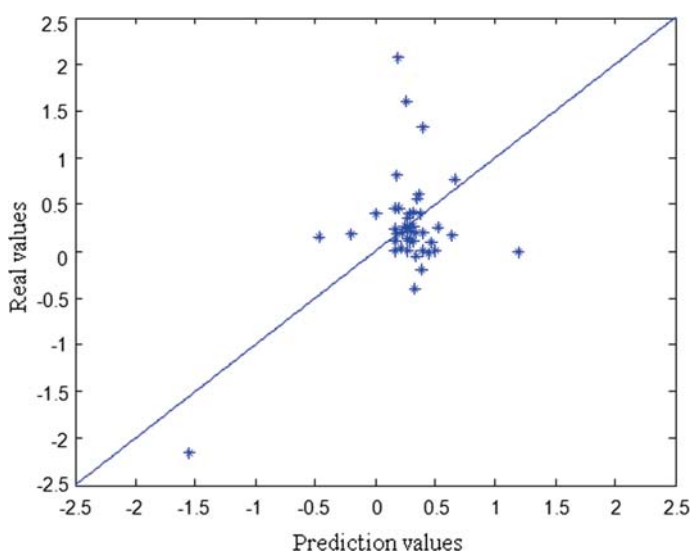


Fig. 3 The prediction diagram of thermal comfort based on the partial least squares method

subjects. In other words, most of the points are close to the diagonal line, which indicates the difference between fitted value and real value is very small. However, for the thermal comfort model based on the PMV model, most of the points are far from the diagonal line, only a small number of points are close to the diagonal line. Clearly, the fitting effect of the thermal comfort model based on the partial least squares method is better than the thermal comfort model based on the classical PMV model.

6 Conclusions

After the physiological parameters of subjects collection and the questionnaire votes of thermal sensation, thermal comfort and sweat rate of subjects collection based on a large number of experiments having been done, the evaluating model of six points skin temperature was established by using partial least squares method, then the multiple physiological modeling based on multiple evaluating models of thermal comfort involving the average skin temperature, skin conductance, and heart rate was established by using partial least squares method, the results indicate the variation trend of human thermal comfort can be tracked by the multiple evaluating models perfectly, which can overcome the disadvantages of subjective evaluation. In addition, the TCV is calculated by the multi- physiological evaluation model based on the obtained average skin temperature, skin conductance and heart rate of the human body, and the recommendation table of human body thermal comfort is formed

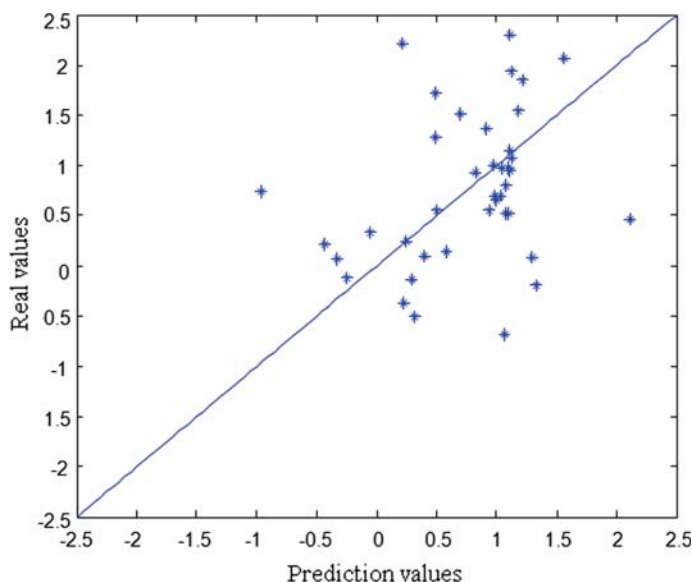


Fig. 4 The prediction diagram of thermal comfort based on the classical PMV model

through decision-making, and the program is programmed into the air-conditioning controller to realize the intelligence control of the air-conditioning, which aims at ensuring indoor personnel achieve superior thermal comfort.

Acknowledgements and Statements This research was funded by the National Key Research and Development Program of China (2017YFC0704100) and the Natural Science Foundation of Anhui Province, China (1508085QF131) and the Major Project on the Integration of Industry, Education and Research, Institute of plasma physics, Chinese Academy of Sciences, China (AJ-CXY-KF-17-36).

This study has been conducted with ethics approval obtained from the ethics committee: Prof. Yinfeng Zhu of the Department of Electronic and Information Engineering, Anhui Jianzhu University, Hefei, China.

Participants are recruited from Anhui Jianzhu University, Hefei, China, and all participants have given their consent to use the dataset and disclose information relevant for research in this study.

References

1. (Jerry) Yu, Z., Yang, B., Zhu, N., Olofsson, T., Zhang, G.: Utility of cooling overshoot for energy efficient thermal comfort in temporarily occupied space. *Build. Environ.* **109**(15), 199–207 (2016)
2. Lawrence, R., Keime, C.: Bridging the gap between energy and comfort: post-occupancy evaluation of two higher-education buildings in Sheffield. *Energy Build.* **130**(15), 651–666 (2016)

3. Embaye, M., AL-Dadah, R.K., Mahmoud, S.: Numerical evaluation of indoor thermal comfort and energy saving by operating the heating panel radiator at different flow strategies. *Energy Build.* **121**, 298–308 (2016)
4. Huang, Y., Niu, J.: A review of the advance of HVAC technologies as witnessed in ENB publications in the period from 1987 to 2014. *Energy Build.* **130**(15), 33–45 (2016)
5. He, Y., Li, N., Peng, J., Zhang, W., Li, Y.: Field study on adaptive comfort in air conditioned dormitories of university with hot-humid climate in summer. *Energy Build.* **119**, 1–12 (2016)
6. Chuen, L., Sears, D., McAdams, S.: Psycho-physiological responses to auditory change. *Psycho-physiology* **53**(6), 891–904 (2016)
7. Alfano, F.R.D., Palella, B.I., Riccio, G.: Thermal environment assessment reliability using temperature humidity indices. *Ind. Health* **49**(1), 95–106 (2011)
8. de Paulo, J.M., Barros, J.E.M., Barbeira, P.J.S.: A PLS regression model using flame spectroscopy emission for determination of octane numbers in gasoline. *Fuel* **176**, 216–221 (2016)
9. Hussin, L.X.A., Salleh, E., HY Chan, S. Mat.: The reliability of predictive mean vote model predictions in an air-conditioned mosque during daily prayer times in Malaysia. *Archit. Sci. Rev.* **58**(1), 67–76 (2015)
10. Schweiker, M., Shukuya, M.: Adaptive comfort from the viewpoint of human body energy consumption. *Build. Environ.* **51**, 351–360 (2012)

Study on Thermal Comfort Model Based on Genetic Algorithm with Backpropagation Neural Network



Yalong Yang, Dejian Hong, Rui Zhang, Qiansheng Fang, Xulai Zhu and Wenmiao Wu

Abstract To reduce fossil fuel energy consumption and improve the using efficiency, it is of great significance to study the thermal comfort model based on multiple physiological parameters. Compared with the classical models, comfort can be reflected more accurate by using the thermal comfort model based on multiple physiological parameters. In this paper, the experiments were performed to verify the effectiveness of the thermal comfort model. In particular, to verify the practicability of the thermal comfort model based on multiple physiological parameters, the established thermal comfort model based on the genetic algorithm with a backpropagation neural network and the classical PMV were compared. The results indicate that the established thermal comfort model is reasonable, which provides a feasible option for achieving a comfortable indoor environment. Finally, it puts forward further study on the thermal comfort model based on more physiological parameters.

Keywords Thermal comfort model · Multiple physiological parameters · Genetic algorithm

1 Introduction

With the increasing population and the development of society, energy saving is necessary for building industry [1]. To reduce air pollution and greenhouse effect, it is essential to study the thermal comfort of buildings [2]. Because the thermal comfort

Y. Yang · D. Hong · R. Zhang (✉) · Q. Fang · X. Zhu · W. Wu
Anhui Province Key Laboratory of Intelligent Building and Building Energy Saving, Anhui
Jianzhu University, Hefei 230022, China
e-mail: rzhang0551@foxmail.com

Y. Yang · D. Hong · R. Zhang · Q. Fang · X. Zhu · W. Wu
School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230022,
China

Y. Yang
Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031, China

degree of buildings can be evaluated by using multiple physiological parameters, the change of environmental parameters can be achieved by using human physiological parameters indirectly [3].

In recent years, great progress has been made on the research of thermal comfort. In 2000, Nguyen had collected the datum of natural ventilation from a lot of long-term houses in Southeast Asia, then the adaptive natural ventilation model was established [4]. In 2008, the thermal comfort model was established by Prof. Qiansheng Fang at the Key Laboratory of Intelligent Building, Anhui Jianzhu University, Hefei, China, which was based on a 5000-questionnaire survey about the building thermal comfort of hot summer and cold winter. In 2011, according to the environmental monitoring and questionnaire survey, the indoor ventilation adaptation model was established by Baizhan Li and Zhenxing Jin in Chongqing University, Chongqing, China [5].

Generally, the physiological parameters related to the thermal comfort include skin temperature, skin conductance, heart rate, pulse rate, electrocardiogram, electroencephalography, blood oxygen saturation, electromyography, perspiration rate, respiratory rate, metabolic rate. The good physiological correlation between selected physiological parameters and thermal comfort is required. In this paper, the thermal comfort model will be established based on multiple physiological parameters involving electrocardiogram (ECG), blood pressure (BP), and galvanic skin (GSR).

2 Establishment of Thermal Comfort Model

In this paper, three types of human physiology parameters were chosen as the input of the thermal comfort model, which include electrocardiogram (ECG), blood pressure (BP), and galvanic skin (GSR) of tested people. Firstly, the environmental parameters were changed to induce the alternations of human physiology variables. As a result, the relationship between the physiological signals and the thermal comfort was obtained. Then, the physiological characteristics were classified and recognized using a backpropagation (BP) neural network [6], furthermore, the thermal comfort model was established based on genetic algorithm optimization. Finally, the thermal comfort model and the classical predicted mean vote (PMV) were compared, which aims at verifying the feasibility and practicability of the established thermal comfort model. Establishment of the thermal comfort model is as shown in Fig. 1.

2.1 Physiological Signal Extraction

The physiological signal should be extracted before establishing the thermal comfort model. The wavelet packets transform (WPT) is used to decompose the original signals involving low frequency and high frequency. The frequency band can be chosen according to the signal characteristics and requirements. Therefore, WPT

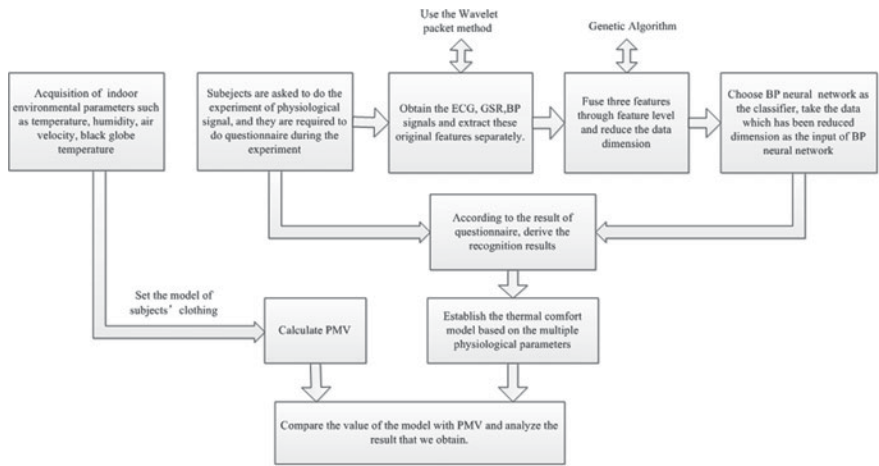


Fig. 1 Establishment of the thermal comfort model

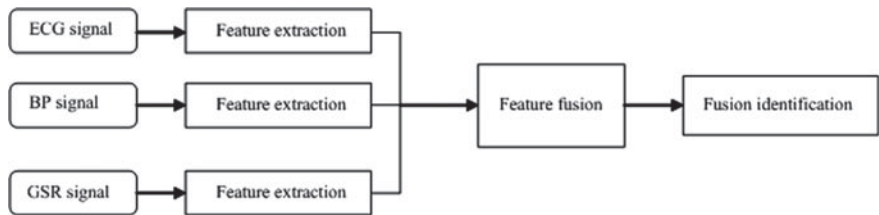


Fig. 2 Feature fusion process of the three physiological signals

is appropriate for time–frequency analysis, and the essential characteristics of the extracted signals can be reflected by using WPT.

2.2 Feature Fusion of Physiological Signals

Fusion of the three physiological signals was regarded as multi-biometric feature fusion, it is important to select the method of fusion, especially the choice of fusion level. In this paper, the feature layer fusion of three physiological signals is selected by the mean impact value (MIV) algorithm. The feature fusion process of the three physiological signals is shown in Fig. 2.

2.3 The Feature Optimization Based on Genetic Algorithm

Due to large dimensions of the new feature vector, measures should be taken to classify and recognize the physiological features better, so the new feature vector should be optimized, which aims at deleting the redundant information of physiological features and choosing the feature vectors that can reflect the relationship between input and output.

The feature optimization is performed by using a genetic algorithm, the basic principle is as follows: After the defect feature parameters needed to be optimized are encoded, the population is initialized with a certain scale, so each individual of the population was regarded as solutions. Then the fitness value of each individual was calculated by fitness function and the genetic operation was decided, consequently, the population is selected, crossed and mutated with a certain probability until it satisfies the terminal condition [7].

2.3.1 Encoding the Defect Features

The chromosome (individual) of each population is represented by using a binary signed digit number—{0, 1}. After extracted, each of the three physiological signals was converted into 24 dimensions, which corresponds to 72 dimensions of physiological signals, obviously, the length of the chromosome is 72. Every value of a gene is named whether ‘1’ or ‘0’. When the gene of the chromosome is ‘1’, it means the feature vector corresponding to the gene is selected, whereas it is not selected.

2.3.2 Calculation of Fitness Function

In the genetic algorithm, fitness is used to measure the individuals in the population close to the degree of optimal solutions, the higher fitness of an individual, the higher probability of heredity to the next generation. The function used to measure individual fitness is called fitness function, which is expressed by Eq. (1).

$$f(X) = \frac{1}{SE} = \frac{1}{sse(\hat{T} - T)} = \frac{1}{\sum_{i=1}^n (\hat{t}_i - t_i)^2} \quad (1)$$

where $\hat{T} = \{\hat{t}_1, \hat{t}_2, \hat{t}_3, \dots, \hat{t}_n\}$ is the predicted results of test set; $T = \{t_1, t_2, t_3, \dots, t_n\}$ is the actual values, and n is the test set of quantity of the samples.

2.3.3 Selection Operations

Proportional selection operator is chosen for selection operation. First, the fitness of all individuals is added together, which is expressed by (2).

$$F = \sum_{k=1}^{n_r} f(X_k) \quad (2)$$

Then the relative fitness of each individual in the population is calculated by using (3):

$$P_K = \frac{f(X_k)}{F} \quad k = 1, 2, \dots, n_r \quad (3)$$

Finally, the individual fitness is determined and arranged from high to low according to the fitness values obtained in the first step. 10% of the excellent individuals in the population are replicated to the next generation, and the rest have to be selected in all chromosomes through the proportion of ‘roulette’, which aims at ensuring the optimal group in the next generation is not worse than the one in the last generation.

2.3.4 Crossover Operations

New chromosome is formed through recombination of two homologous chromosomes from each other and then the new individuals are generated. According to the pre-set crossover probability named P_c , two individuals as parents are selected from the population by using single-compound crossover operator. For two individuals that pair up with each other, the position followed a gene (without including the first position) was random set as the intersection, so the possible positions of intersection are $L - 1 = 71$.

2.3.5 Mutation Operations

The mutation operations on the gene number of individual code strings are performed based on basic bit mutation strategy. The locus of mutative genes is selected with a pre-set mutation probability named P_m , which means the gene of each specified mutation point is inverted, and the change of genes is from ‘0’ to ‘1’ or from ‘1’ to ‘0’, so new individual was generated.

2.3.6 Terminal Conditions

Currently, the genetic algebra named the number of iterations is used as the terminal conditions. In this paper, the number of iterations is $T = 100$. The design procedure is as shown in Fig. 3.

Fig. 3 The design procedure of the genetic algorithm

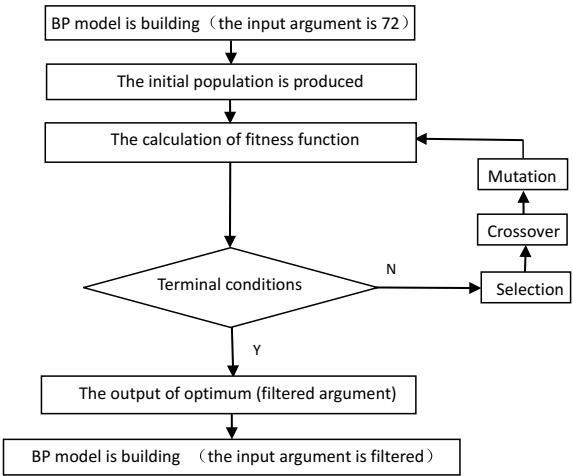
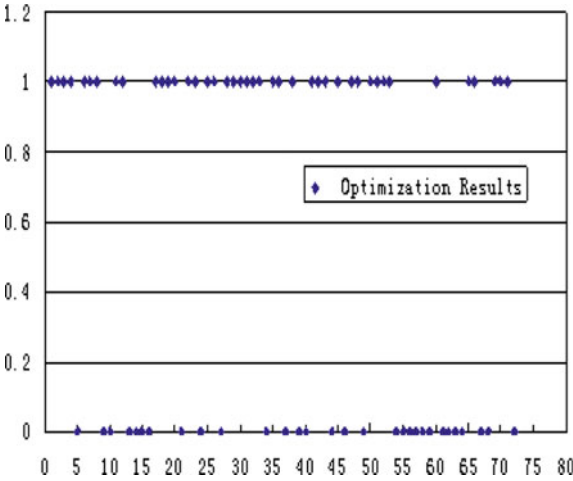


Fig. 4 Optimal binary encoding of chromosome



2.3.7 Optimal Solutions and the Results of Filter

The optimal binary encoding of chromosome is shown in Fig. 4, the horizontal axis is the dimension of feature dimensions, and the vertical axis is the optimization results. So the numbers of input filtered argument are as follows:

- 1, 2, 3, 4, 6, 7, 8, 11, 12, 17, 18, 19, 20, 22, 23, 25, 26, 28, 29, 30, 31, 32, 33, 35, 36, 38, 41, 42, 43, 45, 47, 48, 50, 51, 52, 53, 65, 69, 70, 71.

2.4 The Principles and Construction of BP Neural Network

BP neural network consists of input layers, hidden layers, and output layers [8], which is used to establish the relationship between physiological signals and thermal comfort based on multiple physiological parameters. According to Kolmogorov theorem [9], the key issues of classification can be fulfilled by using the hidden layers.

2.5 Identification of the Numbers of Neurons in Hidden Layers

At present, there is no sound theory to determine the number of neurons in the hidden layers of a BP neural network. So the number is often set with the given performance indicators based on theoretical analysis and experiments. As shown in Eqs. (4), (5), and (6), three empirical equations are chosen to compute the numbers of neurons in the hidden layers.

$$y = 2x + 1 \quad (4)$$

$$y = \sqrt{x + z} + a, a \in [1, 10] \quad (5)$$

$$y = \log_2 x \quad (6)$$

In Eqs. (4–6), x is the number of neurons in the input layer; y is the number of neurons in the hidden layer; and z is the number of neurons in the output layer. According to the obtained results, the most suitable number of the neurons in the hidden layer is 20.

2.5.1 Transfer Function

As shown in (7) and (8), the nonlinear Logsig function and the transferring function are usually chosen as the Transfer function. The output results are described by “0” and “1”. So Logsig function is chosen in the output layer, and the transferring function is chosen in the hidden layer.

$$\text{Logsig}(x) = \frac{1}{1 + \exp(-x)} \quad (7)$$

$$\text{Tansig}(x) = \frac{2}{1 + \exp(-2x)} - 1 \quad (8)$$

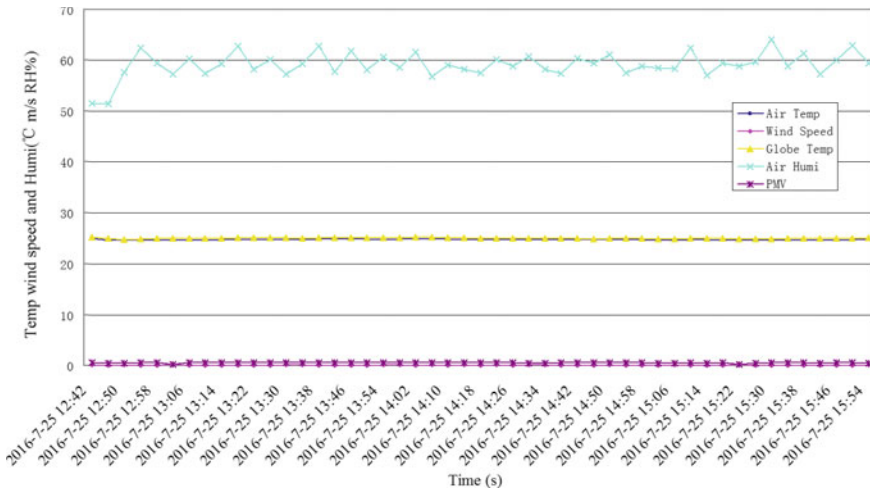


Fig. 5 Environmental parameters and calculating values of PMV

3 Experiment and Results

3.1 Subjects and Facilities

The subjects are 86 local and healthy undergraduate, and all of them wear the same type of clothes. The clothing insulation value is approximately 0.5 clo. The indoor environmental parameters are controlled by the precision air conditioner named HADC0191 and numbered HADC0191, the indoor humidity is 40 ~ 65%, and the indoor wind velocity is less than 0.3 m/s. In the laboratory, each subject is required to seat and maintain emotional stability. The physiological parameters are measured by PowerLab Biological Signal Collection System comes from AD Instruments company of Australia, the environmental parameters and values of PMV were measured by SSDZY-1 thermal comfort recorder. All the measured environmental parameters and calculating values of PMV are shown in Fig. 5.

3.2 Data Preprocessing

Totally, there are 86 groups of acquired data, but only 82 groups of acquired data are reasonable, which include 65 groups of trained data and 16 groups testing data. The learning rate and deviation are adjusted by calculating until the deviation meets the requirements.

3.3 Establishment of BP Neural Network Model

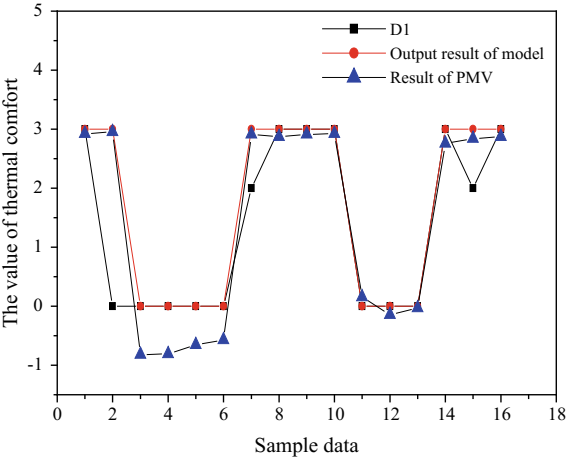
The 72-dimensional feature vectors and 40 filtered feature vectors by genetic algorithm are used as the input level of BP neural network [10], the choice of the number of neurons on the output level corresponds to the level of comfort. In this paper, five comfort grades are used to represent the feelings of hot, warm, moderation, cool and cold, which correspond to 100, 011, 010, 001, 000, respectively. When the thermal values are expressed by three-bit binary number, the number 3 of output layer neurons can be decided based on the neural network. The calculating results and the experimental results indicate that the suitable number of hidden level for modeling is 23.

3.4 Verification of Experimental Results

The experimental results of the BP neural network model need to be verified. To verify if the model conforms to thermal comfort sensation of subjects, the output results of modeling are compared with the thermal comfort sensation of subjects, meanwhile, the classical PMV is also compared with the thermal comfort sensation of subjects, so the advantages and disadvantages of the thermal comfort model are obvious. The results are shown in Fig. 6.

As shown in Fig. 6, the circle represents investigation, the point represents calculation, and the cross represents the results of output modeling. For the value of thermal comfort calculated by using PMV and investigation, the maximum deviation, the minimum deviation, and the average value are 2.9515, 0.0326, and 0.3196, respectively. The maximum deviation between the result of output modeling and

Fig. 6 Scatter plot among classic PMV, the result of output modeling and investigation



investigation of thermal comfort is 3.0, while the minimum deviation is 0, and the average value is 0.3125.

Compared with the subjective well-being of subjects, the accuracy of the thermal comfort model based on multiple physiological parameters is higher than the results calculated by PMV. So it is feasible for the thermal comfort modeling.

4 Conclusions and Discussion

- (1) Compared with the PMV, the establishment of the thermal comfort model based on the multiple physiological parameters is absolutely necessary for building energy conservation.
- (2) The performance of the thermal comfort model established by the BP neural network is improved based on the optimized feature vector by genetic algorithm, which can provide a valuable reference on the research of a comfortable indoor environment.
- (3) In this paper, three physiological parameters including ECG signals, BP signals, and GSR signals are studied, however, for the other physiological parameters which can better reflect thermal comfort, need to be investigated further.
- (4) To obtain more accurate results, the number of tested subjects should be large enough. In addition, the sex ratio should be taken into account, the percentages of boys and girls should be balanced as much as possible.

Acknowledgements and Statements This research was funded by the National Key Research and Development Program of China (2017YFC0704100), the National Key Technology Support Program of China (2015BAJ08B03), the Natural Science Foundation of Anhui Province, China (1508085QF131) and the Major Project on the Integration of Industry, Education and Research, Institute of plasma physics, Chinese Academy of Sciences, China (AJ-CXY-KF-17-36).

This study has been conducted with ethics approval obtained from the ethics committee: Prof. Yinfeng Zhu of the Department of Electronic and Information Engineering, Anhui Jianzhu University, Hefei, China.

Participants are recruited from Anhui Jianzhu University, Hefei, China, and all participants have given their consent to use the dataset and disclose information relevant for research in this study.

References

1. Jiang, Y.: China building energy conservation strategy study. *Eng. Sci.* **13**(6), 30–38 (2011) (In Chinese)
2. Arslanoglu, N., Yigit, A.: Experimental and theoretical investigation of the effect of radiation heat flux on human thermal comfort. *Energy Build.* **113**, 23–29 (2016)
3. Cui, W., Ouyang, Q., Zhu, Y.: Field study of thermal environment spatial distribution and passenger local thermal comfort in aircraft cabin. *Build. Environ.* **80**, 213–220 (2014)
4. Humphreys, M.A., Nicol, J.F.: Outdoor temperature and indoor thermal comfort: raising the precision of the relationship for the 1998 Ashrae database of field studies. *Ashrae Trans.* **106**(2), 485–492 (2000)

5. Liu, H., Tan, Q., Li, B., Tan, M., Ma, X.: Impact of cold indoor thermal environmental conditions on human thermal response. *J. Central South Univ. Technol.* **18**(4), 1285–1292 (2011) (In Chinese)
6. Wang, J., Cheng, C.-M.: Formulation of estimation models for wind force coefficients of rectangular shaped buildings. *J. Appl. Sci. Eng.* **20**(1), 55–62 (2017)
7. Tang, B., Kong, J., Wang, X., Hou, Y.: Feature dimensions reduction and its optimization for steel strip surface defect based on genetic algorithm. *J. Iron Steel Res.* **23**(9), 59–62 (2011) (In Chinese)
8. Bing, W., Han, S., Xiao, J., Xiaoguang, H., Fan, J.: Error compensation based on BP neural network for airborne laser. *Optik (International Journal for Light and Electron Optics)* **127**, 4083–4088 (2016)
9. Datta, S.: Robust cycles in Kolmogorov-Lotka-Volterra class of models with intraspecific co-operation. *Nonlinear Anal. Real World Appl.* **32**, 98–110 (2016)
10. Liu, W.W., Lian, Z.W., Zhao, B.: A neural network evaluation model for individual thermal comfort. *Energy Build.* **39**, 1115–1122 (2007)

An Optimization Model for PV and CCHP-Supplied Power System in Buildings



Jin Zhao and Jing Yong

Abstract The combined cooling, heating, and power (CCHP) system is beneficial to energy conservation and pollution elimination. With the increased application of the distributed photovoltaic (PV) system, the CCHP and PV combined system is attracting the attention of energy users. Due to the randomness of PV power, the operation dispatch of the combined CCHP and PV system becomes even complicated. This paper proposes an optimization model for this system to achieve the minimization of the cost, by considering the seasonal change of the electric power and heating demands. An improved particle swarm optimization (PSO) is adopted in the optimization model to enhance effectiveness. The case study for an office building is also presented in this paper. And the results show more than 10% of the cost is saved.

Keywords Photovoltaic · Combined cooling heating and power · Optimization model · Particle swarm optimization

1 Introduction

The CCHP system, also called the trigeneration system, has been widely accepted and utilized as the power supplies for industrial, commercial and residential customers. In the recent decade, renewable energy becomes a hot topic and it indeed improves energy conservation. Among various renewable energies, PV is the most pronounced one, due to the mature technology, decreased installation cost and low maintenance needs. Combined CCHP and PV system may be a trend in the future.

The CCHP system for general industrial, commercial, or residential buildings is always classified into the small-scale CCHP system. Once the PV is introduced, the combined system will encounter uncertainties brought not only by the load variation,

J. Zhao (✉) · J. Yong
School of Electrical Engineering, Chongqing University, Chongqing, China
e-mail: well_zhaojin@163.com

J. Yong
e-mail: yongjing@cqu.edu.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_4

but also the PV power, since the PV output is determined by the solar irradiance which varies with time.

There are many studies on the CCHP systems, mainly focusing on how to schedule the units or subsystem. Xianya and Dajun [1] and Yanfei et al. [2] proposed a mathematical model on the economical operation of the CCHP. Jie et al. [3] and Youyin [4] analyzed the impact of different dispatch modes on microgrid economic operation. In [5–7], multi-objective optimizing models of CCHP were developed to satisfy the economy and pollution requirements. However, the research on the system dispatch strategy is considered as a MILP problem in [1, 2] because the authors assumed that the performance parameters of equipment were constant, which was inconsistent with the reality. To solve the nonlinear problem in optimizing the operation of the CCHP system, [8] adopted PSO. But the main shortcoming of this algorithm is getting trapped in a local optimum solution because of the lack of diversity of swarm [9].

Responding to this need, this paper proposes an optimization model for the combined CCHP and PV system to achieve the minimization of the cost and proposes an improved PSO to solve this optimization model. This remainder of this paper is organized as follows. Section 2 presents and discusses the models of each unit associated with the combined CCHP and PV system. Section 3 proposes an improved PSO developed for the particular system. Section 4 demonstrates the process of using the proposed algorithm and the case study results. Section 5 concludes the work.

2 Models of Individual Unit

Figure 1 shows a diagram of the combined CCHP and PV system. In the diagram, the electric energy and thermal energy are integrated by various devices including PV unit, the internal combustion engine (ICE), gas furnace (GF), the waste heat recovery unit (WHRU), the lithium bromide absorption chiller (LiBr AC), etc. This system should be well managed to satisfy both the load demands and the low-cost requirements. To achieve this, the PV model must be included in the conventional CCHP model.

2.1 Model of PV Unit

Theoretically, the output power of PV is determined by solar irradiance, ambient temperature, wind speed, ambient humidity, air cleanliness, etc. For simplification, only solar irradiance and ambient temperature are considered in this paper, because the PV modules on the present market have been encapsulated and the other factors do not significantly affect the PV's output.

Therefore, the actual power output of the PV unit can be expressed as (1) [10].

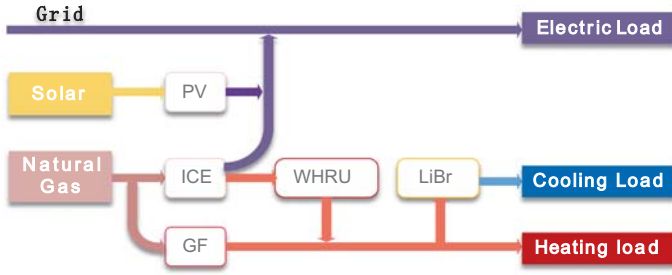


Fig. 1 The diagram of the combined CCHP and PV system

$$P_{PV} = P_{STC} \frac{G}{1000} [1 + k(T_{\varepsilon} - 25)] \quad (1)$$

In this equation,

$$T_{\varepsilon} = T_{amd} + 30 \times \frac{G}{1000} \quad (2)$$

where P_{PV} indicates the output power of PV unit; P_{STC} is the maximum output power of PV unit under standard testing condition when solar irradiance is 1 KW/m² and ambient temperature is 25 °C simultaneously; G is solar irradiance; k is the power temperature coefficient; T_{ε} and T_{amd} are the operating temperatures of PV panel and ambient temperature, respectively.

The PV output is uncontrollable owing to the indeterminacy of environmental factors. It is, therefore, assumed that the PV unit always keeps its output in maximum condition.

2.2 Model of ICE

The operation model of ICE is described as (3)–(6).

$$V_{ICE} = \frac{\sum P_{ICE} \Delta t}{\eta_e H} \quad (3)$$

$$Q_{ICE} = \frac{P_{ICE}(1 - \eta_e - \eta_l)}{\eta_e} \quad (4)$$

$$Q_{EG} = \eta_g \lambda_1 Q_{ICE} \quad (5)$$

$$Q_{JW} = \eta_w \lambda_2 Q_{ICE} \quad (6)$$

where V_{ICE} is the volume of natural gas consumed by ICE during operating; P_{ICE} and Q_{ICE} are the output electric power and output waste heat, respectively, of ICE; Δt is the operating time; η_e is the generating efficiency of ICE; H is the caloric value of natural gas; η_l is the energy loss coefficient of ICE; Q_{EG} and Q_{JW} are the recovered waste heat, from flue gas and waste heat of cooling water of cylinder sleeve respectively; λ_1 is the ratio of waste heat of flue gas in Q_{ICE} and λ_2 is the ratio of waste heat of cooling water in Q_{ICE} ; η_g is the heat exchange efficiency between flue gas and water as well as η_w is the heat exchange efficiency between water and water in WHRU.

Actually, η_e is related to the loading rate (k_l) of ICE, which is shown by the fitting Eq. (7).

$$\eta_e = 0.1824k_l^3 - 0.8150k_l^2 + 0.8977k_l + 0.0245 \quad (7)$$

where

$$k_l = \frac{P_{ICE}}{P_N} \quad (8)$$

where P_N is the maximum output power of ICE.

2.3 Model of GF

In this paper, GF acts as a standby device of ICE. Heat efficiency (η_{gf}) is a key performance indicator of GF, which is defined as (9) [2].

$$\eta_{gf} = \frac{Q_{GF}}{Q_c} \quad (9)$$

where

$$Q_c = H V_{GF} \quad (10)$$

and Q_{GF} is output heat of GF; Q_c is the total heat generated when fuel combusts; V_{GF} is the volume of natural gas consumed.

2.4 Model of LiBr AC

In comparison with the electric chiller, LiBr AC has better performance but lower cost. Generally, coefficient performance (COP) represents the energy efficiency level

of LiBr AC. Equation (11) shows the relation between input thermal power (Q_{IN}) and output cooling power (Q_{AC}).

$$Q_{AC} = Q_{IN} \cdot COP \quad (11)$$

3 Optimization Algorithm for System Dispatch

3.1 Optimizing Dispatch Modeling

The optimizing dispatch model for the combined CCHP and PV system is established based on the forecast curves for electric, thermal and cooling loads.

(1) Optimization Objective Function

The optimization objective is to minimize the operating costs shown in (12), considering fuel costs C_g , maintenance costs C_{om} , costs of interacting with the grid C_{grid} and subsidy of PV generation C_{pv} , i.e.,

$$\min C = C_g + C_{om} + C_{grid} - C_{pv} \quad (12)$$

where

$$C_g = C_{gas}(V_{ICE} + V_{GF}) \quad (13)$$

$$C_{om} = \sum_{i=1}^N (k_{omi} P_i \Delta t) \quad (i = 1, 2, \dots, N) \quad (14)$$

$$C_{grid} = C_p E_p - C_s E_s \quad (15)$$

$$C_{pv} = C_b P_{PV} \Delta t \quad (16)$$

where C_{gas} is the natural gas price; N means the number of equipment; k_{omi} and P_i are the maintenance coefficient and the output power, respectively, of equipment i ; C_p is the price consumers purchasing electric energy (E_p) from the grid and C_s is the price consumers selling electric energy (E_s) to the grid; C_b is the subsidy rate of PV.

(2) Operating Constraints

The operating constraints consist of below two parts: the power balance constraints as expressed in (17)–(19), and the capacity limitations as (20).

$$P_L = P_G + P_{PV} + P_{ICE} \quad (17)$$

$$Q_{HE} = Q_{GF} + Q_{EG} + Q_{JW} \quad (18)$$

$$Q_{CO} = Q_{AC} \quad (19)$$

$$P_i^{\min} \leq P_i \leq P_i^{\max} \quad (20)$$

where P_L , Q_{HE} and Q_{CO} are the electricity demand, heating demand and cooling demand, respectively. P_G is the electric power supplied by the grid.

3.2 Algorithm of Optimizing Dispatch

The CCHP system design normally adopts the heat demand based scheme or the electric demand based scheme. The former is used in the paper. Thus, the proposed model is a nonlinear programming problem in fact. The Particle Swarm Optimization method (PSO) is very suitable for solving this type of problems. The principle of PSO is to search the optimal value of fitness by updating the positions of particles in one after another iteration. It is better for the whole searching when inertia weight is lesser while it is better for local searching when inertia weight is more. This paper improves the PSO as shown in (21).

$$\omega_k = \omega_{\text{start}} - (\omega_{\text{start}} - \omega_{\text{end}}) \left(\frac{k}{M} \right)^2 \quad (21)$$

where ω_k is the inertia weight k th iterating; ω_{start} and ω_{end} are the inertia weights at the start and at the end, respectively, and $\omega_{\text{start}} = 0.9, \omega_{\text{end}} = 0.4$ in this paper; M means the maximum iteration times.

Figure 2 shows the flow diagram of improved PSO. The iterative equations of particles' positions and speeds are depicted in Eqs. (22)–(23).

$$v_k = \omega_k v_{k-1} + c_1 r_{1,k} (p_{k-1} - n_{k-1}) + c_2 r_{2,k} (g_{k-1} - n_{k-1}) \quad (22)$$

$$n_k = n_{k-1} + v_k \quad (23)$$

where v_k is the speed of particles in k th iteration. c_1 and c_2 are the acceleration factors of individual particle and global particles, respectively. p_{k-1} and g_{k-1} are the positions

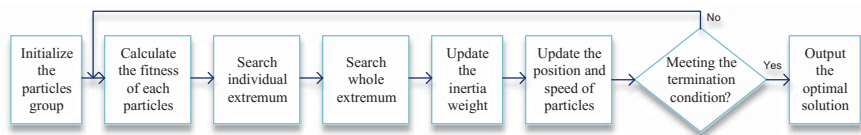


Fig. 2 The flow diagram of improved PSO

of individual extremum particle and global extremum particles, respectively, in k th iteration. n_{k-1} is the position of the particle in k th iteration.

4 Algorithm Implementation and Results of the Case Study

4.1 Algorithm Implementation

The algorithm is implemented by MFC in Visual C++ 6.0, including three steps:

- (1) Parameters setup: Input the forecast values of electric load, heating load and cooling load based on historical load data. Then, set the price parameters, such as the natural gas price, the price consumers purchasing electric energy from the grid, the price consumers selling electric energy to the grid, etc. Finally, choose season.
- (2) Algorithm operation: According to the given parameters, the optimal dispatch of each energy generator will be obtained by the proposed model.
- (3) Results output: The optimal dispatch results will be demonstrated on the interface, which is not shown here for the sake of space.

4.2 Case Study

A 2,000 m² office building with electric, heat and cooling loads is used to conduct the case study.¹ The power grid voltage is 0.4 kV with TOU pricing. The solar PV panels with a maximum rated power of 50 kW and one ICE rated at 100 kW are installed. The natural gas has a fixed price. Local energy prices are shown in Table 1. Figures 3 and 4 show the optimization results for the four seasons.

Table 1 The price parameters of different kinds of energies

	Time (h)	Price (¥)
Natural gas	0–24	1.98
Price of purchasing	Peak period: 8:00–11:30 and 18:30–23:00	0.713
	Plain period: 11:30–18:30	0.49
	Valley period: 0:00–7:00 and 23:00–24:00	0.267
Price of selling	0–24	0.85
Subsidy standard of PV	0–24	0.42

¹This study has been permitted by School of Electrical Engineering, Chongqing University Chongqing, China, for carrying out in the office building.

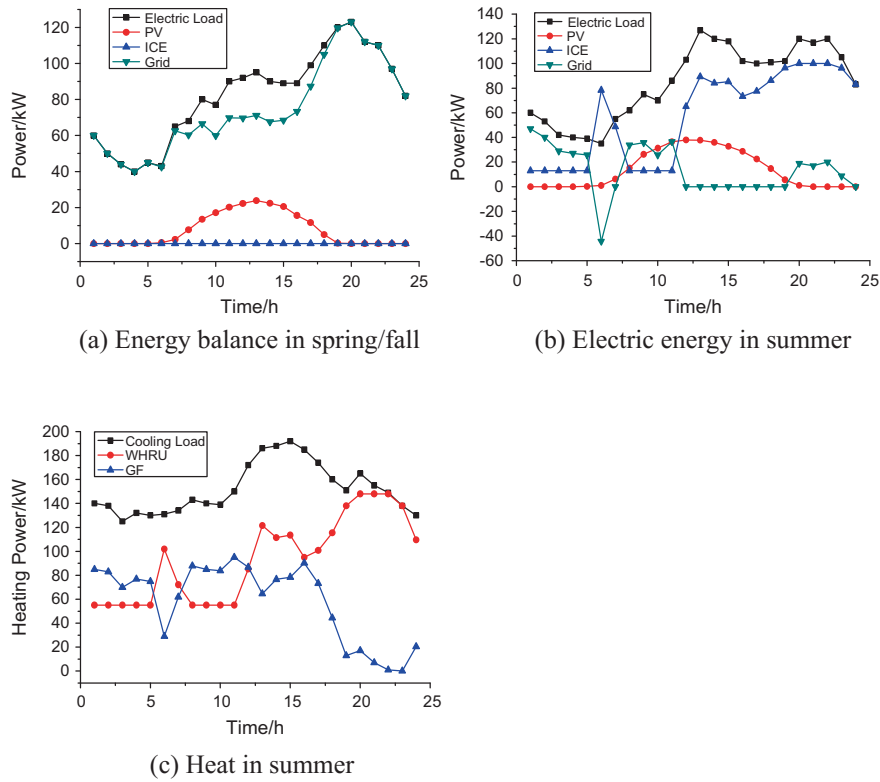


Fig. 3 Energy balance in spring/fall and summer

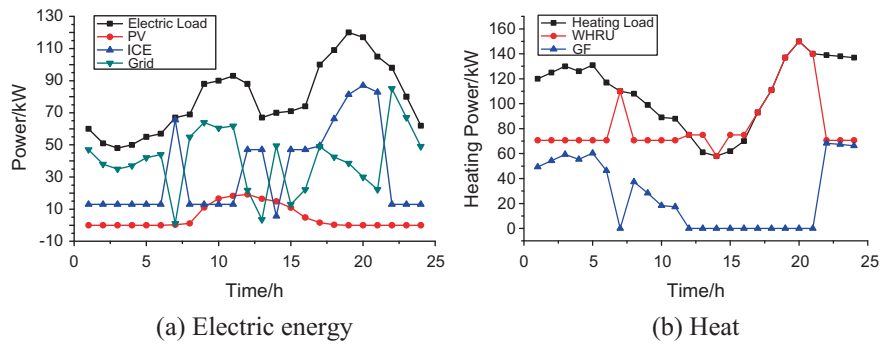


Fig. 4 Energy balance in winter

Table 2 The contributions of different energy generators in different seasons

Energy types	Energy generators	Spring/Fall	Summer (%)	Winter (%)
Electric	PV	10%	19	6
	ICE	0	63	42
	Grid	90%	18	52
Heat	ICE	0	60	76
	WHRU	0	40	24

In spring and fall, ICE does not participate in scheduling because the users in this building have no heating and cooling demand as shown in Fig. 3a. In this case, the cost for electric demand is more expensive using ICE in comparison with purchasing electric power from the grid directly when the same power is required.

In summer, as shown in Fig. 3b, ICE keeps constant output during 0:00–5:00 which is in the valley period. At the 6th hour, the output power of ICE increases dramatically when electric demand is least in this day and during this period, users sell electric energy to the grid. During 7:00–11:00 of the peak period, ICE still keeps the same output as 0:00–5:00. Users neither purchase nor sell electric energy to the grid during 12:00–19:00. For the heat, the output heating power and output electric power of ICE change simultaneously, while the GF supplements the remaining heating demand, as shown in Fig. 3c.

In winter, as shown in Fig. 4, users do not sell electric energy to the grid in a whole day. During 0:00–6:00 and 8:00–11:00, ICE keeps a constant output power which is in the valley period. At 7:00, users purchase neither electric energy nor sell to the grid. After 22:00, users purchase electric energy from the grid due to the low price. However, to satisfy users' electric demand, the produced heat by ICE is more than users' heat demand at 13:00 and 15:00, which has to be abandoned. Furthermore, GF does not work from 12:00 to 21:00.

Table 2 presents a summary of the case study for the contributions of each energy generators in different seasons. In summer, ICE plays an important role in energy supply which provides users with 63% of electric energy and 60% of thermal energy. In winter, because of less solar irradiance, PV just satisfies 6% of electrical demand of the building and more than half of electrical demand is satisfied by the grid.

In Table 3, the cost with the optimization model is compared to that with the traditional energy supply method in 1 day. In a spring or fall day, 16.8% of the cost is saved with the proposed optimization model. 18.5 and 12.6% of the cost are saved on a summer day and on a winter day separately.

Table 3 Comparisons of daily cost with optimization model and without optimization model

Seasons	Cost with optimization model (¥)	Cost without optimization model (¥)	Percentage of cost saving (%)
Spring/Fall	1065.2	886.3	16.8
Summer	1640.8	1336.7	18.5
Winter	1403.7	1226.5	12.6

5 Conclusions

This paper proposes an improved optimization model for the combined CCHP and PV system. This model is able to deal with the optimization problem caused by both varied PV and load demands, including electric, heat, and cooling. This algorithm is especially suitable for the small-scale CCHP combined with local installed PV. A case study shows that the proposed algorithm presents concise results for day-ahead management of this type of system.

References

1. Dai, X.Y., Mao, D.J.: Operation optimization method of distributed system for combined cooling heating and power. *J. Electric Power Sci. Technol.* **32**(1), 55–64 (2017). (in Chinese)
2. Hu, Y.F., Wu, J.Y., Li, S.: Optimal operation analysis of combined cooling heating and power system. *J. North China Electric Power Univ.* **37**(1), 5–9 (2010). (in Chinese)
3. Chen, J., Yang, X., Zhu, L., et al.: Comparison of microgrid economic operation among different dispatch modes. *Electric Power Auto. Equip.* **33**(8), 105–113 (2013). (in Chinese)
4. Jing, Y.Y., Bai, H., Zhang, J.L.: Multi-objective optimization design and operation strategy analysis of a solar combined cooling heating and power system. *Proc. CSEE* **32**(20), 82–87 (2012). (in Chinese)
5. Liu, X.Y., Wu, H.B.: A control and operation optimization of combined cooling heating and power system considering solar comprehensive utilization. *Auto. Electric Power Syst.* **39**(12), 1–6 (2015). (in Chinese)
6. Wang, F., Zhou, L., et al.: Optimized operation model for stand-alone CCHP microgrid. *Appl. Sci.* **7**(754), 1–15 (2017)
7. Chen, P., Guan, L., et al.: An optimal planning method for combined cooling heating and power system. *Energy Proc.* **103**, 123–128 (2016)
8. Yang, G., Zhai, X.Q.: Optimization and performance analysis of solar hybrid CCHP systems under different operation strategies. *Appl. Therm. Eng.* **133**(25), 327–340 (2018)
9. Lorestani, A., Ardehali, M.M.: Optimal integration of renewable energy sources for autonomous tri-generation combined cooling, heating and power system based on evolutionary particle swarm optimization algorithm. *Energy* **145**(15), 839–855 (2018)
10. Li, L., Xiao, X.N., Chen, P.W.: Improved photovoltaic model and its application in reliability evaluation of microgrid. *Adv. Technol. Electrical Eng. Energy* **35**(11), 65–71 (2016). (in Chinese)

Control and Optimization of Indoor Environmental Quality Based on Model Prediction in Building



Anjun Zhao, Meng Zhou, Junqi Yu, Junlin Zhang and Xiong Yang

Abstract Control and optimization of the quality of the indoor environment are necessary to ensure indoor comfort and reduce building energy consumption. Indoor environmental quality that contains a variety of uncertainties and nonlinear factors is difficult to be described by the traditional linear system. In this paper, by defining the linear relationship between physical parameters and control parameters of the indoor environmental quality, the control, and energy consumption optimization modeling is established according to the data measured based on a bilinear model. On this basis, this study proposes a model predictive control system coupled with an intelligent optimizer for indoor environmental quality control. Ant colony optimization (ACO) is utilized to optimize the building energy management. Experimental results show that the proposed intelligent control system successfully manages indoor environmental quality and energy conservation.

Keywords Indoor environmental quality · Bilinear model · Model predictive control · Colony algorithm

1 Introduction

At present, people spend 80% of their time indoors, so indoor environmental quality has a very important impact on human health [1]. Indoor air quality can be improved

A. Zhao (✉) · M. Zhou · J. Zhang · X. Yang
School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi, China
e-mail: zhao_anjun@163.com

J. Zhang
e-mail: 393798181@qq.com

J. Yu
School of Information and Control Engineering, Xi'an University of Architecture and Technology Smart City Research Center, Xi'an, Shaanxi, China
e-mail: junqiyu@126.com

by ventilation, but will also have impacts on the indoor thermal comfort and result in the waste of energy. At the same time, most of the air conditioners adopt simple adjustment methods and cannot be adjusted according to the change of the outdoor temperature. Therefore, how to ensure good indoor environmental quality and meet the requirements of lower energy consumption has become the main problem of indoor environmental quality control and optimization study [2].

At present, for controlling the quality of the indoor environment, Hellwig proposed a method of perceived control, that is, controlling according to the body's own feelings, and emphasizing the importance of personal control from a psychological point of view [3]. Liu Chang proposes two important goals of indoor environmental control are indoor thermal comfort and energy saving. In order to achieve comfort control, the predicted mean vote (PMV) is designed as the control goal and uses the precise formula of PMV to control the optimization. On the other hand, PMV is designed based on a genetic algorithm to achieve better performance and overcome nonlinear characteristics [4].

As the most successful advanced control strategy, the Model Predictive Control (MPC) is a powerful technique for generating feasible trajectories and has attained great success in the field of control engineering over the previous decades [5]. Moria proposes the use of a recently developed robust, adaptive MPC as a possible solution for handling this model uncertainty [6].

The conventional iterative optimization methods are very sensitive to the initialization of the algorithm and usually lead to unacceptable solutions due to the convergence to local optima [7], so the intelligent evolutionary algorithms are more suitable for the optimization in MPC, such as genetic algorithm, ant colony optimization, and particle swarm optimization [8, 9].

In this paper, a bilinear model is used to reflect the relationship between indoor environmental quality control and energy consumption optimization based on split-type air conditioning and independent ventilation equipment. Bilinear model is a class of the nonlinear model, which has the most simple form and also closest in the linear model [10]. By identifying the parameters of the model, the model can better reflect the relationship between indoor environmental quality outdoor climate and building characteristics. At the same time, MPC and ACO are utilized to complete the control and optimization of indoor environmental quality.

2 The Model of Control System

2.1 Notations

$CO_{2in}(k)$	the indoor carbon dioxide concentration at the moment of k
$T_{in}(k)$	the indoor temperature at the moment of k
$CO_{2out}(k)$	the outdoor carbon dioxide concentration at the moment of k
ρ	the parameters that need to be identified in the model

$T_{out}(k)$	the outdoor temperature at the moment of k
$w(k)$	the ventilation fan ventilation at the moment of k
$A(k)$	the refrigeration and heating of the air conditioning system at the moment of k
$X(s)$	the comfort set vector
$CO_2(s)$	the set of indoor carbon dioxide concentrations
$T(s)$	the set of indoor temperature
∂_1, ∂_2	the relationship between the first moment of carbon dioxide concentration and the temperature parameter in the outdoor climate and the next moment
β	the model parameters
U_A	the open degree of the air conditioner
U_w	the wind turbine control signal
G	the coefficient matrix of the different control actuators in terms of energy consumption
V	the coefficient matrix of thermal comfort and air quality comfort in comfort parameters
q_0	the adjustable parameters ($0 \leq q_0 \leq 1$)
d_{ij}	the desirability of edge (i, j)
τ_{ij}	the pheromone
Q_K	the collection of control points and estimated step points that the ant k has not yet passed
γ	the pheromone volatile factor
W_{CO_2}	the weight of the CO_2 concentration
W_{UW}	the energy consumption of the air conditioner
W_{UA}	the energy consumption of the air conditioner.

2.2 The Indoor Environment Quality Control Mathematical Model

The block diagram of the model predictive control system is depicted in Fig. 1, where ACT represents the implementation of improving the quality of the indoor environment, including air conditioner equipment and exhaust fans. BEMS is a building energy management system that completes the collection of indoor environmental quality parameters and controls the actuator. Control system of indoor environmental quality based on model prediction can be described as the following equations.

$$\text{State equation: } x(k+1) = f(x(k), u(k), d(k)) \quad (1)$$

$$\text{Reference vector: } y(k) = x(k) \quad (2)$$

$$\text{Control vector: } u(k) = g(x(s), y(k)) \quad (3)$$

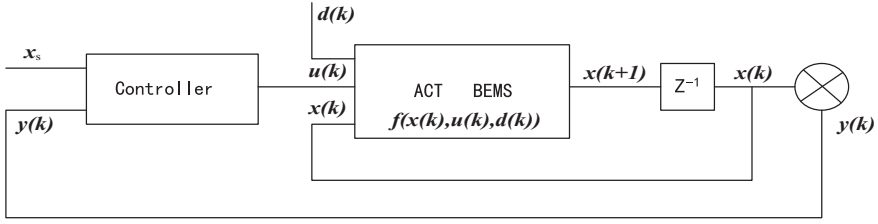


Fig. 1 Block diagram of the indoor environmental quality control system

In order to simplify the identification process of the system, the model of indoor environmental quality can be expressed as follows without considering the existence of noise.

$$x(k+1) = x(k) + \rho_1 f_1(x(k), u(k), d(k)) + \cdots + \rho_n f_n(x(k), u(k), d(k)) \quad (4)$$

$$x(k) = [CO_{2(in)}(k), T_{in}(k), CO_{2(out)}(k), T_{out}(k)] \quad (5)$$

$$u(k) = [w(k), A(k)] \quad (6)$$

$$x_s = [CO_{2(s)}, T_s] \quad (7)$$

The bilinear model is described in different forms for indoor carbon dioxide and temperature state variables.

The indoor carbon dioxide concentration at the moment of $k+1$ is a linear function of the concentration of indoor carbon dioxide and ventilation fan ventilation at the moment of k , it is defined as follows:

$$CO_{2in}(k+1) = CO_{2(in)}(k) + \alpha w(CO_{2out}(k) - CO_{2in}(k)) \quad (8)$$

Among them, α is the model parameter that needs to be identified. The indoor temperature at the moment of $k+1$ is the linear function of the indoor and outdoor temperature, ventilation fan ventilation and cooling and heating of air conditioner at the moment of k :

$$\begin{aligned} T_{in}(k+1) = & T_{in}(k) + \beta_1(T_{out}(k) - T_{in}(k)) + \beta_2 A(k) \\ & + \beta_3 w(k)(T_{out}(k) - T_{in}(k)) \end{aligned} \quad (9)$$

The control system model based on model prediction can be expressed as follows:

$$x(k+1) = Bx(k) + Cx(k)w(k) + Fu(k) \quad (10)$$

$$\text{where } B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \beta_1 & 0 & \beta_1 \\ 0 & 0 & \partial_1 & 0 \\ 0 & 0 & 0 & \partial_2 \end{bmatrix}, C = \begin{bmatrix} -\alpha & 0 & \alpha & 0 \\ 0 & -\beta_3 & 0 & \beta_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, F = \begin{bmatrix} 0 & 0 \\ 0 & \beta_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The outdoor variables can be measured with temperature sensors and carbon dioxide sensors. The ∂_1, ∂_2 can be modeled by dummy equations.

The model identification was performed at the Intelligent Building Laboratory of Xi'an University of Architecture and Technology. The laboratory has 110 m² areas with almost 2.7 m height. The Indoor air conditioning system which model is KF-72LW/Y-Sx (E) by Midea is the split floor type air conditioner with cooling capacity of 7200 W and cooling power of 2820 W. The exhaust fan is FD400E by Baochang with step-less speed regulation and the power of 200 W and the maximum airflow is 4000 m³ per hour. Energy management system uses EHS building energy management and control platform, through the indoor temperature sensor, carbon dioxide concentration sensor to collect the indoor environmental parameters. The outdoor data is provided by the weather station system based on the crossbow.

During the experiment, the air conditioning control is performed by adjusting the set temperature of the air conditioner from 16 to 26 °C. The air volume is constant and set to the medium speed. Corresponding to the control signal of the air conditioner energy consumption management and control platform, the air conditioning output is $U_A = \{-1, -0.9, \dots, 0, 0.1, \dots, 0.9, 1\}$ and the control precision is 1 °C. The exhaust fan control is performed by adjusting the airflow of the fan from 0 to 4000 m³. Corresponding to the energy consumption management and control platform, the fan output is $U_w = \{0, 0.1, 0.2, \dots, 0.9, 1\}$.

As to the summer outdoor climate conditions in Xi'an (China), the energy management and control platform are set to work in a continuous mode in order to obtain the real data. The number of indoor staff is 6. In order to make the model to adapt to different control strategies, during identification process the position air conditioning and exhaust fan is moving gradually from its minimum to its maximum value for refrigeration, heating and exhaust with a 10% step in order to cover the overall range in the 24 h. The indoor environment parameters and control signal of the system as well as the outdoor temperature are observed and recorded. The sampled data are used as the sample data for the identification.

The experiment was performed by every 2 min for sampling from April 11–14, 2015. A total of 2882 sample points are collected as shown in Figs. 2 and 3.

The system identification is based on the least squares estimation [11]. The identification procedure consists of the estimation of the ρ_1 parameters of (4).

Based on (10), the system is described by the bilinear model of the form:

$$x(k+1) - x(k) = \rho_1 f_1(x(k), u(k), d(k)) + \dots + \rho_n f_n(x(k), u(k), d(k)) \quad (11)$$

or

$$x = \varphi \rho \quad (12)$$

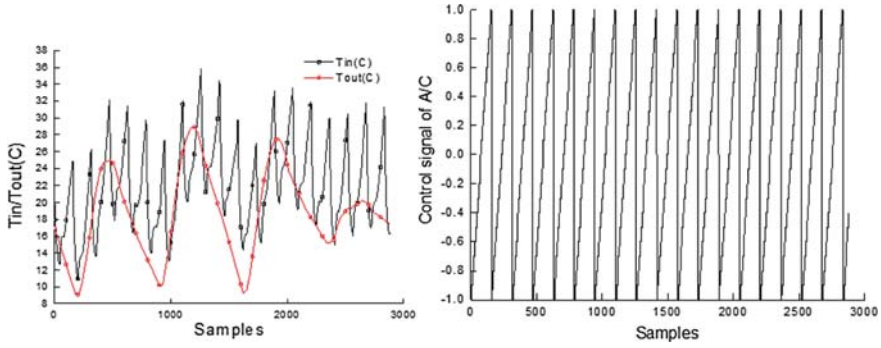


Fig. 2 Indoor and outdoor temperature for different air condition outputs

$$x = \begin{bmatrix} x(2) - x(1) \\ x(3) - x(2) \\ \vdots \\ x(l) - x(l-1) \end{bmatrix} \quad \varphi = \begin{bmatrix} f_1(x(1), u(1)) & \dots & f_n(x(1), u(1)) \\ f_1(x(2), u(2)) & \dots & f_n(x(2), u(2)) \\ \vdots & \ddots & \vdots \\ f_1(x(l-1), u(l-1)) & \dots & f_n(x(l-1), u(l-1)) \end{bmatrix},$$

$$\rho = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{bmatrix}, \quad \hat{\rho} = (\varphi^T \varphi)^{-1} \varphi^T x \text{ is the estimated parameters.}$$

Then, the estimated value of the indoor environmental quality at the next moment is $\hat{x}(k+1)$, which can be expressed by the current state $x(k)$ and the estimated parameters $\hat{\rho}$:

$$\hat{x}(k+1) = x(k) + \hat{\rho}_1 f_1(x(k), u(k)) + \dots + \hat{\rho}_n f_n(x(k), u(k)) \quad (13)$$

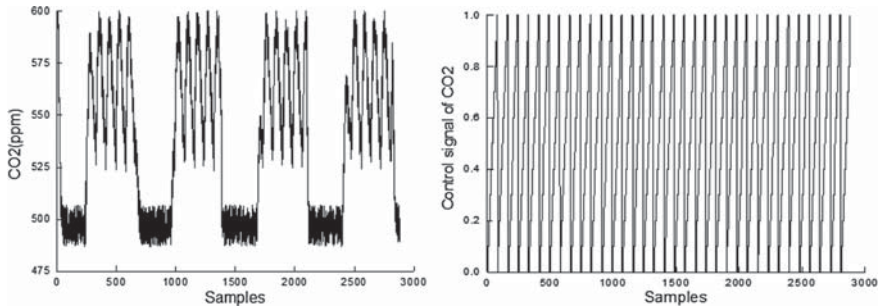


Fig. 3 Indoor CO₂ concentration for different fan opening outputs

Based on the experimental data, the estimated values of the model parameters are obtained by the least square estimation of indoor environmental quality model: $\hat{a} = 0.274$, $\hat{\beta}_1 = 0.0014$, $\hat{\beta}_2 = 0.0937$, $\hat{\beta}_3 = 0.0026$.

3 Analysis and Validation of Model

The experimental environment used in model verification is basically the same as data acquisition. The verification is done by comparing the actual measured indoor environment quality parameters with the predicted values of the model outputs. The real value is collected by the energy management and control platform by the preset of the air conditioner and fan control mode. The experiment was performed from 0:00 to 24:00 on September 17, 2015 and sampled every 2 min for a total of 722 sampling points. After the real data is completed, the predicted data is calculated at the same data of the environment for outdoor temperature and indoor carbon dioxide concentration. The real data and model predicted data are shown in Fig. 4 and Fig. 5 respectively.

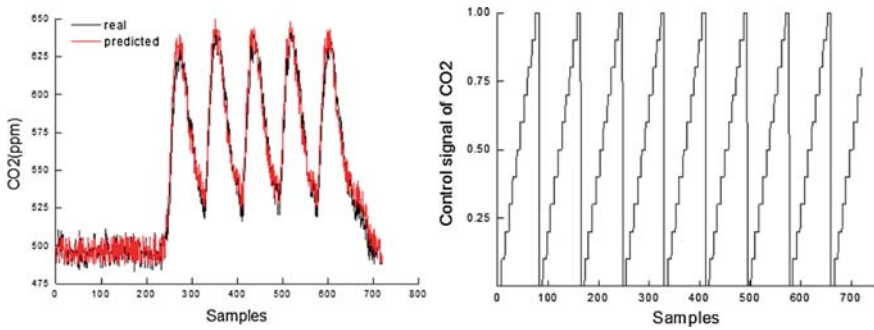


Fig. 4 Real and predicted values of CO₂ concentration for different fan opening outputs

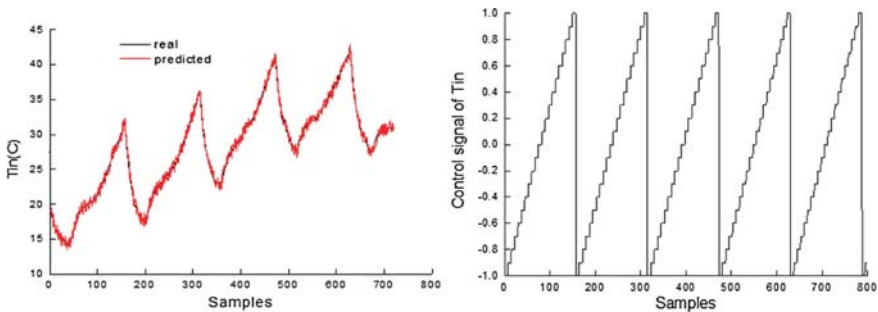


Fig. 5 Real and predicted values of indoor temperature for different air condition outputs

For intelligent building laboratory, in the case of air conditioner off, Fig. 4 shows the comparison of the real data and the predicted data for the concentration of carbon dioxide in different air volumes and in different sampling points. Because there was no staff at night, the concentration of carbon dioxide is changed little in the room between 00:00 and 08:30 and the same as the outside concentration. After 8 o'clock, the laboratory staff kept in six people and indoor carbon dioxide concentration changes periodically in accordance with the opening degree of the fan.

The experimental data show that the predicted values of indoor carbon dioxide concentration are basically consistent with real measurement. The correlation coefficient between the measured value and the predicted value is equal to 0.889, while the root mean square error (RMSE) is equal to 2.789 ppm. Figure 5 shows the comparison of the indoor temperature when the exhaust fan off and only in the case of air conditioner working at different times through the model predictions and the actual measurement of the intelligent building laboratory. It can be seen from Fig. 4, whether it is the real measurement value or the model estimates, the change of the law is affected by the outdoor temperature and the role of indoor air conditioner and show a periodical change. The experimental results show that the indoor temperature of the model predicted and the actual measured values are basically consistent. When the root mean square error is 0.378 °C, the correlation coefficient between the measured value of the indoor temperature and the predicted value is 0.9287.

4 Model Predictive Control and Optimization

The objective of the indoor environment quality control is to make the energy consumption of air conditioners and fans lower in obtaining better indoor comfort. Constructing the objective function [12, 13] as follows:

$$\min_{u \in U} J(k) = V \|x_{in}(k+N) - x_s\|^2 + G \|u(k)\|^2 \quad (14)$$

where $x_{in} = [CO_{2(in)}(k), T_{in}(k)]$, $u \in U$. The discrete control state space of the air conditioner is $U_A = \{-1, -0.9, \dots, 0, 0.1, 0.2, \dots, 0.9, 1\}$, the discrete control state variable of the fan output is $U_W = \{0, 0.05, 0.1, \dots, 0.95, 1\}$, $V = \alpha(w_{CO_2}, w_{T_{in}})$, $G = (1 - \alpha)(w_{U_A}, w_{U_W})$, α represents the weight coefficient of the importance of comfort and energy saving in the optimization process.

The controller parameters in model prediction are optimized by using ant colony algorithm. According to the state space of the control variable and the fixed distance dividing method [14], the control variables U_A , U_W , predictive time domain are divided into 20 equal points, expressed $L_1 = \{L_{11}, L_{12}, \dots, L_{1n}\}$, $L_2 = \{L_{21}, L_{22}, \dots, L_{2n}\}$, $L_3 = \{L_{31}, L_{32}, \dots, L_{3n}\}$, where L_1 represents the Control quantity U_A , L_2 represents the amount of control, U_W , L_3 represents the time domain for the prediction and $n = 20$. Using the equal points, which are divided by the control variables and the estimated time domain as nodes in ant colony algorithm, the con-

nection between them represents the path of the ant walking path. The ant colony walking environment is constructed by setting the starting point and the ending point.

The optimization of the ant colony algorithm is to find the optimal path, which includes the control point of the air conditioner and the fan and the amount of the control points that need to be advanced. Assuming there are m ants in total from the S point of departure to reach the end point T , the path of the cycle is $S \rightarrow L_{1n} \rightarrow L_{2n} \rightarrow L_{3n} \rightarrow T$. The ant selects the next node among list to move, which is based on the following rules [15]:

$$s = \begin{cases} \arg \max_{y \in Q_k} \{ \tau_{iy}(t) J(i, j)^\beta \} & \text{if } q \leq q_0 \\ S & \text{other} \end{cases} \quad (15)$$

where i is a collection of all points on the line, q is an equally distributed random number in region $[0, 1]$, q_0 is an adjustable parameters ($0 \leq q_0 \leq 1$), $D(i, j)$ is the heuristic value, $\tau_{i,j}$ is the pheromone, Q_k represents the collection of control points and estimated step points that the ant k has not yet passed. S is a random variable and given by the following equation:

$$P_{ij} = \frac{\tau_{ij}^\alpha d_{ij}^\beta}{\sum_{i,j=1}^n \tau_{ij}^\alpha d_{ij}^\beta} \quad (16)$$

Global pheromone update means when all ants have completed a tour, select the best path of from all the ants passed and update the path on each side of the pheromone, the update process in accordance with the fellow equation to complete:

$$\tau_{ij}(t) \leftarrow (1 - \gamma) \tau_{ij}(t) + \gamma \Delta \tau_{ij}(t) \quad (17)$$

where $\Delta \tau_{ij}(t) = \begin{cases} \frac{Q}{J(i,j)} & ((i,j) \in \text{globally best tour}) \\ 0 & \text{other} \end{cases}$, Q is a problem parameter, $J(i, j)$ is computed by the indoor environmental quality control.

5 Analysis of Experiment and Result

Optimization control and data acquisition have the same experimental environment, in which the indoor CO_2 concentration is 600 ppm and the temperature is 26° in summer and 20° in winter. So set the variable X_S for $X_S = [600T_{SP}]$.

For the coefficient matrix V , in order to make the contribution of different indoor environment quality parameters to the comfort are same in the process of the optimization calculation, the weight of CO_2 concentration and indoor temperature should be normalized. Normalization is performed by dividing the maximum value from the set-point vector with the set point for each variable and the indoor temperature is

calculated by dividing by 25 °C, which represents the average outdoor temperature condition, so the value of W_{CO_2} and $W_{T_{in}}$ can be shown as follows:

$$W_{CO_2} = \frac{600}{600} = 1, \quad W_{T_{in}} = \frac{600}{25} = 24 \quad (18)$$

The specific indoor temperature weight is as follows:

$$W_{T_{in}} = \begin{cases} 0 & \text{if } 21^\circ\text{C} \leq T_{in} < 26^\circ\text{C} \\ 20 & \text{if } T_{in} \geq 26^\circ\text{C or } T_{in} > 21^\circ\text{C} \end{cases} \quad (19)$$

The weight of CO_2 concentration is as follows:

$$W_{CO_2} = \begin{cases} 0 & \text{if } CO_{2in} < 600 \text{ ppm} \\ 1 & \text{if } CO_{2in} \geq 600 \text{ ppm} \end{cases} \quad (20)$$

For the coefficient matrix G, in the process of the fan and air conditioner, the energy consumption of the fan is smaller than the air conditioner system. The energy consumption of the fan (w_{U_w}) is 1 and the energy consumption of the air conditioner (w_{U_A}) is 8.

In ant colony optimization, the number of ants is 20, the maximum number of iterations is 50 times, the adjustable parameter V is 0.85, the pheromone important factor is 1, the inspired importance factor is 3, the constant Q is 0.8 and the pheromone volatilization factor is 0.75.

Based on Ant Colony Optimization, winter condition experiment was carried out on January 17–19, 2016, summer condition experiment was carried out on June 12–15, 2015. The traditional PID control of the winter experiment was carried out on January 17–19, 2016, the summer condition of the experiment was carried out on June 18–20, 2015.

The daily control time from 8 in the morning to 22 in the evening and the sampling period is 2 min.

Figure 6 shows that in the winter working condition, the predictive control based on ant colony optimization and the traditional PID control in a similar outdoor climate condition. From the control effect, the MPC based on the optimization in the control accuracy is better than the traditional PID control, the control of the room temperature is basically 20 °C in winter and 26 °C in summer and the control error absolute value is below 0.2 °C. However, the traditional PID control under whether in summer or winter condition, the absolute value of the control error is about 1°, the amplitude of the temperature oscillation is larger. At convergence speed, the optimized MPC is slower than the traditional PID control method, the traditional PID control can converge through 10 sampling points and the former requires 15 samples to converge. This is mainly due to the optimization algorithm in determining the size of the control has considered not only the temperature and carbon dioxide concentration, while also taking into account the size of the energy consumption and resulting in a relatively slow convergence rate.

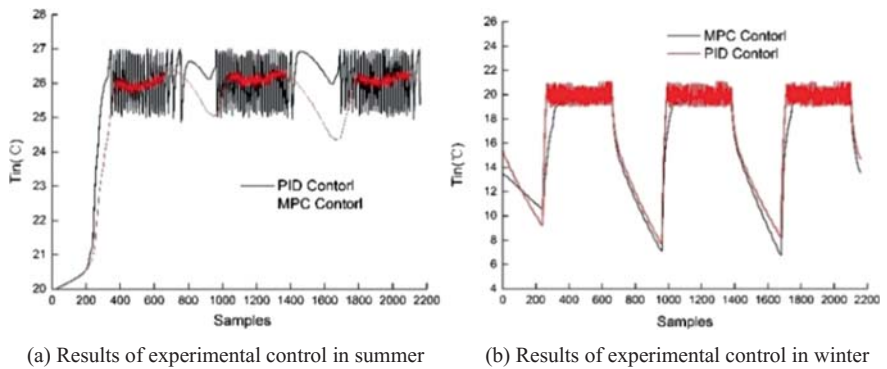


Fig. 6 Control effect in different working conditions

Fig. 7 Control effect of CO_2 concentration

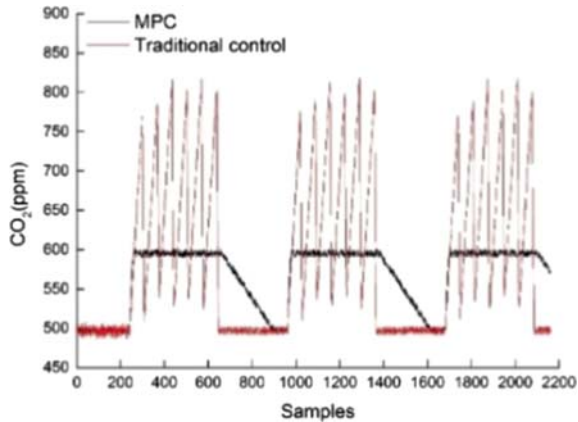


Figure 7 shows the control effect of the concentration of carbon dioxide in the summer condition. The traditional control method uses the timing fan opening way, every 2 h to open the fan once, 20 min each time and the fan running at full speed. The MPC based on ant colony optimization generates a fan speed control signal according to the sampling and estimation data and is controlled by BEMS. In view of the control effect, the timing control method is overshoot. When the fan is turned off, the carbon dioxide concentration increases quickly and exceed over 600 ppm. The concentration of carbon dioxide is decreased rapidly after the fan is opened and it is basically consistent with the outdoor carbon dioxide concentration. Based on the optimal control of the concentration of carbon dioxide keep in the vicinity of the set point after control convergence, the control effect is very ideal.

Figure 8 shows the different control scheme comparison of air conditioner energy consumption. Compared with the outdoor temperature change of the traditional control (PID control) and optimization control (model predictive control based on ant colony optimization), the outdoor temperature of the traditional control experiment

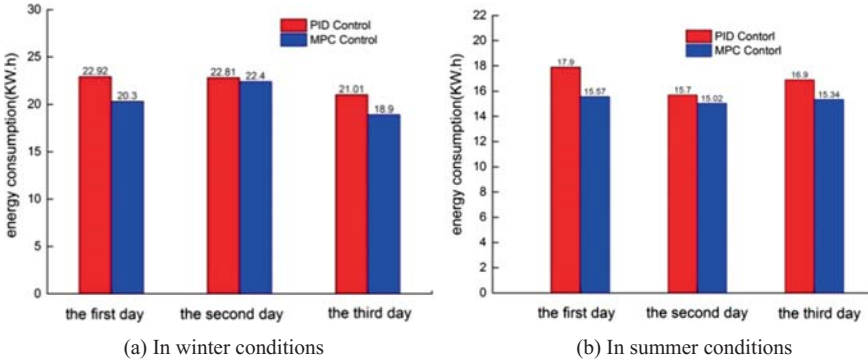


Fig. 8 Comparison of energy consumption in different working conditions

is higher than the optimal control under the winter and summer condition and the energy consumption should be smaller. However, the experimental results show that the energy consumption of the air conditioner system generated by the traditional control method is higher than the optimal control method. Under the traditional control mode, the total energy consumption of the air conditioner system in summer and winter condition was 118.34 kw h, the optimal control mode was 107.53 kw h, and the energy consumption increased by 10%.

6 Conclusions

The control and optimization which based on the model predictive ensure to have a good indoor environmental quality as well as meet the requirements of low energy consumption. At the same time, the control methods have a fast and stable response to outdoor climate change. In the model predictive control, solving the optimal solution takes a long time, using the ant colony algorithm can shorten the solution time, making it better suited to the weak computing ability of embedded systems.

For the energy consumption generated by the fan, the traditional control mode and the optimal control mode is essentially flat or even lower than the optimal control. This is mainly due to the timed open mode of the fan which the open time is relatively short and the total length of time is only 140 min. In two control modes, the fan energy consumption in the experiment is 11.95 kw h. In the total energy consumption, air conditioner energy consumption accounted or 95%, fan energy consumption accounted for 5%, fan energy consumption is much lower than the air conditioner energy consumption, so fan optimization in the control of energy consumption will not affect the operation of the entire system energy conservation.

Finally, the control performance is quite satisfactory and selects the optimum solutions based on energy consumption and the set-point proximity by satisfying the performance of the objective function. Future works need to consider the extended

comfort index, adding indoor light comfort into comfort index and research and analysis other prediction and optimization methods in order to make the control and optimization of performance can be improved.

Acknowledgements This work is supported by the National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques) and the National Natural Science Foundation of China (Grant No. 51508445).

References

1. Xi, Y., Li, D.W., Lin, S.: United Nations, transforming our world: the 2030 agenda for sustainable development (A/RES/70/1), UN General Assembly, New York, 2015. Model predictive control-status and challenges. *Acta Autom. Sinica* **39**(3), 222–236 (2013)
2. Pervez, H., Nursyar, B., Perumal, N., et al.: A review on optimized control systems for building energy and comfort management of smart sustainable buildings. *Renew. Sustain. Energy Rev.* **34**, 409–429 (2014)
3. Hellwig, R.T.: Perceived control in indoor environments: a conceptual approach. *Build. Res. Inf.* **43**, 302–315 (2015)
4. Chang, L., Yifei, C.: Neural computing thermal comfort index PMV for the indoor environment intelligent. In: *Proceedings of SPIE—The International Society for Optical Engineering*, vol. 8768 (2013)
5. Abdul, A., Farronkh, J.: Theory and applications of HVAC control systems a review of model predictive control (MPC). *Build. Environ.* **72**, 343–355 (2014)
6. Tanaskovic, M., Sturzenegger, D., Morari, M., et al.: Robust adaptive model predictive building climate control. *IFAC-PapersOnLine* **50**, 1871–1876 (2017)
7. Bououden, S., Karimi, H.R., Chadli, M.: Fuzzy predictive controller design using ant colony optimization algorithm. In: *IEEE Multi-Conference on Systems and Control*, Antibes, France, 8–10 Oct 2014
8. Núñez, C.E.C., Sáez, D., De Schutter, B., et al.: Multiobjective model predictive control for dynamic pickup and delivery problems. *Control Eng. Pract.* **32**, 73–86 (2014)
9. Lehmann, B., Gyalistras, D., Gwerder, M., et al.: Intermediate complexity model for model predictive control of integrated room automation. *Energy Build.* **58**, 250–262 (2013)
10. Zavala, V.M.: Real-time optimization strategies for building systems. *Industr. Eng. Chem. Res.* **52**(9), 3137–3150 (2013)
11. Chu-an, L.: Distribution theory of the least squares averaging estimator. *Econometrics* **186**(1), 142–159 (2015)
12. Bououden, S., Chadli, M., Karimi, H.R.: An ant colony optimization-based fuzzy predictive control approach for nonlinear processes. *Inf. Sci.* **299**(3), 143–158 (2015)
13. Ma, Y., Borrelli, F., Hencsey, B., et al.: Model predictive control for the operation of building cooling systems. *IEEE Trans. Control Syst. Technol.* **20** (3), 796–803 (2012)
14. Islam, T., Islam, M.E., Ruhin, M.R.: An analysis of foraging and echolocation behavior of swarm intelligence algorithms in optimization: ACO, BCO and BA. *Int. J. Intel. Sci.* **08**(01), 1–27 (2018)
15. Gonzalez-Pardo, A., Jung, J.J., Camacho, D.: ACO-based clustering for ego network analysis. *Future Gen. Comput. Syst.* **66**, 160–170 (2017)

Identifying Abnormal Energy Consumption Data of Lighting and Socket Based on Energy Consumption Characteristics



Liangdong Ma, Yiying Xu, Yugen Qin and Jili Zhang

Abstract The data quality of building energy consumption monitoring platform is generally not high and there are a lot of problem data. This paper proposes an identification method of implicit error energy consumption data based on the overall usage characteristic. In this method, we connect hourly energy consumption data into lines. According to the influencing factors of the building operation, we classify the energy-usage mode. By using the clustering method, we identify partial abnormal data. Then we count the slopes of historical energy consumption data characteristic lines and compare the time-varying characteristic lines of real-time with the historical characteristic lines under the same energy-usage mode. By using the energy consumption data of an office building, we verify the reliability of this method in identifying the abnormal energy consumption data of lighting and socket. This method improves the quality of data and will make the energy monitoring platform more efficient in building energy conservation.

Keywords Energy consumption data of lighting and socket · Energy consumption characteristic · Abnormal data · Identification · Cluster

1 Introduction

Since 2007, many cities in China have established the building energy consumption monitoring platforms in different scales. More than 9,000 buildings have established energy consumption monitoring platforms and accumulated rich building energy consumption data. At present, these data have not been fully utilized, resulting in a waste of a large number of resources and hindering the advancement of building energy management. The main reasons are as follows:

L. Ma (✉) · Y. Xu · J. Zhang

Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian, China
e-mail: liangdma@dlut.edu.cn

Y. Qin

School of Software Technology, Dalian University of Technology, Dalian, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_6

- (1) The classification models, monitoring contents and data storage formats of different energy consumption monitoring platforms are not uniform. And there is a lack of unified data standards and analysis methods for energy consumption monitoring platforms.
- (2) The system perceptual layer is subject to strong electromagnetic interference, the network structure is complex, diverse and the quality of network communication is uneven. The data quality of the building energy monitoring platform is generally not high and there are a large number of problem data, reaching more than 20%. Currently, there are no scientific identification and effective treatments for problem data, the monitoring results of a platform are distorted.

Therefore, realizing the identification and repair of problem data and improving the quality of data are essential for energy consumption conservation.

2 The Classification of Data

In energy consumption monitoring platform, there are two types of data, real data and problem data. As we can see from Table 1, the problem data is further divided into missing data or “0” data, mutational data and implicit error data. The problem data need to be fixed. Although the abnormal data reflect the real situation of energy usage, it does not reflect the actual energy consumption of the building at that time so the abnormal data need to be marked.

Figures 1, 2, and 3 further illustrate the characteristics of missing data and mutation data.

Table 1 Data types and characteristics in energy consumption monitoring platform

Data types		Data characteristics	Whether the data is repaired
Real data	Normal data	The energy consumption is normal, the data reflects the real situation	No
	Abnormal data	The energy consumption is abnormal, the data reflects the real situation	No, marked
Problem data	Mutational data	The data is obviously too large or too small to reflect the real situation	Yes
	Missing data	No data	Yes
	Implicit error data	The data do not meet the actual operating characteristics of the energy-using equipment	Yes

Fig. 1 Missing data

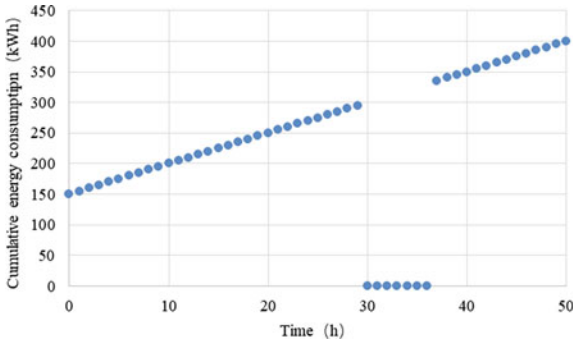


Fig. 2 Mutation data

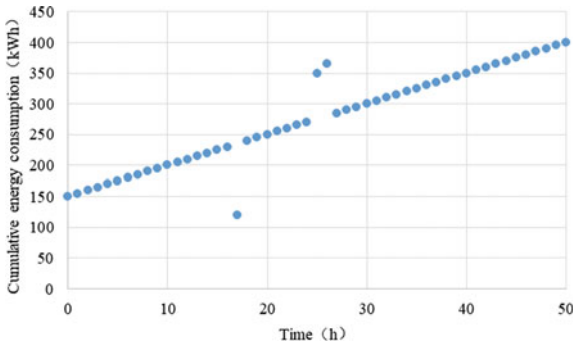
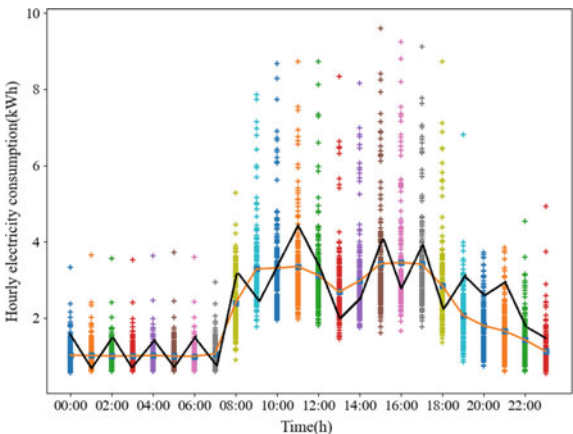


Fig. 3 Implicit error data
(dark line, orange line is
average hourly electricity
consumption)



3 Identification Method of Missing or Mutational Data

Missing data or '0' data are very easy to identify, their cumulative number of meters is 0 or there is no data. Mutational data is roughly divided into three categories, as shown in Figs. 4, 5 and 6. For noncontinuous mutation, according to the monotonically increasing characteristics of the cumulative number in the meter, Yang [1] identified mutation data based on the difference of cumulative data, the positive and negative energy consumption per unit time. It can be seen from Fig. 4 that if a sudden change occurs at a certain time, there must be a negative value in the energy consumption of the previous time or this time. Assuming that a sudden change occurs at time i , the mutation value is A_i and then returns to the normal value at time $i + 1$, there is bound to be $(A_i - A_{i-1})(A_{i+1} - A_i) < 0$.

For continuous up or down mutation, as shown in Figs. 5 and 6, set the upper and lower thresholds of the slope. The slope exceeding the threshold of the odd number of times is regarded as the start of the abnormal zone, the slope exceeding the threshold of the even number of times is regarded as the end. Remove all data in the abnormal zone.

Fig. 4 Noncontinuous mutational data

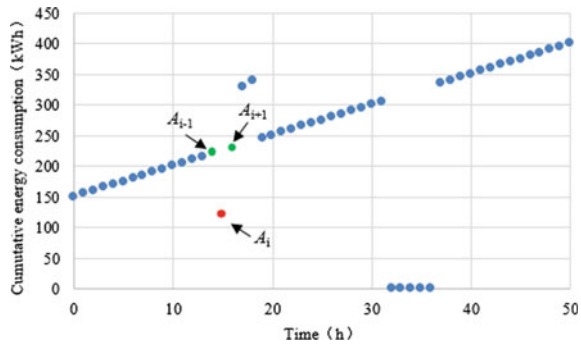


Fig. 5 Continuous up mutational data

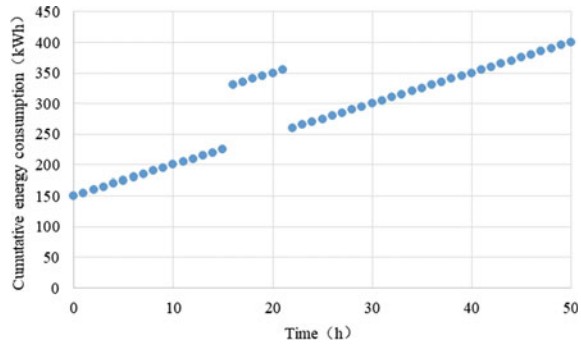
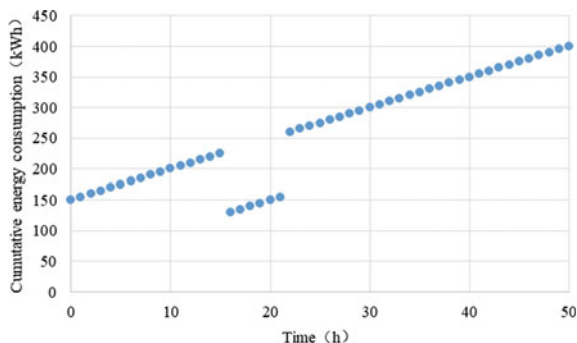


Fig. 6 Continuous down mutational data



4 Identification Methods of Implicit Error Data

The missing or mutational data are easy to identify but there are a lot of implicit error data in monitoring platform, they affect normal statistic of energy consumption and energy audit. It is hard to know the characteristic of implicit error data so they are hard to be identified. The threshold method and clustering method are commonly used and we propose a new identification method based on energy consumption characteristics, the details are shown below.

4.1 Threshold Method

The threshold method [2, 3] analyses the historical data and the experience of the project then set upper and lower limits to exclude abnormal data. The threshold method is simple and easy to understand. However, the threshold is difficult to be determined. If the threshold is too high, a number of false alarms are issued. If the threshold is too low, equipment or system failures go undetected [4]. Without considering the seasonal and regional environmental characteristics, the threshold value neglects the characteristics of the building and cannot dynamically adapt to change. Therefore, it is necessary to combine the characteristics of energy consumption to develop anomaly judgment criteria suitable for own platform. The threshold method can identify a part of dominant abnormal data, such as mutation data and missing data. But implicit error data are often impossible to be identified by the threshold method and it is not feasible to set the threshold for each building.

4.2 Clustering Method

Due to the threshold method can't identify the implicit error data, this study uses the clustering method to identify abnormal energy consumption data. The clustering algorithms that are commonly used are: prototype-based clustering, such as k-means, Learning Vector Quantization, Mixture-of-Gaussian, etc.; density-based clustering, such as DBSCAN; hierarchical clustering, such as AGNES [5]. In general, k-means is simple and applicable compared to other clustering methods.

The main idea of k-means is as follows [6]: the value of k and initial cluster centers whose number is k are given, each sample datum is divided into the nearest cluster by calculating the distance between itself and initial cluster centers. After all sample data have been allocated, the center is recalculated based on all data in this cluster, then the steps of assigning data and updating the cluster centers are iterated until the cluster center is invariable or the specified number of iterations is reached. The distance between two sample data is calculated by the Euclidean distance.

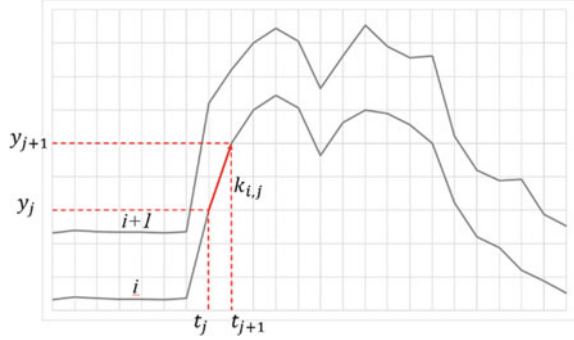
In the k-means algorithm, the choice of k is generally based on experience and multiple test results. For different data, the value of k has no reference. Through literature review and trial, it is found that hierarchical clustering [7] can select the value of k easily, avoiding the influence on clustering effects which caused by artificial selection. The Bisecting k-means clustering [8] avoids the initial clustering center to be selected into one data and overcomes the algorithm to fall into a local optimum state to some extent. So we choose hierarchically clustering to select the value of k and use Bisecting k-means clustering to identify implicit error data.

4.3 Identification Method Based on Energy Consumption Characteristics

Classification algorithms [9] mainly include decision tree, Bayes classifier and support vector machines. For a given building, under the conditions given by the service level, the energy consumption of the building is mainly affected by the building operation pattern (air conditioning equipment is also affected by weather). Therefore, this study classifies energy consumption into different energy-usage modes based on the influence factors. Energy consumption is not related to meteorological parameters, such as lighting and socket energy consumption and power energy consumption, their daily energy consumption can be divided into two types: weekdays and weekends. According to whether it is a weekday and 24 h, hourly energy consumption can be divided into 48 types (2×24); air conditioning energy consumption needs to consider factors of weather such as outdoor air temperature, solar radiation intensity, etc.

For the implicit error data shown in Fig. 3, it is usually not recognized by the above clustering method. Therefore, this study proposes an identification method

Fig. 7 Schematic diagram of the slope of the electric characteristic line



of invisible error data based on energy consumption characteristics. The specific process of the method is as follows:

- (1) Classify the energy-usage mode according to the influencing factors of the building operation;
- (2) Eliminate some abnormal data by clustering method;
- (3) Link hourly energy consumption data into characteristic lines, calculate the slope of historical energy consumption data characteristic line, compare the time-varying characteristic line of real-time with the historical characteristic line under the same energy-usage mode.

Suppose there are historical energy consumption characteristic lines whose number is m . For historical energy consumption characteristic line i , the energy consumption at time t_j is y_j , y_{j+1} at time t_{j+1} . As we can see from Fig. 7, the slope of the energy consumption characteristic line from t_j to t_{j+1} is $k_{i,j}$, the equation is as follows.

$$k_{i,j} = \frac{y_{j+1} - y_j}{t_{j+1} - t_j} \quad (1)$$

From t_j to t_{j+1} , the historical energy consumptions increased or decreased. Assuming that the number of the historical energy consumption characteristic lines which slope are greater than 0 is n . The proportion of characteristic lines whose energy consumptions increase from t_j to t_{j+1} is p , $p = n/m\%$. The proportion of characteristic lines whose energy consumption reduce at the time is q which is equal to $1 - p$. If $p > q$, the historical energy consumption shows an increasing trend from t_j to t_{j+1} , otherwise, it shows a reducing trend.

If there is a real-time energy consumption characteristic line, select historical energy consumption lines under the same energy-usage mode. Determine whether the trend of its energy consumption is consistent with the trend of historical data in each time period, count the number of time periods whose trend characteristic is consistent, assuming the number is a . If $a/24 < 10\%$, the energy consumption characteristic is abnormal (the proportion is determined by user), there are implicit abnormal error data.

5 Case Study

Taking the main building of Dalian University of Technology (Dalian, China) as an example, electricity consumption data on the fourth floor of the building are selected from February 29, 2016 to July 24, 2016 as the sample. Taking 24 h as X-axis, hourly electricity consumption of lighting and socket (kWh) as Y-axis, the electricity consumptions are sequentially drawn in the coordinate system as shown in Fig. 8 (the range of hourly electricity consumption is large).

5.1 Classify the Energy-Usage Mode

Figures 9 and 10 show the electricity consumption of the lighting and socket on weekdays and weekends from February 29 to July 24, 2016. It can be clearly seen from the figures that the hourly electricity consumption on weekends is significantly lower than that on weekdays. The overall electricity consumption characteristic on the weekday is obvious. There is also a high timely electricity consumption on weekends because people work sometimes.

In Fig. 9, the upper density of the sample data are discrete and the lower density is sparse. The working time of the office staffs are highly regular, the electricity consumption of the lighting and socket on weekdays should be reproducible. But we can see from Fig. 9 that there is big difference in hourly energy consumption data especially from 8:00 to 18:00. Therefore, the data are further classified by month to analyze the reason for the increase.

Figures 11, 12, 13, 14 and 15 show the hourly electricity consumption data in different months. The X-axis is 24 h and the Y-axis is the hourly lighting and socket electricity consumption (kWh). It can be seen from the figures that the discretionary of electricity consumption in weekdays becomes larger from June to July.

Fig. 8 The scatter diagram of hourly electricity consumption

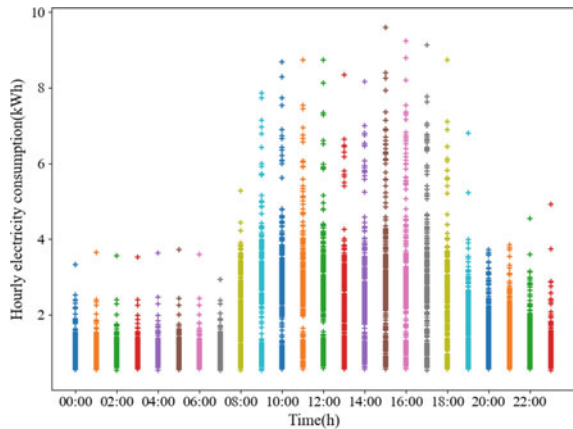


Fig. 9 The scatter diagram of hourly electricity consumption on weekdays

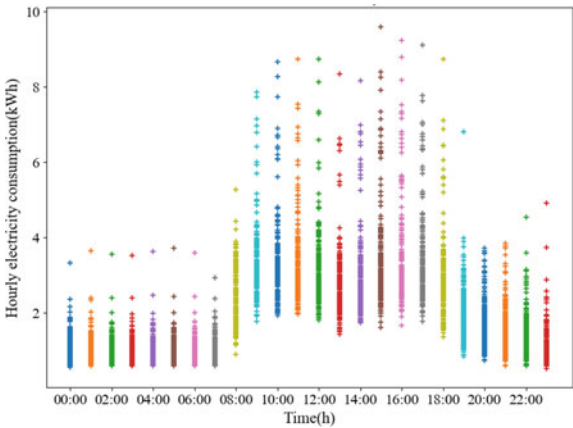


Fig. 10 The scatter diagram of hourly electricity consumption on weekends

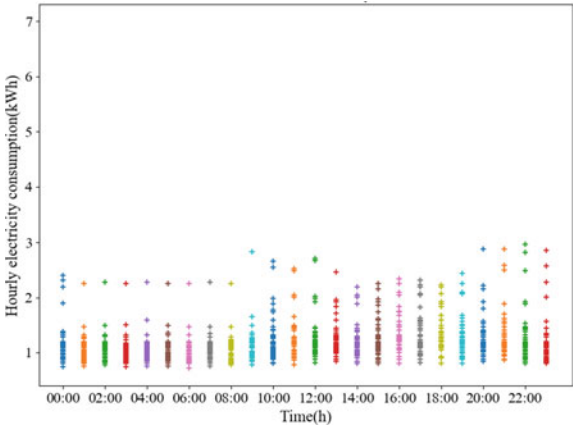


Fig. 11 Electricity consumptions in March

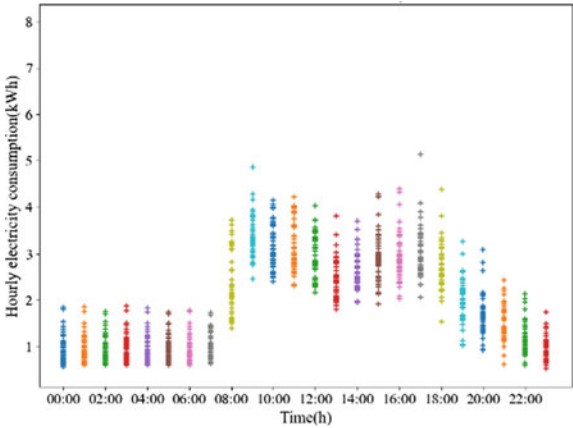


Fig. 12 Electricity consumptions in April

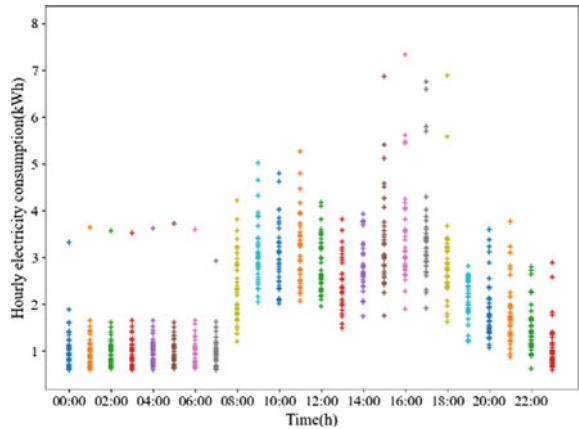


Fig. 13 Electricity consumptions in May

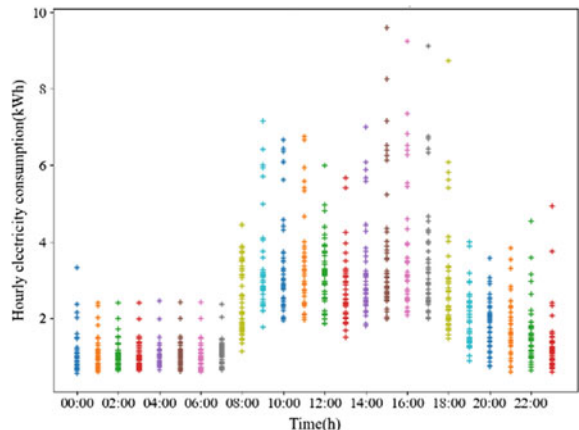


Fig. 14 Electricity consumptions in June

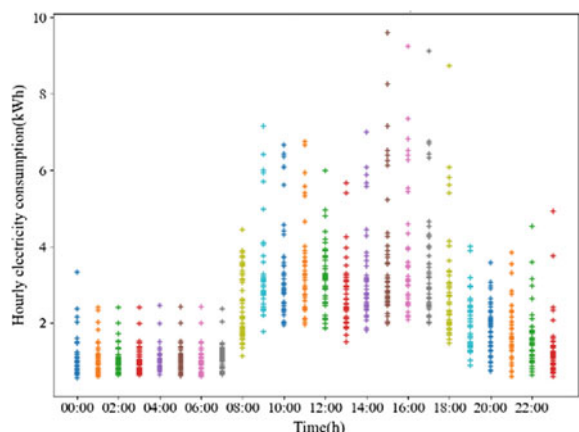


Fig. 15 Electricity consumptions in July

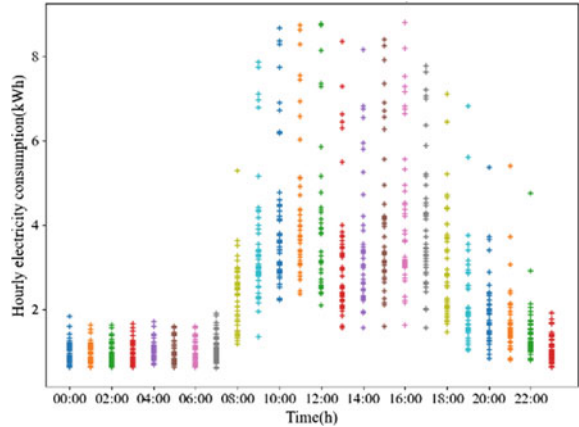


Table 2 PCCs between daily electricity consumption and daily average temperature

Month	PCCs
March	−0.102
April	−0.098
May	−0.001
June	0.383
July	0.491

We went to the site to inspect the end electrical equipment of the lighting and socket wiring. We find that the indoor air conditioner of VRV also connects with a wiring. In June and July, staffs need to use air conditioners, hourly electricity consumption data are larger than the corresponding period of March, April, and May. We study the relationship between the electricity consumption of lighting and sockets and the outdoor temperature through correlation coefficients. The Pearson correlation coefficient (PCCs) is obtained with the equation as follows:

$$\rho_{X,Y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} \tag{2}$$

In statistics, PCCs is used to measure the degree correlation (linear correlation) between two variables X and Y, with values between −1 and 1. PCCs between the daily electricity consumption of the lighting and sockets and the daily average temperature from March to July is shown in Table 2.

It can be seen from the table that PCCs is close to zero in March, April, and May, which increases in June and July due to the use of air conditioning.

Through the above data classification study, we take the electricity consumption data on weekdays of April and May, 2017 as an example. Connect hourly electricity consumption into lines, which are shown in Fig. 16.

Fig. 16 Electricity consumption characteristic lines on weekdays in April and May 2017

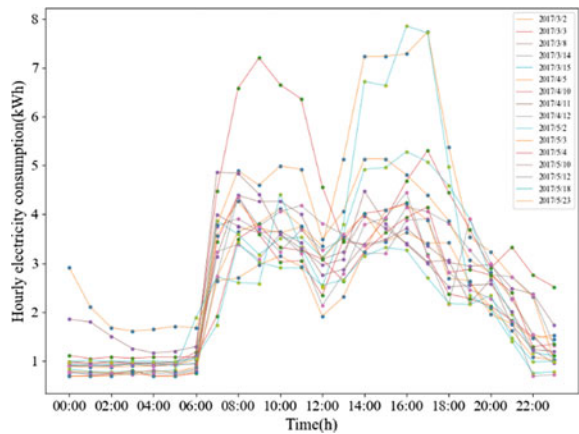
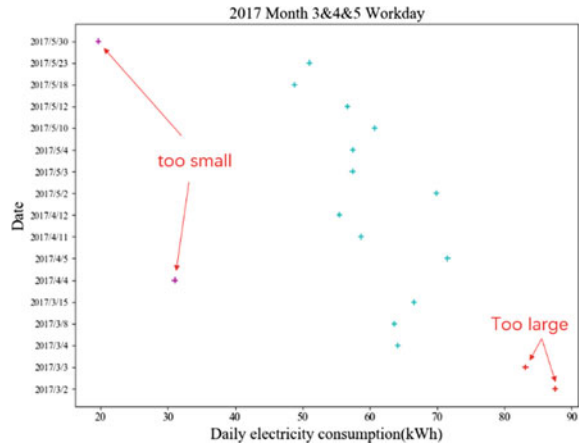


Fig. 17 The clustering diagram



5.2 Eliminate Abnormal Data by Clustering Method

Some abnormal data are eliminated by k-means clustering, and the clustering result of daily electricity consumption is shown in Fig. 17. The anomaly data identified by the clustering method is shown below, the electricity consumption of March 2nd and March 3rd are too large.

Fig. 18 The implicit error data

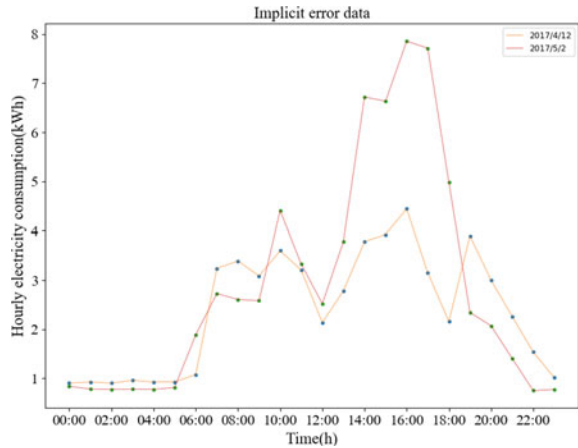
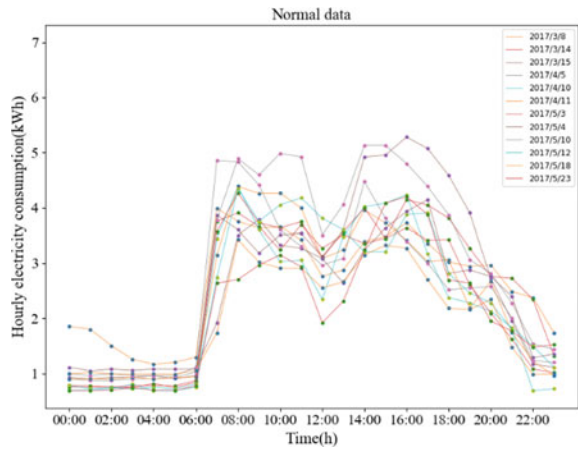


Fig. 19 Normal energy consumption



5.3 Compare with Historic Energy Consumption Characteristic Lines

Select the hourly historical electricity consumption of the lighting and socket in March, April, and May of 2015 and 2016 as historical data to find out the electricity consumption characteristics of weekdays. The invisible error data is further identified by the identification method based on the energy consumption characteristics. The identification result is shown in Fig. 18. After identifying, the normal energy consumption is shown Fig. 19. The electricity of April 12th and May 2nd is abnormal, the reasons need to be ascertained.

By using the identification method based on energy consumption characteristics, we identify four abnormal electricity consumption lines. Besides two lines which are identified by clustering method, we find two more implicit error electricity

consumption lines by comparing with historical data. This new method of identifying implicit error data is more comprehensive and accurate.

6 Conclusions

It can be seen from the case study that the new identification method based on energy consumption characteristics identifies abnormal electricity consumption data correctly and accurately. Instead of identifying individual data, this method links hourly electricity consumption data into lines and try to find the characteristic of lines to identify abnormal lines. It can be commonly used in energy consumption monitoring platform. After identification, the correct energy consumption data will play an important role in energy conversation and audit, etc.

Acknowledgements The study is supported by the National Key R&D Program of China (Grant No. 2017YFC0704200).

References

1. Yang, S., Luo, X., Ming, D.: Data processing method for public building energy consumption supervision platform based on data mining. *J. HV&AC* **2**, 82–86 (2015). (in Chinese)
2. Li, Z., Li, H., Yu, S.: Equivalent energy consumption of building energy efficiency. *J. Tongji Univ. Nat. Sci.* **38**(3), 353–357 (2010). (in Chinese)
3. Lee, W.S., Lee, K.P.: Benchmarking the performance of building energy management using data envelopment analysis. *Appl. Therm. Eng.* **29**(16), 3269–3273 (2009)
4. Seem, J.E.: Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy Build.* **39**(1), 52–58 (2007)
5. Zhou, Z.: *Machine Learning*, 1st edn. Tsinghua University Press, Beijing (2016). (in Chinese)
6. Xue, A., Yao, L., Shiguang, J.: Survey of outlier mining. *Comput. Sci.* **35**(11), 13–18 (2008). (in Chinese)
7. Karypis, G., Han, E.H., Kumar, V.: Chameleon: hierarchical clustering using dynamic modeling. *Computer* **32**(8), 68–75 (2002)
8. Kotsiantis S.B.: Decision trees: a recent overview. *Artif. Intel. Rev.* 1–23 (2011)
9. Kashef, R., Kamel, M.S., Kashef, R.: Enhanced bisecting k-means clustering using intermediate cooperation. *Pattern Recogn.* **42**(11), 2557–2569 (2009)

Part II

Construction Robot and Automation

An Improved Weight Control System for Slender Cigarette Production



Zhonghua Han, Xu Yang, Kaiyuan Bi, Xiaoting Dong and Xixian Sun

Abstract In this paper, a weight control system is developed based on Beckhoff PLC for a ZJ19 cigarette machine, which improved the performance of the existing system, and the control system is also introduced. The system uses the microwave to measure cigarette weight, and uses the digital PID control algorithm to achieve weight control. The system hardware structure is simple, the performance of the system is stable and reliable with quick response set aside the upgrade space. The original machine control system was completely replaced with the designed system to meet the requirements of slender cigarette production precision and the production efficiency and product quality were ensured.

Keywords Slender cigarette · Microwave measure · Weight control · PID

Z. Han (✉) · X. Yang · K. Bi

Faculty of Information and Control Engineering, Shenyang Jianzhu University, Shenyang, China
e-mail: hanzhonghua@sia.cn

X. Yang

e-mail: yangxu@sia.cn

K. Bi

e-mail: 176922905@qq.com

Z. Han

Department of Digital Factory, Shenyang Institute of Automation, CAS, Shenyang, China

X. Yang

Intelligent Detection and Equipment Department, Shenyang Institute of Automation, CAS, Shenyang, China

X. Dong

Department of Electric Engineering, Sichuan College of Architectural Technology, Deyang, China

X. Sun

Shenyang Jingdaxi Machine and Electrical Equipment Co., Ltd, Shenyang, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_7

1 Introduction

The weight of cigarettes is an important indicator of the quality of cigarettes [1]. As people's demand is increasing for the quality of cigarettes, the weight control system becomes more and more important; that strictly controls the average weight of cigarettes and to ensure the weight standard deviation of cigarettes can effectively control the amount of tobacco. This method is beneficial to the factory's production management, and reduces production costs. Therefore, tobacco weight control caused great concern for tobacco companies [2].

Compared with ordinary cigarettes, slender cigarette has less tobacco and less harmful substances such as nicotine and tar. It can not only meet the needs of smoking, but also reduce the harm to the human body. They are popular with many smokers. As the slender cigarette in China is still in its infancy stage of development. It is still in the initial stage of development in China. Tobacco companies consider production costs. Most enterprises transform the old model equipment to produce slender cigarette. The cigarette machine was adjusted greatly during the transformation, so there is a large change in the amount of tobacco during production. The slender cigarette has higher requirements than ordinary cigarettes for average weight and standard deviation. Therefore, it is necessary to develop an efficient and reliable weight control system to ensure the quality of cigarettes.

The weight of a single cigarette in a certain brand is 0.55 g. The weight of 20 cigarettes is 11 ± 0.15 g. The standard deviation is less than 0.014 g. So the accuracy of the request is very high. The new cigarette machine is improved from the old Cigarette machine. The old equipment's weight control system adopts the DSP circuit board, which is difficult to meet the requirement of cigarette weight and standard deviation. The new control system uses Beckhoff PLC as the main controller. Compared to the DSP control circuit, PLC works stable and is easy to expand. It is also easy to upgrade the system later [3]. This system adopts the Beckhoff high-speed oversampling module. It is competent for the measuring task during high-speed production. PLC emulate PID controller to realize the weight control of tobacco [4]. The system responds quickly and moves smoothly. It perfectly meets the requirements for the weight control of the cigarette machine.

2 Structure of Control System

The control system derives the weight of the cigarette by measuring the density of the tobacco. The control of the amount of tobacco is realized by adjusting the height of the leveling plate according to the weight of the cigarette. It adjusts the position of the tight head by adjusting the leveling plate. At the same time, excluding ultra-light cigarettes, overweight cigarettes, hollow cigarettes, caked cigarettes, and other substandard cigarettes. Make sure the cigarette average weight and standard deviation are controlled in the required range.

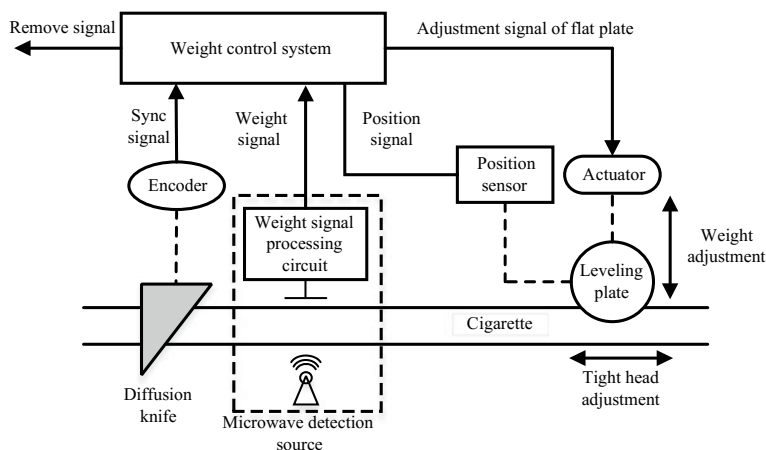


Fig. 1 The control system structure

In this paper, the weight control system consists of a weight measurement unit, control unit, execution unit, and position measurement unit and other components. The control system structure is shown in Fig. 1.

Weight measurement unit measures the density of tobacco by using SIA-MW10-12K microwave source. When the cigarette passes through the microwave cavity, the microwave passes through the cut tobacco and decays. There is a corresponding relationship between attenuation and tobacco density, according to this relationship can calculate the cigarette weight.

In order to improve the accuracy of cigarette weight measurement, it is necessary to ensure that there are enough sampling points on every cigarette. So we can calculate the total weight of cigarettes as well as the weight distribution of cigarettes. The control system uses encoders to acquire the synchronization signal. It can calculate the cigarette weight accurately.

The control unit contains PLC and IPC. PLC is used to complete data acquisition and output control. The host computer is used to provide a friendly man-machine interface and store PLC operating parameters.

(1) Beckhoff CX2020 is selected as the PLC module. It has excellent performance and is fast. It enables fast control below milliseconds (speed control technology). We use CX2100 power module. IO module is:

1. EL1262: The digital input module with oversampling is used to acquire the encoder's synchronization pulse. Its sampling speed is fast enough to meet the requirements of cigarette machine production.
2. EL3702: Analog input module with supersampling function used to obtain the density signal of the tobacco in the microwave detection source. And the position signal of the high sensor output of the flat plate. Then carry out AD conversion.

3. EL2008: The digital output module provides eight-point output signals to adjust the height and phase of the leveling plate and to control the culling valve.
- (2) Siemens industrial PC is selected for the upper computer. The interactive program is developed for data monitoring, parameter modification, and other auxiliary functions. The actuating unit is DC servo motor and valve. The position measurement unit is a resistive sensor.

3 Weight Control Principle

3.1 Weight Signal Acquisition

According to the signal difference between microwave transmitting end and receiving end, we get the voltage signal by the microwave detecting circuit. According to the output voltage of the microwave detection source, we can calculate the cigarette weight. The equation is

$$F = K \cdot \ln(V/V_0) \quad (1)$$

where F is the weight of the cigarette corresponding to the output voltage V , V is the output voltage when a cigarette is detected in the source, V_0 is the output voltage of the detection source when there is no cigarette, K is the proportionality coefficient [5].

The data will cause some interference in the process of collection and transmission, so it is necessary to preprocess the collected data. If the collected value is within the allowable range, it is processed as normal data. If the collected value is outside the normal distribution, the signal is disturbed. At this point, it should be limited to a reasonable range. The specific equation is as follows:

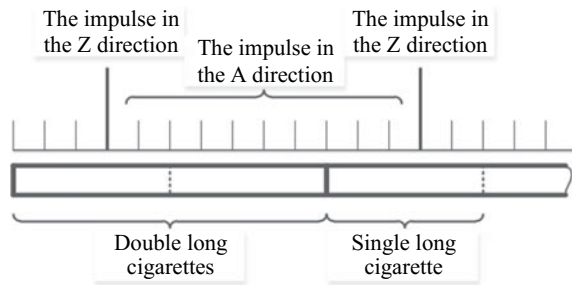
$$m = m_R + \frac{1}{2}(m_0 + d) \quad (2)$$

$$m = m_R + \frac{1}{2}(m_0 - d) \quad (3)$$

where m_R is the original data collected, m_0 is the weight average, d is the allowable range of fluctuation, m is the final valid data.

3.2 Calculation of Cigarette Weight

The above measurement value is the weight of a point on the cigarette. So we need to take a sufficient number of sampling points on a cigarette for measurement. Then the

Fig. 2 Diagram of the tobacco weight calculation

total weight and the weight distribution of the cigarettes were calculated. In order to improve the accuracy of the measurement value and take into account the workload of data processing, taking 256 points on each cigarette whose length is twice as long as the original cigarette. Take 128 cigarettes on a single cigarette.

The control system uses incremental encoders to obtain the synchronization signal. Encoder and cutter plate are connected together. The encoder rotates one revolution and the cutter cuts a double-length cigarette. We choose 256 line encoder. The control system collects A-direction pulses and Z-direction pulses. Phase A pulses are used to synchronize and count acquisition points and Z pulses are used to determine the position of the acquisition point on the cigarette. When receiving the Z pulse, the A-phase pulse value is cleared (Fig. 2).

PLC completes the collection and processing of the weight signal by the EL1262 module and the EL3702 module [6]. In a running cycle of PLC, the two modules to complete the same number of acquisition tasks. To keep each collection time consistent. First check EL1262 whether to receive the pulse jump in a cycle. If a jump is detected, record the time of this acquisition task. Find the weight signal collected by EL3702 and use it as the weight of this collection [7].

We obtained the A-direction pulse value of the encoder on the end of the cigarette. These data points can be corresponding to the specific location of the cigarette. We calculate the total weight of cigarettes and the weight distribution of cigarettes. We can determine the position of the tight head in this way. We can also use this method to determine whether the cigarette is caked or hollow.

3.3 Weight Control Algorithm

The weight of cigarette depends mainly on the amount of tobacco that is involved in the cigarette and the deviation of the tight head position. So the weight control system should ensure that the amount of tobacco in the cigarette is stable. It can be realized by adjusting the height of the leveling plate. At the same time, the position deviation of the tighthead should be controlled effectively. It can be achieved by adjusting the phase of the leveling plate.

Then the basic principle of the weight control system will be introduced. The weight of the cigarette is controlled by adjusting the height of the leveling plate according to the weight of the cigarette in real time. This is a typical hysteresis control system. The control system uses PID control algorithm.

According to the characteristics of the system, it uses a PLC to simulate a PID controller. So the PID needs to be discretized. Digital PID can be divided into two types; position and incremental. The cigarette maker adjusts the height of the leveling disc by calculating the amount of cigarette weight deviation. Incremental PID control is used as follows [8]:

$$\begin{aligned} \Delta u_k = & K_p \left(1 + \frac{T}{T_i} + \frac{T_d}{T} \right) e_k \\ & - K_p \left(1 + \frac{2T_d}{T} \right) e_{k-1} + K_p \frac{T_d}{T} e_{k-2} \end{aligned} \quad (4)$$

Set $A = K_p \left(1 + \frac{T}{T_i} + \frac{T_d}{T} \right)$, $B = K_p \left(1 + \frac{2T_d}{T} \right)$, $C = K_p \frac{T_d}{T}$, (4) can be rewritten as:

$$\Delta u_k = A e_k - B e_{k-1} + C e_{k-2} \quad (5)$$

where e_k is the deviation obtained from the k sampling, e_{k-1} is the deviation obtained from $k - 1$ sampling, e_{k-2} is the deviation obtained from $k - 2$ sampling.

The sampling period T is a constant. As long as the deviation values of A , B and C are determined, the control quantity could be obtained. There is a small amount of computation that is suitable for this system.

In order to improve control precision, it is necessary to filter the collected weight data. According to the characteristics of this system, the method of sliding average filtering is selected for data processing. The basic principles are described below. First, continuously collect the weight of 18 cigarettes (control particle size) and save. The maximum and minimum values are removed and the average weight of the remaining 16 cigarettes is calculated. Then calculate the deviation. The weight of nine cigarettes (the number of sliding samples) is measured continuously. Remove the first nine data from the weight of the last 18 cigarettes and reconstitute 18 data with 9 new data. Then we calculate the average and calculate the new deviation. Repeat this process to provide feedback data for PID control.

Adjust the sensitivity and speed of control by adjusting the size and number of sliding samples. Improve the reliability and stability of the system.

To avoid system instability caused by adjusting the height of the leveling plate frequently. If the adjustment is within the tolerance, the system will not adjust. At the same time, in order to judge the leveling height in real time, the PLC reads the level sensor signal, preventing the plate from moving when it is in the extreme position, and to avoid motor overload and protect the motor. When the weight of cigarettes is too bigger for a long time, level plate reaches the limit position and cannot be adjusted. The control system will give an alarm [9].

The weight of a single cigarette will be affected by a tight head position deviation. If the head position deviation exceeds the allowable range, the weight of the two cigarettes will be disqualified. It will affect the standard deviation of cigarettes. As the leveling disk and the cutter disk rotate synchronously, the tight head position does not need to be adjusted frequently. If the head is shifted in the same direction several times in a row, it will cause its offset to exceed the allowable range. Then its position needs to be adjusted. The corresponding relationship between the output of PLC and the phase of the plate is found out. Then adjust it quantitatively.

As the control system adopts the hysteresis control method, it cannot completely eliminate the quality of substandard cigarettes. Therefore, it is necessary to calculate the weight of each cigarette in real time and determine whether it is within the allowable weight range. Then remove unqualified cigarettes, which can effectively control the average weight and standard deviation. At the same time, according to the distribution of the weight of cigarettes, we can determine whether the cigarettes are loose, caking, and other issues. And the quality of cigarettes is improved by rejecting substandard products.

4 Experiment and Result Analysis

In order to verify the comprehensive performance of the weight control system, the system is integrated into the transformed cigarette machine unit for installation and debugging. After a period of operation, the effect is good. Sampling at the site and measuring various indicators. The indicators can meet the design requirements.

According to the factory's measurement requirements, we measure 30 cigarettes at a time. We get the average weight, standard deviation, and coefficient of variation of 30 cigarettes by measuring. Ten sets of data needs to be measured. Measuring instrument is OM-VL-type comprehensive test station provided by Beijing America Technology Co., Ltd. The weight is set to 550 mg. The allowed range of the average weight is 550 ± 0.075 mg. The allowable range of the average weight is 550 ± 0.075 mg. The specific data is shown in Table 1.

It can be seen from the data in the table that the average is better than the design requirements. The standard deviation is less than 0.014 g, which meets the requirements. And the coefficient is small. This shows that the system is stable and reliable and can achieve high control accuracy. It can meet the needs of normal production perfectly.

5 Conclusions

In this paper, we design a weight control system for slender cigarette production based on Beckhoff PLC. The new control system can completely replace the original cigarette machine control system. The system uses a common PLC module. The

Table 1 Tobacco weight measurement data

No.	Average weight	Standard deviation	Coefficient
1	0.554	0.009	1.709
2	0.551	0.011	1.946
3	0.546	0.011	2.041
4	0.553	0.009	1.571
5	0.550	0.013	2.314
6	0.554	0.011	1.899
7	0.552	0.008	1.508
8	0.555	0.012	2.195
9	0.549	0.011	1.957
10	0.553	0.013	2.279

new control system has the advantages of simple structure, short development cycle, good stability and can be upgraded and expanded. Meanwhile, the maintenance cost of the equipment is reduced, and the data acquisition and control precision are improved. The control algorithm adopts PID algorithm, and the accuracy and real time are better. It meets the requirement of weight control in the production process of slender cigarette. The system is of great significance to improving the production efficiency and reliability of the cigarette machine, as well as improving the product quality and controlling the cost.

Acknowledgements The work was supported by the Natural Science Foundation of Liaoning Province of China (201602608) and National High-tech R&D Program of China (2015AA043901).

References

1. Wei, H.: Application of microwave detecting technology in cigarette weight controlling system. *Tobacco Sci. Technol.* **3**, 15–18 (2006). (in Chinese)
2. Gongming, W.: Design of cigarette weight control system based on microwave detecting technology. *Food Mach.* **31**(06), 113–116 (2015). (in Chinese)
3. Milik, A., Hryniewicz, E.: Synthesis and implementation of reconfigurable PLC on FPGA platform. *Int. J. Electron. Telecommun.* **58**(1), 313 (2012)
4. Kocian, J., Koziorek, J., Pokorny, M.: An approach to PLC-based fuzzy expert system in PID control. *ICIEIS 2011 Inf. Eng. Inf. Sci.* **225**, 195–205 (2011)
5. Rongxiang, L.: The realization of PID weight control system of cigarette machine. *J. Baoji Univ. Sci. Technol. (Natural Science Edition)* **25**(03), 221–223 (2005). (in Chinese)
6. Rui, G., Xiaoe, W., Xu, Y.: Application of the extreme fast control technology in weight control and quality inspection system for cigarettes. *Process Auto. Instrument.* **37**(02), 1–32 (2016). (in Chinese)
7. Yunkai, C., Yiqin, X., Jinsong, D.: Design of the cigarette weight control system based on the super sampling technology of bechhoff. *Control Eng. China* **22**(05), 831–834 (2015). (in Chinese)

8. Yun, X., Deng, Z., Wang, X., Yang, C.: Design of incremental PID controller based on FPGA. *Appl. Mech. Mater.* **10**(4028), 830–833 (2014)
9. Kopček, M.: Embedded PLC webserver and possibilities of its utilization. *Res. Papers Faculty Mater. Sci. Technol. Slovak Univ. Technol.* **24**(39), 812 (2016)

Design of Building Environment Mobile Monitoring and Safety Early Warning Robot



Guoqing Yang, Yuhao Wang and Bing Chen

Abstract In view of the characteristics of two-wheeled robots, such as the small size and flexible movement, and the increasing demand for environment and security, this design implements a building environment mobile monitoring and security early warning robot. On the basis of the self-balancing of two wheels, the robot can monitor the building environment and alert people with voice intelligently. The system combines automatic control technology, PWM DC motor control technology and sensor technology. It applies the STM32 as the controller and uses the PID control algorithm to meet the requirements of its characteristics, such as multiple variables, nonlinearity, strong coupling, parameter uncertainty, and others. The wireless sensing and control are also established. For this, people can use the mobile phone or PC terminal to realize the wireless mobile monitoring, early warning, and alarm of building environment including temperature, humidity, illumination, air pressure, altitude, air quality, fire source detection, etc., based on mobile machines.

Keywords Self-balance · PID algorithm · Environmental monitoring · Robot

1 Introduction

For the higher requirements of the environment and the quality of life, mobile robots are facing a more complex environment and tasks. Using robots to replace human beings requires the robot to use only two wheels to realize the self-balance of movement. The two-wheeled robot is a highly unstable, multivariable, nonlinear, and strongly coupled system and it needs to control the motor to maintain an attitude balance. The self-balancing of the two-wheel robot is realized by the coordination of motor-drive wheels, sensors, microprocessor, and the robot's own mechanical device. The robot is a complex nonlinear system with various functions such as environmental perception, dynamic decision-making and planning, behavior control, and

G. Yang (✉) · Y. Wang · B. Chen

School of Control and Mechanical Engineering, Tianjin Chengjian University, Tianjin, China
e-mail: tjcj2008@126.com

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_8

execution. At the same time, it can realize the wireless mobile monitoring, timely security early warning and alarm of building environment to solve the environment security monitoring and early warning problem of the region, where humans cannot reach or 24 h mobile requirements.

2 Balance Principle Design of Two-Wheeled Robot

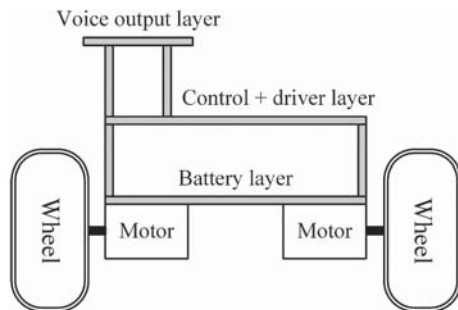
2.1 Mechanical Structure Design of Two-Wheeled Monitoring Robot

The structure of the two-wheeled monitoring robot is designed mainly in three layers of architecture [1, 2], including the voice output layer, control and driver layer, and battery layer from top to bottom as shown in Fig. 1. The voice output layer is composed of a voice module and a speaker for early warning and environmental parameters broadcast; the control and driver layer integrates the controller, sensors, display, and drivers; and the battery layer is used to supply the 11.1 V lithium battery for the entire system.

2.2 Two-Wheeled Robot Motion and Balance Analysis

The intuitive experience of controlling the balance of two-wheeled robots comes from people’s daily life experience. Two conditions are required to keep a stick upright on the tip of the finger: one is to move the palm, and the other is the angle and trend (angular velocity) of the stick’s inclination [3, 4]. The movement of the palm can remove the angle and trend of inclination to keep the stick upright. These two conditions are indispensable. In fact, they are negative feedback mechanisms in control.

Fig. 1 Robot mechanical structure



Because the two-wheeled robot has two wheels, and the robot will only tilt in the direction of the wheel rolling, its balance control is realized by a similar negative feedback mechanism [5, 6]. That is, controlling wheel rotation and counteracting the trend of inclination on one dimension can keep the robot in balance. When the bottom wheels of two-wheeled robot are controlled to accelerate, the stress of the inverted pendulum on the two-wheeled robot is analyzed (the non-inertial system, with the wheel as the origin of the coordinate) [7]. The inverted pendulum will be subjected to the additional inertia force, which is proportional to the wheel's acceleration in the opposite direction [8, 9]. Thus, the resilience of the inverted pendulum is as follows:

$$F = mg \sin \theta - ma \cos \theta \approx mg\theta - mk_1\theta \quad (1)$$

where m is the mass of the two-wheeled robot; g is the gravitational acceleration; a is the acceleration of wheels; θ is the angle of inclination; k_1 is the coefficient of proportionality.

Because θ is very small, the equation can be linearized. Suppose the wheel acceleration a is proportional to the deflection angle θ in the negative feedback control and the ratio is k_1 . If $k_1 > g$, the direction of restoring force is the opposite direction of displacement.

In addition, in order to stabilize the two-wheeled robot as fast as possible in the vertical position, the damping force is also needed. Because damping forces such as air resistance and friction are relatively small, the additional control damping force is needed. The increased damping force is directly proportional to the speed of the inclination in the opposite direction. Therefore, equation constraint (1) becomes

$$F = mg\theta - mk_1\theta - mk_2\theta' \quad (2)$$

where k_2 is the coefficient of proportionality and θ' is the angular speed.

According to the above control method, the inverted pendulum model can be changed into a pendulum model, which can be stabilized in a vertical position. So the algorithm to control the acceleration of two-wheeled robots is proposed.

$$a = k_1\theta + k_2\theta' \quad (3)$$

The sum of these two terms is used as the control quantity of wheel acceleration. As long as the conditions $k_1 > g$, $k_2 > 0$ are guaranteed, the two-wheeled robot can be maintained as upright as a pendulum. There are two control parameters which determine whether the robot can be stable to the vertical equilibrium position. The wheel acceleration must be greater than the acceleration of gravity. The parameters determine the damping coefficient of the robot back to the vertical position, and the proper damping coefficient can be selected to ensure the robot to be stable in the vertical position as soon as possible.

From the motion state analysis and balance principle of two-wheeled robot, it is known that the acceleration control algorithm of two-wheeled robot can be abstracted

as a PD controller (proportional derivative control). Through the IIC communication protocol, the algorithm reads the angle θ (Angle-Balance) from the MPU6050-DMP and the angular velocity θ' (Gyro-Balance) from the MPU6050-gyroscope. Then the two values are multiplied, respectively, by the proportional coefficient P (Angle) and the derivative coefficient D (Angle), and the results are added together to output a PWM signal to control the speed of the motor.

The action of the upright ring is to speed up the system by adjusting the PWM speed of the motor so as to keep the system upright. However, the acceleration is generated by the speed, so it is inevitable that the speed of the system has reached its limit, but the angular velocity of this direction is still needed to increase to keep the balance. At this time, the system will fall due to the failure to provide acceleration or to provide enough acceleration.

2.3 Walking Upright Control Algorithm

The upright ring can make the robot hold its balance only for a short time, but not for a long time and it even can not control the speed and the direction. Therefore, the speed loop and the steering ring need to be introduced to make the robot keep the upright state for a long time and control the speed and direction [3].

In the speed controller, the PI control is the classical control algorithm. The integral helps to eliminate the residual error with the target speed, which is of great significance for keeping the system upright and speed regulating. In order to ensure the stable operation of the system, the rate of speed change is reduced to avoid interference of upright control and the program is used to realize the function of first-order low-pass RC filter to process the data.

The first-order low-pass filtering algorithm is as follows:

$$Y(n) = \alpha X(n) + (1 - \alpha)Y(n - 1) \quad (4)$$

where α is the filter coefficient; $X(n)$ is the current sampling value; $Y(n - 1)$ is the previous filter output; $Y(n)$ is the current filter output.

In the first-order low-pass filtering algorithm, the current sampling value is weighted by the previous filter output, and the effective filtering value is obtained, so that the output has a feedback effect on the input. After the first-order low-pass filtering algorithm, the PI control is carried out. In the P control, the proportional coefficient P is multiplied by the filtered value. In the I control, the integrated data are added up which magnitudes are limited, and the results are multiplied by the integral coefficient I.

In the control of the steering ring, it is divided into two states. One is upright and straight. In this state, the proportional control is closed and the pure derivative control is adopted. The goal is to use Z-axis gyroscope and to keep the system to walk in a straight line as far as possible and to restrain the drift. The other is the steering state.

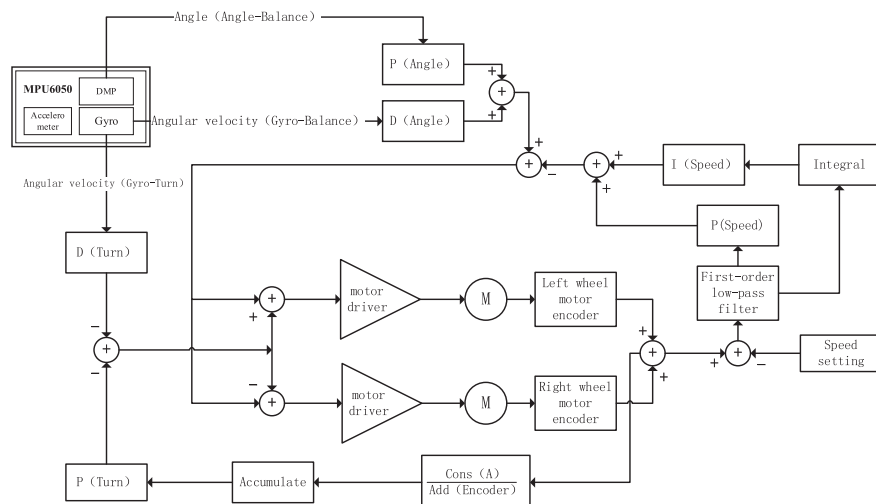


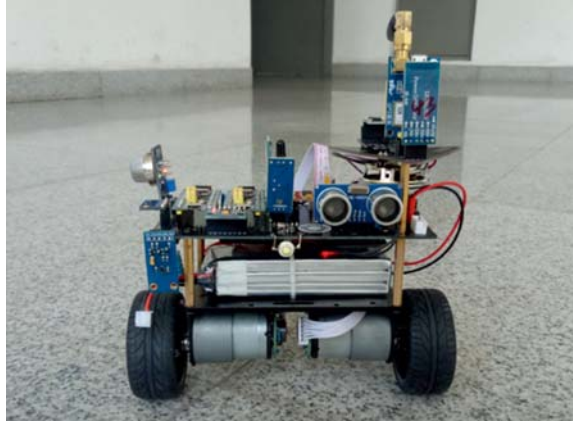
Fig. 2 Upright walking control algorithm

In the steering state, the derivative control is closed and the pure proportional control is used. According to the current speed of system, the inverse proportion operation is carried out. The result is multiplied by the coefficient Cons (A) and the speed step value is obtained. Its accumulation is multiplied by the proportional coefficient P. The PD control of upright ring, the PI control of speed ring and the PD control of steering ring is integrated to calculate the final PWM signal and then the signal is sent to the motor drive chip TB6612FNG to control the motor speed and motion state, as shown in Fig. 2.

3 System Architecture Design

This robot system mainly consists of two parts: motion control and environmental monitoring. The motion control uses a STM32 (Cortex-M3) as the main control chip, equipped with a 3S lithium battery and a DC geared motor with an encoder; it uses the 6-axis attitude sensor MPU6050 to measure the attitude and the TB6612FNG as the drive circuit. And it carries the ultrasonic distance measurement system, the GPS positioning system, the RFID radio-frequency identification system, and the Bluetooth wireless transmission system. The environmental monitoring applies another STM32 (Cortex-M3) as its controller, and the transducers, such as temperature and humidity sensor DHT11, barometer MS5611, illuminometer BH1750, air quality sensor MQ135, fire source detector, are used to monitor environmental parameters. The light-operated lighting system, broadcast system, OLED display system are also designed in this part. The computer or mobile terminal can display the oper-

Fig. 3 Building environment mobile monitoring and safety early warning robot



ating parameters, environment parameters, map, and the robot's GPS location. The building environment mobile monitoring and safety early warning robot is shown in Fig. 3.

4 System Software Program Design

4.1 Motion Control Program Design

After power-on reset, the robot program defines the global variables in the control chip at first, and then initializes Delay, USART, NVIC, LED, Key, PWM output, Encoder, IIC protocol, MPU6050, MPU6050-DMP, OLED, IWDG and set their working modes. After monitoring the legitimate ID, MPU6050 initializes 5 ms timer interruption, USART3 used for ultrasonic and Bluetooth communication, USART2 used for dual-CPU asynchronous serial communication.

After this initialization, the program enters the system cycle, constantly deals with the data in the dual-CPU communication, sends data to the host computer, updates the OLED data, and feeds the watchdog. MPU6050 performs the initialization of 5 ms timer interruption. The system is interrupted and enters the interrupt handler. In the interrupt handler, the pitch angle, the encoder, the distance of the ultrasonic distance, and the state of the key are obtained and the PWM control signals of the left and right motors are calculated by the PID controller and are sent out.

Using 5 ms external interruption of MPU6050, the program reads a movement attitude from DMP and gyroscope through the IIC communication protocol per 5 ms, including DMP data on Y-axis and gyroscope data on Y- and Z-axis. In addition, the program reads the encoder data, the key scan, and the ultrasonic distance measurement. The PWM signal is calculated by the PID controller and is sent to the motor driver. The control process flow is shown in Fig. 4.

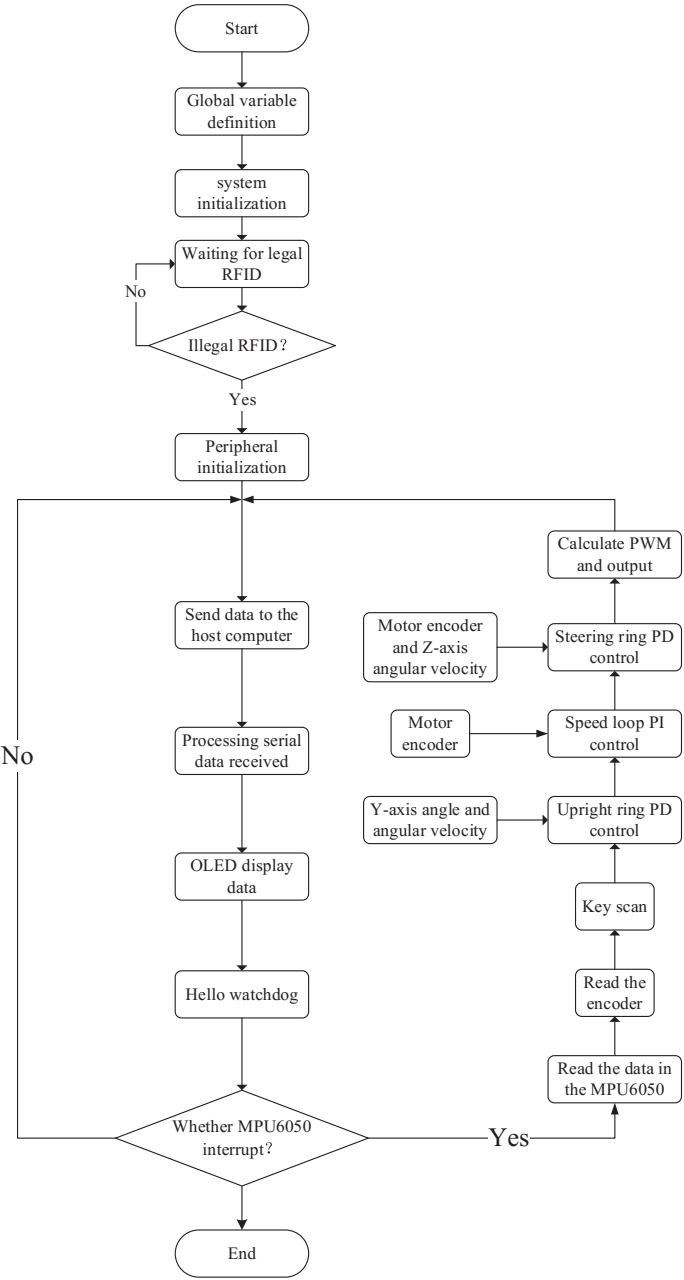


Fig. 4 Motion control microcontroller running process

4.2 Program Design of Environmental Monitoring Software

After power-on reset, the global variables of the hardware are defined at first. Delay, USART2 for dual-CPU communication, USART1 for the voice control output, NVIC, LED, Key, IIC protocol, ADC, and OLED are initialized in turn. Then the peripherals, temperature and humidity sensor, illuminometer, barometer MS5611, air quality sensor, and fire source detector are initialized. Finally, the timer and the watchdog are initialized. After initialization, the program enters the system cycle, continuously processes data in dual-CPU communication, updates OLED display data, and feeds the door dog. The environmental monitoring is strictly controlled by timer 1 with the timing of 500 ms. In the timer interrupt function, temperature and humidity, illumination, fire condition, atmospheric pressure, altitude are obtained and data processing and format control output are accomplished.

The timer is used for timing acquisition and the interrupt signal is produced per 500 ms. The data is collected by the communication protocol which is corresponding to each sensor, and then data processing and formatting output are carried out. The barometer MS5611 communicates with MCU through the IIC communication protocol. By comparing with the standard atmospheric pressure, the altitude of the location is obtained. The temperature and humidity sensor communicates with MCU through a single bus and five-byte data are read out. The temperature and humidity are output after the data are checked without error. The air quality sensor MQ-135 communicates with MCU through the ADC mode. The sensor is more sensitive to common air quality indicators (smoke, benzene, aldehydes and etc.) and it defines six kinds of air quality grades: A+, A, B, C, D, E. Compared with the values acquired by ADC, the current air quality grade is obtained. Since the interface standard voltage of STM32 is 3.3 V and some pins can accept 5 V, the voltage of 3S lithium battery is collected by means of resistance divider. The illuminance sensor communicates with the microcontroller through IIC communication protocol, and the illuminance is directly read out by MCU. The fire source detector uses direct numerical output. When the flame is monitored, the signal level turns over, and the microcontroller calls the user. Finally, hardware and software are tested together. The result shows that building environment mobile monitoring and early warning robot operation are normal and stable.

5 Conclusions

The building environment mobile monitoring and safety early warning robot has the characteristics of small size and flexible movement, which is suitable to work in a narrow and dangerous environment. The key of its design is the attitude control strategy to realize self-balancing, motion in every direction, rotation, speed regulation, stop of the robot. The design focuses on the integration of two-wheeled balance system and environmental monitoring to realize the control of the robot and param-

eters monitoring in mobile phone or PC terminal. Applying DMP of MPU6050, the design obtains the optimal approximate solution of the robot tilt angle, which is convenient for controller to process. The upright ring PD controller, the speed loop PI controller, and the steering ring PD controller are designed in this design. Through the experiment, the algorithm basically satisfies the design requirements and the self-balancing control of the two-wheeled robot is realized. With the rapid development of the economy and the popularization of mobile terminals, higher requirements for the environment and quality of life are required. Environmental monitoring and safety early warning robot not only can monitor environmental information and alarm early warning, but also can display environment data and warning information on mobile terminals. Furthermore, it still can control the robot wirelessly. In addition, it is of great theoretical and practical significance to use robots to replace people for reducing personnel damage, getting environmental parameters and preventing danger in a dangerous environment.

References

1. Chen, Y., Luo, G., Mei, Y.: UAV path planning using artificial potential field method updated by optimal control theory. *Int. J. Syst. Sci.* **47**(6), 1407–1420 (2016)
2. Peng, T., Qian, J., Zi, B.: Mechanical design and control system of an omni-directional mobile robot for material conveying. *Procedia CIRP* **56**, 412–415 (2016)
3. Jose, K., Pratihari, D.K.: Task allocation and collision-free path planning of centralized multi-robots system for industrial plant inspection using heuristic methods. *Robot. Auton. Syst.* **80**, 34–42 (2016)
4. Drenner, A., Burt, I., Chapeau, B.: Multi-robot systems: from swarms to intelligent automata. In: *Proceedings from the 2002 NRL Workshop on Multi-Robot Systems*, p. 141. Springer Science & Business Media (2013)
5. Digani, V., Sabattini, L., Secchi, C.: Towards decentralized coordination of multi robot systems in industrial environments: a hierarchical traffic control strategy. In: *2013 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 209–215. IEEE (2013)
6. Hu, C., Hu, C., He, D.: A new ROS-based hybrid architecture for heterogeneous multi-robot systems. In: *2015 27th Chinese Control and Decision Conference (CCDC)*, pp. 4721–4726. IEEE (2015)
7. Cai, D., Wu, S.: Distributed multi-robot dynamic network topology control while preserving connectivity. In: *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 2, pp. 51–55. IEEE (2016)
8. Yan, Z., Jouandeau, N., Cherif, A.A.: A survey and analysis of multi-robot coordination. *Int. J. Adv. Robot. Syst.* **10**(12), 399 (2013)
9. Hess, W., Kohler, D., Rapp, H.: Real-time loop closure in 2D LIDARSLAM. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1271–1278. IEEE (2016)

Application of Probabilistic Reasoning Algorithm in Indoor Positioning Based on WLAN



Meng Li and Honglin Wang

Abstract To improve the efficiency of resource utilization of large public buildings, indoor positioning and navigation is more and more possible in recent years. But the navigation accuracy is poor because of the complex environment in the buildings and the failure to use the satellite navigation signal. To solve this problem, a navigation and integrated information system is designed and developed by using the existed wireless network signals. A position fingerprint location method is adopted in order to ensure the finite complexity of the algorithm. A weighted K-nearest neighbor algorithm based on probabilistic reasoning is proposed to meet the requirements of navigation precision. This system will be widely used in large public buildings such as market, museum, library, and transport hubs.

Keywords Indoor positioning · Wireless network · Probabilistic reasoning

1 Introduction

With the development of architecture technology and the improvement of people's living standards, the scale of public buildings is expanding and the functions are becoming more and more complex. Therefore, it is necessary to navigate in large public buildings such as traffic hubs, shopping malls, libraries, cultural centers, and so on. However, due to the limitation of indoor environment, GPS signal is unusable. With the extensive deployment of wireless local area networks (WLAN) in the indoor environment, the maturity of the IEEE 802.11 series protocols, and the popularization and application of mobile intelligent devices such as mobile phones, the indoor positioning technology based on WLAN has developed rapidly. In recent

M. Li (✉)

Control and Mechanical Engineering College, Tianjin Chengjian University, Tianjin 300384, China

e-mail: limeng@tcu.edu.cn

H. Wang

Power China Construction Group Ltd, Beijing 100011, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_9

years, a lot of domestic and foreign universities and scientific research institutes have been carried out the research of indoor positioning technology. ORL laboratory of Cambridge University developed the Active Badge System, using infrared signals to communicate a person’s location [1]. Other systems such as RADAR positioning system of Microsoft company in USA, Cricket System of MIT [2], and Calamari system of University of California [3] have promoted the development of indoor positioning technology. In [4], a navigation and integrated information system is designed and developed by using the existed wireless network signals, based on localization technology using wireless networks. The use of the existing WLAN equipment without adding additional hardware and equipment has the advantages of low deployment cost and easy realization, which has become the mainstream of the research of indoor positioning technology at present.

In large public buildings, due to the complex structure of the building, electrical equipment, and many other characteristics, it will cause some interference to the WLAN signal. The traditional algorithm based on signal propagation model is not robust enough to meet the requirement of location accuracy. To solve this problem, a positioning algorithm based on probabilistic reasoning for indoor wireless networks is proposed.

2 Indoor Navigation System

The indoor navigation system includes navigation and information server, wireless network access points (AP), and intelligent mobile devices. The system structure is shown in Fig. 1.

The mobile device receives signal from the wireless AP, estimates the current position by positioning algorithm based on the intensity of WIFI signal, and plans the path according to the map. A navigation and information server developed on an embedded Linux system based on ARM9 is used. The hardware structure is shown in Fig. 2. The server includes a map information database to collect, analyze, manage,

Fig. 1 System structure

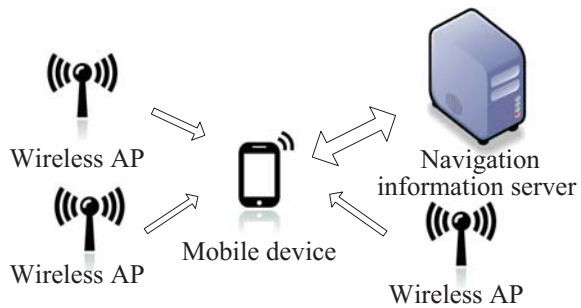


Fig. 2 Hardware structure of the navigation and information server



and release information. It also receives the position of the intelligent device, receives the requirements, and analyzes and stores the information from the device by sending the broadcasting information.

3 Positioning Method Based on Fingerprint

The WLAN signal always gets interference in large public buildings because of the complex structure and the massive electrical devices. Different methods of positioning have been used, and the character of each method is shown in Table 1. The location fingerprint method has low cost, does not need extra external hardware, and has a high positioning accuracy, which is suitable for public buildings. Although in the experiment of Ref. [5], it is shown that the fingerprint method relies on the structure of the building, the structure of a public building is relatively stable, and the parameters can be revised in debugging, so that the fingerprint method based on the signal intensity of WLAN can be used in this system.

The fingerprint method is performed in two steps. First, the location fingerprint database is established by measuring the intensity vector of WLAN signal on different points. Second, the position of the point to be measured is got by matching the signal intensity and the data in the database [6].

Table 1 Comparison of methods for indoor positioning technology based on WLAN

Method	Cost	Hardware required	Location of AP	Robustness	Positioning accuracy
AOA	High	Required	Required	High	High
TOA	High	Required	Required	High	High
TDOA	High	Required	Required	High	High
Signal propagate model	Low	Non-required	Required	High	Low
Nearest AP	Low	Non-required	Required	Low	Low
Location fingerprint	Low	Non-required	Non-required	High	High

4 Probabilistic Reasoning Algorithm

The number of wireless hotspot AP is L , named $AP_l (l = 1, 2, \dots, L)$. The signal intensity of the l th AP at the n th fingerprint is noted as s_{ln} . The signal intensity vector at the n th fingerprint is defined as $\mathbf{s}_n = [s_{1n}, s_{2n}, \dots, s_{Ln}]$. Similarly, the signal intensity vector at the point to be measured is defined as $\mathbf{s} = [s_1, s_2, \dots, s_L]$, where s_l is the signal intensity of the l th AP at this point.

The Pearson correlation function is selected to calculate the correlation between the sectors $\mathbf{s}_n$ and $\mathbf{s}$. The Pearson correlation function R is defined as

$$R(\mathbf{s}, \mathbf{s}_n) = \frac{\sum_{l=1}^L (s_l - \bar{s})(s_{nl} - \bar{s}_n)}{\sqrt{\sum_{l=1}^L (s_l - \bar{s})^2} \sqrt{\sum_{l=1}^L (s_{nl} - \bar{s}_n)^2}} \quad (1)$$

where $R \in [-1, 1]$. When $R = 1$, the sectors are positive correlation. When $R = -1$, the sectors are negative correlation. When $R = 0$, the sectors are uncorrelated.

Hypothesis H_0 is defined as the correlation coefficient between the signal intensity vector at the point to be measured and the vector at the fingerprint is 0.

Here, a hypothesis of independence and normality is made:

For each fingerprint point FP_n , the likelihood probability $P(\mathbf{s}|FP_n)$ can be expressed as

$$P(RSS_{AP1} = s_1, RSS_{AP2} = s_2, \dots, RSS_{APL} = s_L | FP_n) \quad (2)$$

where the general term $P(RSS_{APl} = s_l | FP_n)$ is the marginal probability that the signal intensity vector of AP_l equals to the signal intensity vector s_l of the fingerprint.

For each group of (AP_l, FP_n) , the values of s_l and s_n satisfy normal distribution.

The test statistic T is defined as

$$T = \frac{R\sqrt{L-2}}{\sqrt{1-R^2}} \quad (3)$$

According to the previous hypothesis, T obeys t distribution with $L-2$ degree of freedom.

The hypothesis test is made to N groups of vectors. Each group is composed of the signal intensity vector $\mathbf{s}$ measured online at the point to be measure and the signal intensity vector $\mathbf{s}_n$ at the n th fingerprint. The correlation coefficient R_n is obtained from Eq. (1), and the test statistics T is obtained from (2). The p -value $p\text{-value}^{(n)}$ is calculated and compared with α . If $p\text{-value}^{(n)} < \alpha$, the hypothesis H_0 is refused. Test every fingerprint and select refused ones. The number of fingerprints selected is k . Then the position estimation is

$$\hat{p} = \frac{\sum_{n=1}^k (p\text{-value}^{(n)})^{-1} p_n}{\sum_{n=1}^k (p\text{-value}^{(n)})^{-1}} \quad (4)$$

5 Experimental Results

Numerical experiment was carried out in a simulation environment. Four wireless routers are used as access points to provide WLAN signal. In order to analog mobile device usage environment, an application is designed based on Android to collect signal intensity data. Forty-five even-distributed points are selected as fingerprints.

In order to verify the effectiveness and superiority of the algorithm, two common algorithms (NN algorithm and KNN algorithm) are selected for comparative experiments. The results of the experiments are shown in Fig. 3. From the error distribution curve, the error of FIWKNN algorithm is obviously reduced. The probability of error within 2 m is increased from less than 50% of NN algorithm to more than 80%. It is proved that the FIWKNN algorithm is superior to the traditional NN and KNN algorithms in positioning accuracy.

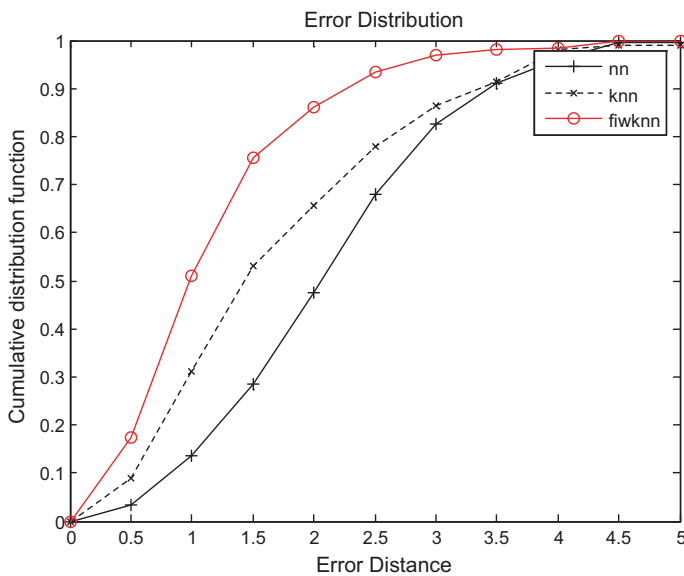


Fig. 3 Experimental results

6 Conclusions

In order to improve the accuracy of positioning based on WLAN signal intensity in large public buildings, various positioning methods are analyzed. Considering the performance and complexity, a fingerprint location method is selected as the location method. An algorithm based on probability reasoning is proposed to calculate the position by the coordination of fingerprints. By calculating the correlation between the pending points and the fingerprints, the T distribution hypothesis testing is used to determine the reliability parameters of each fingerprint point, and the location of the points to be measured is calculated. The effectiveness and superiority of the algorithm are verified by comparative experiments.

References

1. Want, R., Hopper, A., Falcao, V., et al.: The active badge location system. *ACM Trans. Inf. Syst.* **10**(1), 91–102 (1992)
2. Shen, M., Yang, X., Jun, W.U.: A method of identifying beacons of cricket system based on phase superposition. *Chin. J. Sensors Actuators* **26**(7), 981–985 (2013)
3. Whitehouse, C.D.: The design of calamari: an ad-hoc 788 localization system for sensor networks. University of California at Berkeley, Berkeley (2002)
4. Li, M.: Development of the indoor navigation system for large public buildings. In: *Proceedings of 2017 International Conference on Computer, Electronics and Communication Engineering (CECE2017)*, p. 5. Science and Engineering Research Center (2017)
5. Farshad, A., Li, J., Marina, M.K., et al.: A microscopic look at WiFi fingerprinting for indoor mobile phone localization in diverse environments. In: *2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Busan, Korea, pp. 1–10. IEEE (2013)
6. Luo, Y., Hoeber, O., Chen, Y.: Enhancing Wi-Fi fingerprinting for indoor positioning using human-centric collaborative feedback. *Hum.-Centric Comput. Inf. Sci.* **3**(1), 1–23 (2013)

Multiple Rotorcrafts Environment Map Fusion for Atmosphere Monitoring



Pengxiang Bao, Lei Cheng, Xin Wang, Qin Liu and Qiuyue Yu

Abstract In order to improve the atmospheric environment monitoring mechanism and realize the construction of environmental maps, this paper proposes a proximity factor concentration fusion method for the problem of sub-map fusion to the construction of multiple rotorcrafts maps. The method refines the neighboring sub-maps to coincide with the boundary concentration factor and calculates the concentration factor of the overlapping area based on the factor mean algorithm to obtain a complete gas concentration map. The fusion of two sub-concentration maps is taken as an example in this paper. The two sub-concentration maps with short time difference before and after are fused into a map, and then the feasibility of the fusion method is verified by Fluent and Matlab simulation experiments, which provides the research foundation for the fusion of multiple gas concentration maps.

Keywords Multiple rotorcraft · Sub-map fusion · Factor mean algorithm · Factor concentration fusion

P. Bao (✉) · L. Cheng · X. Wang · Q. Liu
School of Information Science and Engineering, Wuhan University of Science and Technology,
Wuhan, China
e-mail: 1162873073@qq.com

L. Cheng
e-mail: chenglei@wust.edu.cn

X. Wang
e-mail: 1264266809@qq.com

Q. Yu
School of Mechanical Engineering, City College, Wuhan University of Science and Technology,
Wuhan, China

1 Introduction

At the 19th National Congress of the Communist Party of China, General Secretary Xi Jinping pointed out “We will get everyone involved in improving the environment and address environmental issues at the root. We will continue our campaign to prevent and control air pollution to make our skies blue again.” It is of great practical significance to develop an intelligent rotorcraft system that can realize the construction of a three-dimensional space environment gas map for improving the national atmospheric environment monitoring mechanism and enhancing the accuracy of atmospheric monitoring and early warning.

Map fusion plays an important role in the autonomous composition of robots. How to fuse the individual maps created by each robot is the key to improve the collaborative exploration efficiency of multiple mobile robots. Researchers all over the world have conducted relevant research on the fusion of robot maps. The existing map fusion methods are mainly proposed for geometric feature maps and grid maps. A grid map stitching method of image registration significance was proposed in [1]. This method regards the raster map splicing problem as the image registration problem to use a series of image registration principles and methods to achieve grid map stitching. For multi-robot map stitching problem, a multi-robot raster map stitching method based on Speeded-Up Robust Features (SURF) was studied in [2]. This method expresses the raster map stitching problem with the minimization problem of image registration and establishes a mathematical model by performing rigid body transformation on the robot motion coordinate system. Finally, experiments were carried out using the public dataset and the TurtleBot 2 mobile robot. The results showed that the proposed method could achieve reliable stitching of the grid maps established by multiple robots. Ergin et al. [3, 4] shared and fused local maps according to the encounter of two robots, and proposed an effective online map fusion method, but this method requires that the robots can accurately locate each other. By analyzing the mathematical model of mobile robots, a factor graph fusion algorithm based on Covariance Intersection (CI) was proposed in [5]. It attempts to calculate the pose information and covariance of fusion nodes and the covariance of the next level factor after fusion through CI fusion algorithm and covariance, then it can achieve map fusion, reduce fusion error, and improve positioning accuracy.

Liu et al. [6] adopted an exploration strategy of independent exploration and focused construction to propose a fusion algorithm, which established a dissimilarity function based on the concept of map similarity. The improved evolutionary algorithm search strategy was used to quickly seek the maximum overlap between local maps, and then the multi-robot system grid map fusion was realized, which effectively solved the problem of map creation under unknown relative position. The fusion of local maps requires that the communication network between the robots is a fixed topological structure, and the fusion of maps is triggered and implemented synchronously between the robots [7], which has certain limitations. Yuhua et al. [8]

proposed a dynamic map fusion algorithm based on the principle of information gain consistency to solve the problem affecting global map construction caused by the lack of information because of the limit of communication distance or the change of network topology in the map fusion. And the validity of the algorithm was verified by the RGB-D (color-depth) data of the actual environment. Ma et al. [9] proposed a grid map fusion method based on immune adaptive genetic algorithm. The algorithm used the optimization function that reflected the degree of dissimilarity of the overlapping regions of two grid maps as an antigen, and each possible translation, rotation plane conversion corresponded to an antibody. Finally, the simulation experiments showed that the algorithm could search for the best overlapping area of the two grid maps to achieve map fusion at faster convergence speed and stronger global search capabilities.

Most of the current research involves the fusion of laser maps and visual maps. However, it is rarely involved in the fusion of maps constructed by multiple rotorcrafts. Therefore, a factor mean algorithm is proposed to fuse the sub-concentration maps constructed by two rotorcrafts in this paper. The two-dimensional concentration map model is established through Fluent, and then the two-dimensional concentration map is constructed according to the concentration of CO₂ and the location information of the corresponding collection points along the planned path of the two rotorcrafts. The feasibility of this method is proved by the simulation experiments, which provides a basis for the fusion of three-dimensional concentration maps. The combination of map fusion can realize the rapid construction and determination of gas concentration maps of multiple rotorcrafts in a three-dimensional space environment. It can be applied to monitor the air quality over the city at any time, which is conducive to the construction of smart cities.

In this paper, the establishment of Fluent's two-dimensional concentration map model is introduced in the second section, and the factor mean algorithm based on the fusion of rotorcraft maps is proposed in the third section. Then the concentration maps are fused through simulation experiments in the fourth section, and the conclusion is drawn in the fifth section.

2 Model Establishment of Fluent's Two-Dimensional Concentration Map

The Fluent software is used to draw the concentration points collected by the two rotorcrafts into a complete concentration map in this paper, and the CFD software developed by Fluent of the United States is used to simulate the plume model with turbulent flow characteristics in a single-side wind environment. Fluent software can effectively describe two-dimensional fluid mechanics through numerical model and image display, and thus a two-dimensional gas concentration map can be established.

The establishment of the mathematical model of Fluent software requires a series of basic conservation equations. These basic conservation equations include mass conservation equations, momentum conservation equations, energy conservation equations, component conservation equations, and turbulence equations. Mass conservation equations, momentum conservation equations, energy conservation equations, and component conservation equations are, respectively, shown as in (1)–(4).

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) = 0 \quad (1)$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i u_j) = -\frac{\partial \rho}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\mu \frac{\partial u_j}{\partial x_i} \right) + (\rho - \rho_a) g_i \quad (2)$$

$$\frac{\partial(\rho T)}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i T) = \frac{\partial}{\partial x_i} \left(\frac{k}{c_p} \frac{\partial T}{\partial x_i} \right) + S_T \quad (3)$$

$$\frac{\partial(\rho c_s)}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i c_s) = \frac{\partial}{\partial x_i} \left(D_s \frac{\partial(\rho c_s)}{\partial x_i} \right) \quad (4)$$

Among the above equations, the gas studied ρ in this paper is an incompressible fluid, so the density is actually a constant; ρ_a is the air density; u_i, u_j , respectively, represents the velocity in the x, y direction μ, c_p, T, k , and S_T are fluid dynamic viscosity, specific heat capacity, temperature, heat transfer coefficient, and viscous dissipation. $c_s, \rho c_s$, and D_s correspond to the volume concentration, the mass concentration, and the diffusion coefficient of component s .

Standard $k-\varepsilon$ two-parameter model is a good choice to express the turbulent flow equation for it has strong stability and high precision. Turbulent flow equation is as in Eq. (5):

$$\rho \frac{Dk}{Dt} = \frac{\partial}{\partial x_i} \left(\left(\mu + \frac{\mu_t}{\mu_k} \right) \frac{\partial k}{\partial x_i} \right) + G_k + G_b - \rho \varepsilon - Y_M \quad (5)$$

Dissipation rating ε equation is as in Eq. (6):

$$\rho \frac{D\varepsilon}{Dt} = \frac{\partial}{\partial x_i} \left(\left(\mu + \frac{\mu_t}{\mu_k} \right) \frac{\partial \varepsilon}{\partial x_i} \right) + C_{1z} \frac{\varepsilon}{k} (G_k + C_{3z} G_b) - C_{2z} \rho \frac{\varepsilon^2}{k} \quad (6)$$

where $C_{1z}, C_{2z}, C_{3z}, k$, and ε are all constants. G_k is the turbulent kinetic energy generated by average velocity gradient, while G_b is generated by the effect of buoyancy.

The steps to solve the numerical simulation of the indoor complex environment studied by Fluent software are shown in Fig. 1.

In this paper, the position coordinates and the concentration information of CO_2 of each point have been measured by two rotorcrafts. Then, the Fluent software is used to plot the CO_2 concentration map of the two-dimensional horizontal plane.

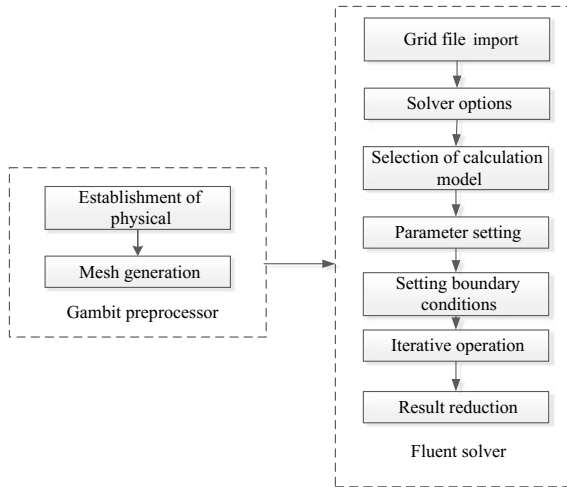


Fig. 1 Solution step flowchart

3 Mean Factor Algorithm

The factor mean algorithm is used to achieve the fusion of CO₂ sub-concentration maps in this paper. The factor mean algorithm is the mean method, which is generally used to calculate the mean value in mathematics and widely used in various fields. Mean value is an important concept in statistics. Generally speaking, the mean value refers to the arithmetic mean, which is the quotient of the sum of a group of data divided by the number of data in this group. The arithmetic mean in statistics is often used to represent the general level of statistical objects, and it is a statistic that describes the degree of data concentration. Therefore, in this paper, the factor means algorithm can be used to calculate the concentration of overlapping regions in two adjacent sub-maps, so that the two sub-concentration maps are fused into a complete map.

The factor mean algorithm can be used to represent a group of data visually and concisely, it is also very simple to calculate, such as a group of data $x_1, x_2, \dots, x_n$ are obtained, adding the sum of this group of data, and then the sum is divided by the number of the number n , so the mean of this group of the number is obtained. For example, the mean $\bar{x}$ of n numbers can be expressed as in Eq. (7).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (7)$$

In this paper, the flight paths of two rotorcrafts are planned within the two adjacent $5 \text{ m} \times 10 \text{ m}$ regions. Then, the concentration of the sensors collected by the sensors installed on the rotorcrafts when the rotorcrafts are following the planned path is collected at certain time intervals, and the concentration of each scanning position is obtained after the entire area has been searched. Then, the position and concentration information collected by rotorcraft P and Q , respectively, for each collection point are entered into the Fluent and Matlab software so that the two sub-concentration maps P_1 and Q_1 can be drawn, respectively. The sub-concentration maps P_1 and Q_1 are placed in the same coordinate system (the adjacent concentration map boundaries correspond to the adjacent area boundaries). In maps P_1 and Q_1 , the two groups of points $A_1, A_2, \dots, A_{21}; B_1, B_2, \dots, B_{21}$ are taken, respectively.

Equation (8) for solving the mean concentration is obtained according to Eq. (7), which can be used to achieve concentration map fusion.

$$\bar{C}_i(x_i, y_i) = \frac{1}{2} [C_{A_i}(x_i, y_i) + C_{B_i}(x_i, y_i)] \quad (8)$$

4 Simulation Experiment

In this paper, taking two adjacent rotorcrafts as an example to monitor the concentration of CO_2 at a fixed height, the search trajectories of the rotorcrafts have been pre-planned in the environment where there are two unilateral wind unilateral winds. According to investigations, there are positioning errors and control errors during the flight of a rotorcraft in a designated area, and therefore it is not possible to accurately fly along a planned path. Under normal circumstances, in order to cover the entire area, the rotorcrafts in neighboring areas will have overlapping scanning areas when scanning adjacent boundaries, which is called an overlap area. The rotorcraft P scans the adjacent boundary area about half a minute earlier than the rotorcraft Q in order to avoid the rotorcrafts colliding in the adjacent boundary areas.

The simulation experiment steps in this article are as follows.

- (1) The search trajectory of the rotorcraft has been planned in advance in a unilateral wind environment. The concentration of CO_2 is measured by two rotorcrafts which measured at the two adjacent planar areas (P_1 and Q_1 areas) of a fixed height of two meters with a certain sampling frequency. Rotorcraft P scans the overlapping area approximately at 30 s and rotorcraft Q scans the overlapping area approximately at 60 s.
- (2) The two-dimensional horizontal concentration map of CO_2 of the rotorcraft is constructed in Fluent and Matlab software according to the collected concentration of CO_2 , Eqs. (1)–(6), and the flowchart of Fig. 1. The maps of regions P_1 and Q_1 simulated by Fluent and Matlab software are shown in Fig. 2 and Fig. 3, respectively.

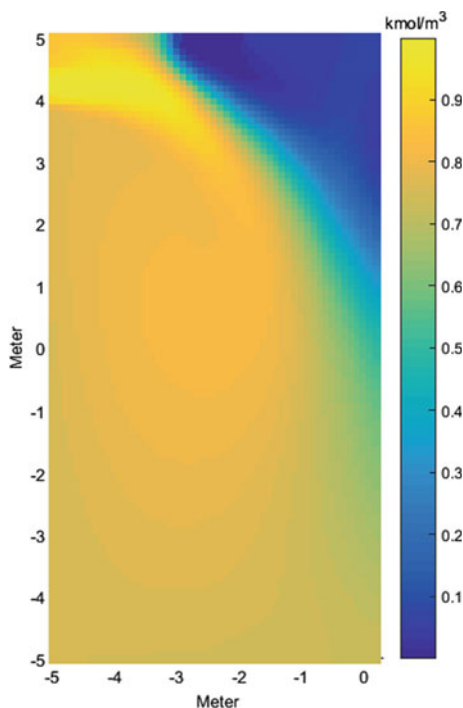


Fig. 2 Map of the concentration of CO_2 of the area scanned by rotorcraft P

- (3) The boundary points of the overlapping area are taken as shown in Figs. 4 and 5 according to the concentration of CO_2 in Figs. 2 and 3. The mean algorithm (such as Eqs. (7), (8)) is used to calculate the concentration of the overlapping area, and Figs. 1 and 2 are merged into Fig. 6.
- (4) All concentration values of CO_2 are collected by rotorcraft P and rotorcraft Q which are put in the same simulation environment. The construction of the concentration map of CO_2 is completed in Fluent and Matlab software as shown in Fig. 7, according to Eqs. (1)–(6), and Fig. 1 shows the flowchart of the steps.

According to many measurements, the rotorcraft P crosses the boundary by approximately 18 cm error into the area where the rotorcraft Q is located. Similarly, the rotorcraft Q passes the boundary with the error of nearly 16 cm into the area where the rotorcraft P is located. In order to facilitate the fitting and calculation, the overlapping area takes the 16 cm distance crossed by the adjacent boundary as the standard.

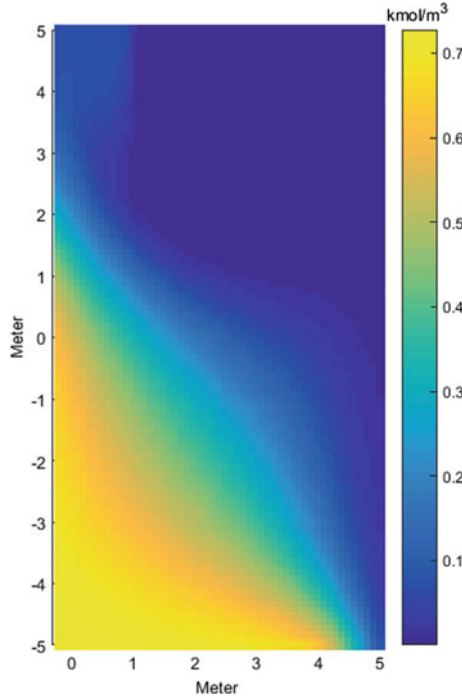


Fig. 3 Map of the concentration of CO_2 of the area scanned by rotorcraft Q

The two groups of points, $A_1, A_2, \dots, A_{21}; B_1, B_2, \dots, B_{21}$, respectively, taken from the overlapping area of rotorcraft P and Q searches, are taken as an example. $A_i (i = 1, 2, \dots, 21)$ is a list of points whose interval between the ordinates is 0.5 m and near the right side in Fig. 4 at the distance of 0.16 m (a list of points with a 0.00 abscissa in Fig. 4). $B_i (i = 1, 2, \dots, 21)$ is a list of points whose interval between the ordinates is 0.5 m on the left boundary in Fig. 5 (a list of points with a -0.16 abscissa in Fig. 5). The factor mean algorithm is used to fuse the concentration map constructed by rotorcraft P (Fig. 2) with the concentration map constructed by rotorcraft Q (Fig. 3), and the map shown in Fig. 6 is obtained. Table 1 shows the coordinates and concentration values of A_i and B_i in Figs. 4 and 5 and the average concentration $\bar{C}_i$ of point A_i and point B_i calculated by mean algorithm.

Similarly, the two groups of points $C_1, C_2, \dots, C_{21}, D_1, D_2, \dots, D_{21}$ are shown in Fig. 4 and Fig. 5, respectively. $C_i (i = 1, 2, \dots, 21)$ is a list of points whose interval between the ordinates is 0.5 m and near the right side in Fig. 4 at the distance of 0.15 m (a list of points with a 0.01 abscissa in Fig. 4). $D_i (i = 1, 2, \dots, 21)$ is a list of points whose interval between the ordinates is 0.5 m and near the left side in Fig. 5 at the distance of 0.01 m (a list of points with a -0.15 abscissa in Fig. 5). The

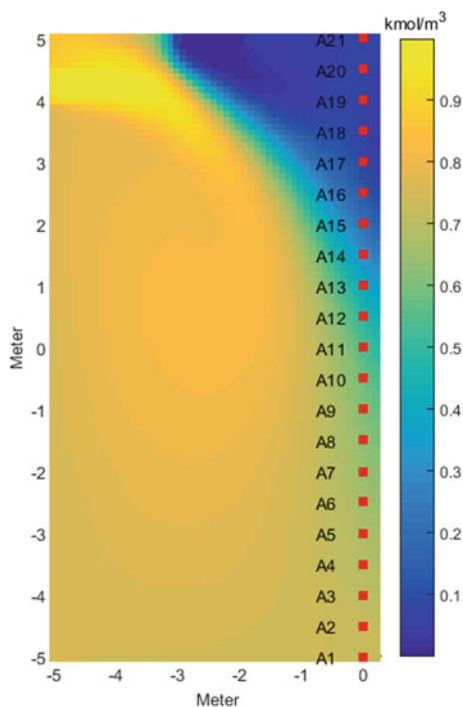


Fig. 4 Points taken from the map of the concentration of CO_2 of the area scanned by rotorcraft P

concentration values of these two sets of points are also fitted by the mean algorithm. All the CO_2 concentration values of the overlapping area are, respectively, taken in Fig. 4 and Fig. 5, that is, the value is taken from a column of points with a horizontal coordinate of 0.00 to the right boundary at an interval of 0.01 m in Fig. 4, and the value is taken at intervals of 0.01 m from the left boundary to the right in Fig. 5. Until the last set of numbers that the abscissa in Fig. 4 is 0.16 and the abscissa in Fig. 5 is the value 0.00, their ordinates are a set of points spaced by 0.5 m that are taken. Finally, the maps that are fused by the mean algorithm are shown in Fig. 6.

All concentration values of CO_2 are collected by rotorcraft P and rotorcraft Q , which are put in the same simulation environment. The construction of the concentration map of CO_2 is completed in Fluent and Matlab software as shown in Fig. 7. Compared with Fig. 6, it can be seen that Fig. 7 is very similar to Fig. 6. There are only some traces of fitting in the coincident area with an ordinate of 3–5 m, and there is almost no difference in other areas.

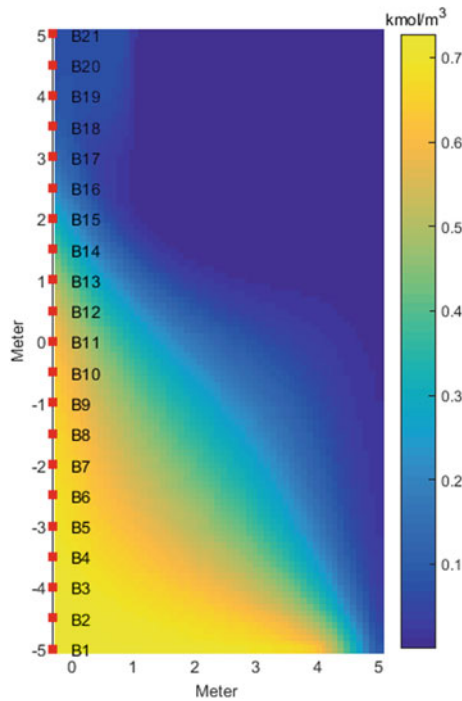


Fig. 5 Points taken from the map of the concentration of CO₂ of the area scanned by rotorcraft *Q*

5 Conclusions

For the fusion of plane sub-concentration maps based on aerial rotorcraft, this paper proposes a factor mean algorithm for fusion of two-dimensional concentration maps. The mean algorithm, having a low computational complexity, is easy to operate and widely used. The concentration in the overlapping area collected by the two rotorcrafts at a time interval of approximately half a minute is plotted in Fluent and Matlab software through simulation experiments. The mean value algorithm is used to fit the concentration of CO₂ for the map overlapping area caused by positioning error and control error. In addition, all the concentration values of CO₂ collected by the two rotorcrafts are placed in the same simulation environment to obtain a concentration map, which is very similar to the fusion map fitted with the mean algorithm, with only slight differences, which proves that the factor mean algorithm can be used for concentration map fusion and lays a foundation for the fusion of three-dimensional spatial concentration maps. This method can be applied to the smart city's atmospheric monitoring and the construction of the gas environment map.

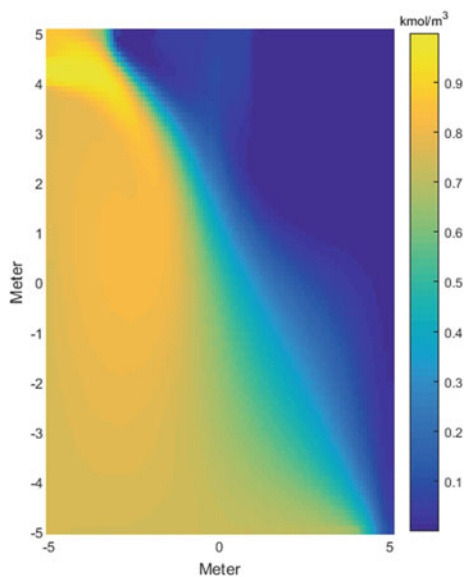


Fig. 6 Map fitted with the mean algorithm. There are only a few traces of fusion in the map after fusion with the mean algorithm, indicating that the fusion algorithm with the mean algorithm is very good

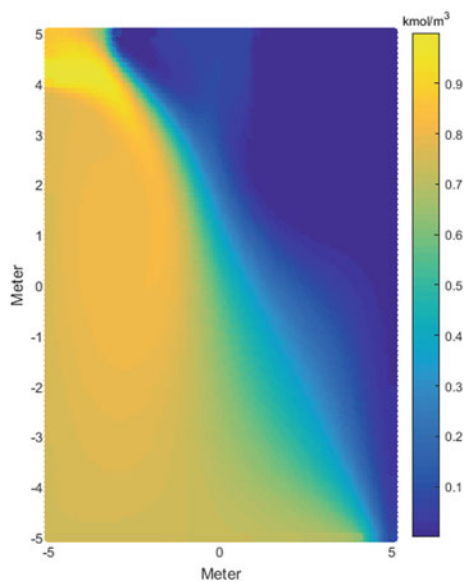


Fig. 7 Map drawn from all concentration points. The concentration map of CO_2 obtained by placing all coordinates and concentration of CO_2 measured by the two rotorcrafts in the same simulation environment

Table 1 Coordinates and concentration table

Point A_i coordinate	$C_{A_i}(x_i, y_i)$	Point B_i coordinate	$C_{B_i}(x_i, y_i)$	$\bar{C}_i(x_i, y_i)$
$A_1(0.00, -5.00)$	0.7365	$B_1(-0.16, -5.00)$	0.7302	0.7335
$A_2(0.00, -4.50)$	0.7259	$B_2(-0.16, -4.50)$	0.7176	0.7218
$A_3(0.00, -4.00)$	0.7162	$B_3(-0.16, -4.00)$	0.7104	0.7133
$A_4(0.00, -3.50)$	0.7095	$B_4(-0.16, -3.50)$	0.7074	0.7085
$A_5(0.00, -3.00)$	0.6887	$B_5(-0.16, -3.00)$	0.6798	0.6843
$A_6(0.00, -2.50)$	0.6573	$B_6(-0.16, -2.50)$	0.6483	0.6528
$A_7(0.00, -2.00)$	0.6341	$B_7(-0.16, -2.00)$	0.6279	0.6355
$A_8(0.00, -1.50)$	0.6172	$B_8(-0.16, -1.50)$	0.6094	0.6133
$A_9(0.00, -1.00)$	0.5791	$B_9(-0.16, -1.00)$	0.5696	0.5744
$A_{10}(0.00, -0.50)$	0.5672	$B_{10}(-0.16, -0.50)$	0.5573	0.5623
$A_{11}(0.00, 0.00)$	0.5386	$B_{11}(-0.16, 0.00)$	0.5307	0.5346
$A_{12}(0.00, 0.50)$	0.4975	$B_{12}(-0.16, 0.50)$	0.5022	0.4999
$A_{13}(0.00, 1.00)$	0.3824	$B_{13}(-0.16, 1.00)$	0.3790	0.3807
$A_{14}(0.00, 1.50)$	0.3371	$B_{14}(-0.16, 1.50)$	0.3276	0.3324
$A_{15}(0.00, 2.00)$	0.2881	$B_{15}(-0.16, 2.00)$	0.3045	0.2963
$A_{16}(0.00, 2.50)$	0.2142	$B_{16}(-0.16, 2.50)$	0.2098	0.2120
$A_{17}(0.00, 3.00)$	0.1179	$B_{17}(-0.16, 3.00)$	0.1265	0.1222
$A_{18}(0.00, 3.50)$	0.0903	$B_{18}(-0.16, 3.50)$	0.1047	0.0975
$A_{19}(0.00, 4.00)$	0.0795	$B_{19}(-0.16, 4.00)$	0.0892	0.0844
$A_{20}(0.00, 4.50)$	0.0706	$B_{20}(-0.16, 4.50)$	0.0773	0.0740
$A_{21}(0.00, 5.00)$	0.0750	$B_{21}(-0.16, 5.00)$	0.0735	0.0743

Acknowledgements This work is supported by four Projects from National Natural Science Foundation of China (60705035, 61075087, 61573263, 61273188), Scientific Research Plan Key Project of Hubei Provincial Department of Education (D20131105), and Project supported by the Zhejiang Open Foundation of the Most Important Subjects, also supported by Zhejiang Provincial Natural Science Foundation under Grant LY16F030007 and Hubei Province Science and Technology Support Project under Grant 2015BAA018.

References

1. Zhu, J.H., Zhou, Y., et al.: Grid map mosaicing method based on image registration. *Acta Automatica Sinica* **41**(02), 285–294 (2014)
2. Tang, H.W., Sun, W., et al.: Multi-robot raster map stitching method based on SURF feature. *J. Electron. Meas. Instrum.* **31**(06), 859–868 (2017)
3. Ergin O.N., Levent, A.H.: Cooperative multi-robot map merging using fast-SLAM. In: *Proceedings of Robolup* (2009)
4. Monica, B., Arturo, G., Oscar, R., et al.: Alignment of visual maps in multi-robot Fast SLAM. In: *Proceedings of the 8th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics* (2009)
5. Wang, W., Pu, Y.M., Li, W.: A multi-robot 2D map fusion algorithm based on CI factor graph. *Robotics* **39**(06), 872–878 (2017)
6. Liu, L.M., Cai, Z.X.: Research on multi-robot map fusion method. *Microcomput. Syst.* **33**(09), 1934–1937 (2012)
7. Aragues, R., Cortes, J., Sagues, C.: Distributed consensus on robot networks for dynamically merging feature-based maps. *IEEE Trans. Robot.* **28**(4), 840–854 (2012)
8. Zou, Y.H., Chen, W.H., Wang, J.H., Wu, X.M.: Multi-robot map fusion algorithm based on information gain consistency. *Robot* **36**(05), 619–626 (2014)
9. Ma, X., Song, R., Guo, R., Li, Y.B.: Robotic grid map fusion based on immune adaptive genetic algorithm. *Control Theory Appl.* **26**(09), 1004–1008 (2009)

Simulated Tests of Feedforward Active Noise Control (ANC) for Building Noise Cancellation



Tongrui Peng, Quanmin Zhu, M. Osman Tokhi and Yufeng Yao

Abstract This paper presents an understanding of active noise control (ANC) systems especially focus on the physical geometries of the detector and the observer and develops a series of simulation experiments for different types of signals to test the cancellation performance using the feedforward active noise control procedure with the fixed controller proposed by Leitch and Tokhi (IEE Proc PS-A 134(6):525–546, 1987). This study demonstrates the effectiveness of the proposed control approach and also builds up a foundation for applying the proposed new ANC algorithms in future studies.

Keywords Building noise • Low-frequency noise • Active noise control • Feedforward ANC control • ANC with fixed controller

1 Introduction

Building noise and many engineering processes are important in our daily life, and its consequences will have the significant impact on people's living quality both in physical and psychological aspects [1], especially in relation to both human comfort and indoor environmental quality [2, 3].

Currently, there are two distinct methods used to reduce the sound pressure levels or the sound intensity levels of unwanted noise, i.e., the passive noise control (PNC) and the active noise control (ANC). Compared to PNC, ANC is an advanced and efficient technology for the cancellation of the unwanted noise in low-frequency range

T. Peng · Q. Zhu (✉) · Y. Yao

Department of Engineering Design and Mathematics, University of the West of England,
Bristol BS16 1QY, UK
e-mail: Quan.Zhu@uwe.ac.uk

M. O. Tokhi

Department of the Electrical and Electronic Engineering, London South Bank University, London
SE1 0AA, UK

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_11

115

as the wavelengths at low frequencies become larger compared with the conventional acoustic absorbers which are the main method to implement passive noise control [4, 5].

Lueg proposed the first concept of ANC in 1936 [4], and several literature reviews have already summarized the development of ANC [5–7]. The ANC systems can be broadly classified into two categories: the feedforward ANC system and the feedback ANC systems. For a feedforward ANC system, a reference microphone is used for detecting unwanted noise and the loudspeaker is placed at a flexible distance to the microphone to generate the secondary acoustic source, while for a feedback ANC system, there is no reference microphone to provide detected primary noise source.

In 1987, Leitch and Tokhi [6] discussed the optimal configuration of feedforward ANC systems as there was little information provided prior to this publication. There are three aims of discussing optimal configuration. First, in order to achieve the aim of optimal cancellation, the controller is required with a frequency-dependent transfer function that can exactly produce the secondary noises. The characteristics of the controller depend on characteristics of the secondary noise sources, propagation path and electronic transducers. Second, the acoustic feedback from the loudspeaker to the detector and the distance ratio from the detector to primary sources and secondary sources cause the instability for the whole control system. Therefore, how to set up the system can avoid instability. Finally, the adjustment of the separation between the primary source and the secondary source affect the reinforcement or cancellation area at the observation point.

In this paper, a special treatment called the fixed controller will be introduced in the feedforward ANC systems, and four input signals containing different frequency components will be tested on MATLAB/SIMULINK platform to demonstrate the performance of proposed framework. Furthermore, discussions and guidelines about the selection of the fixed controller and the setting up of observer and detector will be given.

2 Feedforward and Feedback ANC Systems

Figure 1 gives the geometry configuration about feedback ANC system proposed by Leitch and Tokhi [6].

Here, in Fig. 1,

- D the physical distance between primary sources and secondary sources;
- R_1 the physical distance between primary sources to observer/detector; and
- R_2 the physical distance between secondary sources to observer/detector.

The primary source emits a pressure wave, which is detected by the detector, located at a distance of R_1 relative to the primary source. The microphone transfers the acoustic wave to the electrical signal and passes it to the controller. The controller here is required to adjust the amplitude and phase of each component of the detected

Fig. 1 A schematic diagram of the feedback ANC system [6]

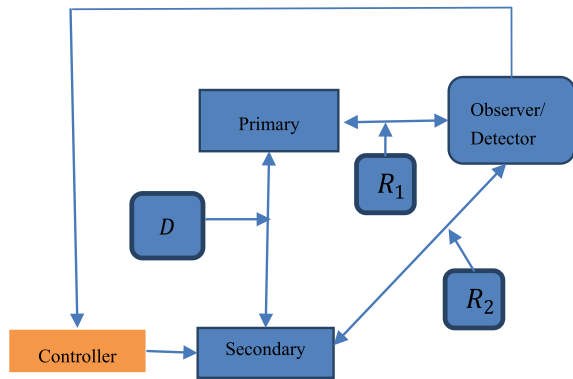
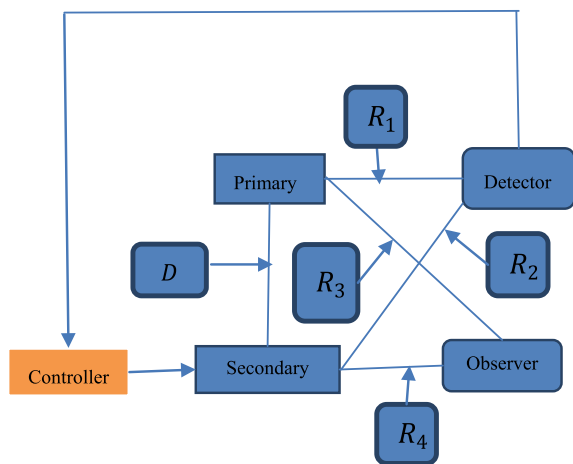


Fig. 2 A schematic diagram of the feedforward ANC system [6]



signal, and later it will be emitted to drive the loudspeaker to generate the secondary acoustic source. It was found that the optimal situation is that at the observation point, the noise levels can be reduced to zero. The characteristics of the feedback ANC system are that observer coincides with the detector.

Here, in Fig. 2,

- D the physical distance between primary sources and secondary sources;
- R_1 the physical distance between primary sources to the detector;
- R_2 the physical distance between secondary sources to the detector;
- R_3 the physical distance between primary sources to the observer; and
- R_4 the physical distance between secondary sources to the observer.

Figure 2 illustrates the geometry configuration of feedforward active noise control (ANC) system proposed by Leitch and Tokhi [6]. Compared with the feedback ANC system shown in Fig. 1, the detector and the observer are placed at separated locations with different distances relative to the primary source and the secondary source.

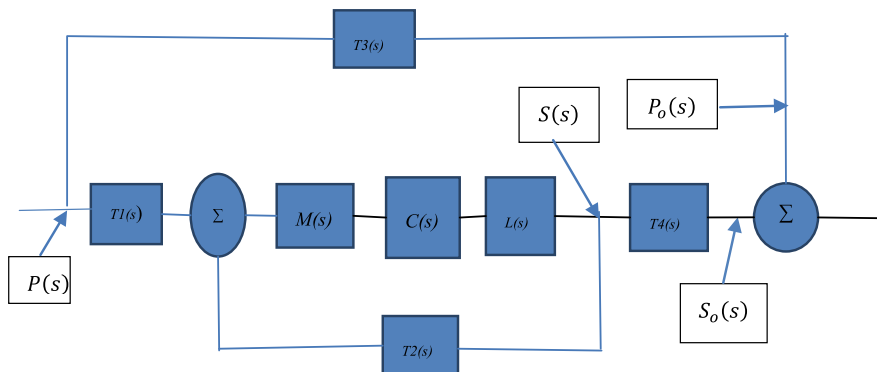


Fig. 3 Block diagram of the feedforward ANC system [6]

$T1(s)$ the transfer characteristics of path R_1 ;
 $T2(s)$ the transfer characteristics of path R_2 ;
 $T3(s)$ the transfer characteristics of path R_3 ;
 $T4(s)$ the transfer characteristics of path R_4 ;
 $M(s)$ the transfer characteristics of the microphone;
 $L(s)$ the transfer characteristics of the loudspeaker; and
 $C(s)$ the transfer characteristics of the controller.

From Fig. 3, the primary source at the observation point

$$P_o(s) = T3(s)P(s) \quad (1)$$

The secondary source at the observation point is

$$S_o(s) = T4(s)S(s) \quad (2)$$

The mathematical expression for the secondary source is at the loudspeaker

$$S(s) = [T2(s)S(s) + T1(s)P(s)]M(s)C(s)L(s) \quad (3)$$

The expected results are at the observation point

$$P_o(s) + S_o(s) = 0 \quad (4)$$

Finally, the required controller continuous-frequency transfer function can be obtained as

$$C(s) = \frac{T3(s)}{M(s)L(s)[T2(s)T3(s) - T1(s)T4(s)]} \quad (5)$$

Equation (5) is the required transfer function for the optimal cancellation for the feedforward active noise control system, and it is clear that the system controller depends on transfer characteristics of the propagation medium and the electronic transducers.

In Fig. 2, if the detector or the observer is located at equidistance relative to the primary source and the secondary source, the whole controller is called the fixed controller and Eq. (6) below represents the controller's transfer function when the observer is located at equidistance from two sources:

$$C(s) = \frac{1}{M(s)L(s)[T2(s) - T1(s)]} \quad (6)$$

Theoretically, the optimal cancellation will occur if both transducers are located at equidistance from the primary source and the secondary source, or the distance ratio for both transducers relative to sound sources is equal. However, the corresponding controller under the condition of the equal distance ratio requires for an infinite larger gain and it is impossible to implement physically.

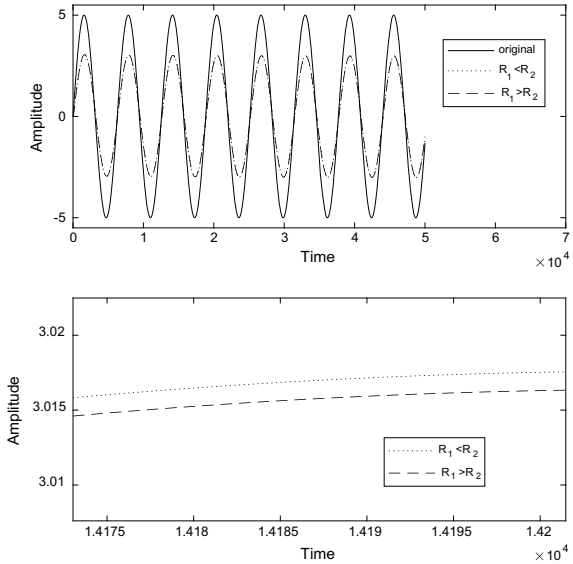
3 Simulation Studies

In this section, two computational experiments are implemented on MATLAB/SIMULINK platform to compare the cancellation performance of two different types of ANC system with the fixed controller. The first case uses a narrowband signal as the undesired noise and the second case uses a mixed signal consisting of a narrowband signal and a broadband signal as the undesired noise. For all computational experiments, the fundamental time step is $1e-4$ and the overall simulation period is 10 s.

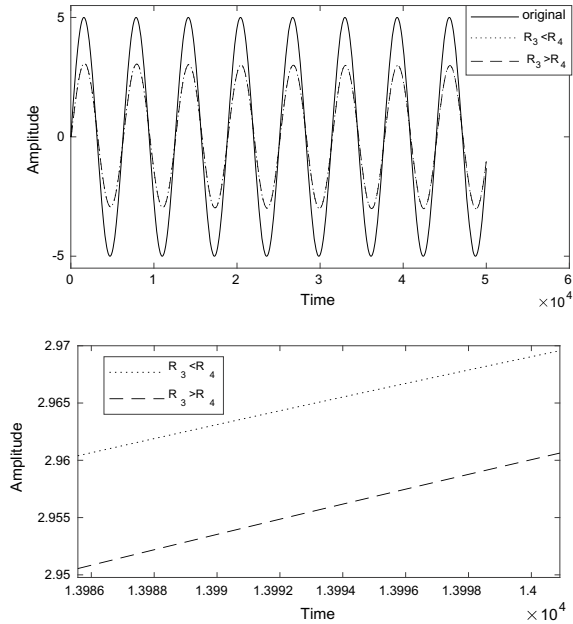
3.1 Narrowband Signal

A sine wave $y = \sin(t)$ is used as the primary (unwanted) noise, and the corresponding simulation results are displayed in Fig. 4.

Fig. 4 Comparison results



(a) $R_3 = R_4, R_1 < R_2 \& R_1 > R_2$



(b) $R_1 = R_2, R_3 < R_4 \& R_3 > R_4$

In Fig. 4a, the simulation result is for the condition that $R_3 = R_4$. The dotted line and the dash line denote to the condition that $R_1 < R_2$ and $R_1 > R_2$, respectively. In Fig. 4b, the simulation result is under the condition that $R_1 = R_2$. The dotted line and the dash line represent the condition that $R_3 < R_4$ and $R_3 > R_4$, respectively. For the narrowband noise, the proposed system can achieve the aim of the cancellation but the degree of performance depends on the configuration of transducers. It clearly reflects that for both conditions, $R_3 = R_4$ and $R_1 = R_2$, the performance will be better if the transducer is placed closer to the secondary source.

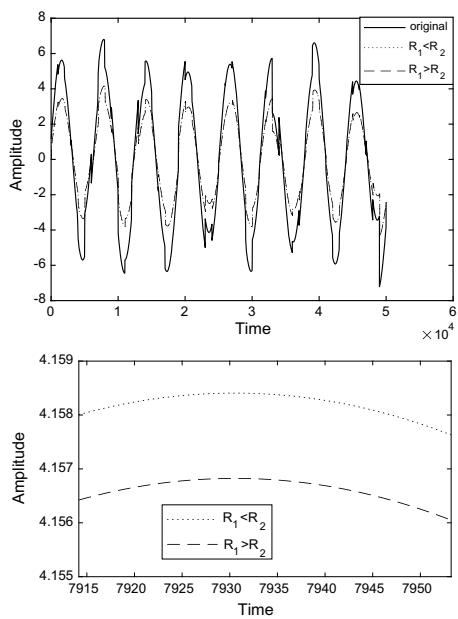
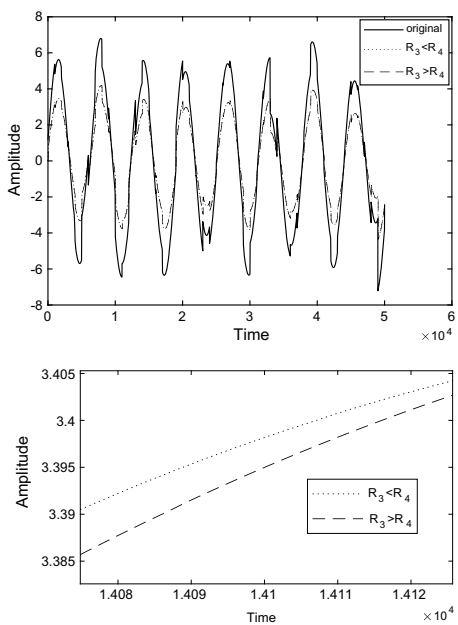
3.2 The Mixed Noise

For this simulation case, a mixed noise consists of a narrowband noise and a broadband noise is used as the undesired noise source. The corresponding result is depicted in Fig. 5.

In Fig. 5a, the simulation result is under the condition that $R_3 = R_4$. The dotted line and the dash line denote the condition that $R_1 < R_2$ and $R_1 > R_2$, respectively. In Fig. 5b, the simulation result is under the condition that $R_1 = R_2$. The dotted line and the dash line represent the condition that $R_3 < R_4$ and $R_3 > R_4$, respectively. The simulation results demonstrate the same conclusion that the performance will be better if the transducer is placed closer to the secondary source.

4 Conclusions

Study on the feedforward Active Noise Control (ANC) system with the fixed controller is presented in this paper. Simulation results from two case studies demonstrate the cancellation capability of the proposed system by Leitch and Tokhi (1987) [6]. Besides, for the feedforward ANC system with the fixed controller, if the transducer placed close to the secondary source, the cancellation performance is much better. The reason for this is that for the secondary source, the controller has enough time to adjust the amplitude and phase of each component of the coming electrical signal and the cancellation can be fully achieved at the observation point. Otherwise, if the path for electrical signal passing is longer than the path for the primary source passing, the secondary source will be later when arriving at the observation point and definitely, the performance is not good enough. Further improvement could, for example, use the FIR filter to replace the fixed controller to achieve the aim of cancellation.

(a) $R_3 = R_4, R_1 < R_2 \& R_1 > R_2$ (b) $R_1 = R_2, R_3 < R_4 \& R_3 > R_4$ **Fig. 5** Comparison results

References

1. Mak, C.M., Wang, Z.: Recent advances in building acoustics: an overview of prediction methods and their applications. *Build. Environ.* **91**, 118–126 (2015)
2. Cheuk, F.N.: Effects of building construction noise on residents: a quasi-experiment. *J. Environ. Psychol.* **20**(4), 375–385 (2000)
3. Sang, H.P., Pyoung, J.L., Jeong, H.J.: Effects of noise sensitivity on psychophysiological responses to building noise. *Build. Environ.* **136**, 302–311 (2018)
4. Lueg, P.: Process of Silencing Sound Oscillations, US Patent 2043416, June 9, 1936
5. Elliott, S.J., Nelson, P.A.: Active noise control. *IEEE Signal Process. Mag.* **10**(4), 12–35 (1993)
6. Leitch, R.R., Tokhi, M.O.: Active noise control systems. *IEE Proc. PS-A* **134**(6), 525–546 (1987)
7. George, N.V., Panda, G.: Advances in active noise control: a survey with emphasis on recent nonlinear techniques. *Sig. Process.* **93**(2), 363–377 (2013)

Research on Production Layout Design of Concrete Prefabricated Units Based on SLP



Yang Liu and Ding Li

Abstract Recently, with the rapid development of the domestic prefabricated concrete units industry in China, the layout of the beam field plays an exceedingly significant role in the efficiency and cost of the production on the prefabricated concrete. The research builds the assessment system based on the analysis of the process flow of the prefabricated box beam and combines with some relative factors on workshops of the prefabricated concrete, which designs the layout scheme of three workshops according to the unit area. The author gets the optimal layout design scheme of the prefabricated box beam by using the module of AHP (Analytical Hierarchy Process) and yaahp software which are used to calculate and analyze three layout schemes.

Keywords Production layout design · SLP · Concrete prefabricated units · AHP

1 Introduction

Prefabricated concrete structure [1] is the structural form of assembling on the spot after molding the concrete members in the factory, which not only improve the quality of the steel and concrete product but also guarantee the construction time limit for the project, and it is widely used in the developed countries, including some European countries and America, as the main form of new constructions.

In recent years, China has made great progress in studying prefabricated concrete production; meanwhile, the researchers at home and abroad carry out the study centering about the technology of concrete units, the formation, and prevention of cracks in prefabricated concrete production cycle [2], and the innovative technology in production and maintenance of prefabricated concrete units [3]. However, few

Y. Liu (✉)

School of Transportation Engineering, Shenyang Jianzhu University, Shenyang, China

e-mail: 21027819@qq.com

D. Li

Department of Business Administration, Liaoning Broadcasting and Television University, Shenyang, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_12

study focused on the production of prefabricated concrete units, particularly the production workshop layout of prefabricated concrete. Therefore, in this paper, it mainly demonstrates the analysis of prefabricated concrete production layout design based on SLP (Systematic Layout Planning) [4].

2 Original Condition

The position of the beam field is DK319+000 – DK319+540 on the northern side of the line, covering the area of 224 mu with convenient transportation and flat terrain. The field, on the south of line G311, is connected to the line X014, which will help for transporting the resource and meeting the requirement of annual output of prefabricated beam up to 544 pieces after completion.

According to the process flow and feature, the beam yard can be divided into 14 work units, including storing the materials, preparing, processing, and assembling the products. There are 10 beam pedestals, and 104 beam seat deposits. Set the beam pedestals and format model one with one. Two gantry cranes in the beam making area are used for moving the heavy equipments. A matching scheme is composed of one pedestal and one formwork. The girder area is equipped with two 120T gantry cranes, which are used as steel bars for heavy lifting. The prefabricated beam hole is transported from the girder pedestal to the beam storage pedestal by two 450T beam machines. When the beam is erected, two 450T girders are used for loading and the beam is transported by four 250T beam transporting vehicles to the line girder, and then the beam is carried to the bridge. With 900T transporting vehicle and bridge machine, the beam hole is transported to the designed position. An HSL120*2 mixing station is set up to ensure the concrete supply.

The prefabricated sequence of beam is based on the overall requirements of gantry crane beam [5]. Prefabrication technology uses the processing of post-tensioned prestress beam production. Beam field centralized prefabrication is used for beam piece construction. The process flow of box beam prefabricated is shown in Fig. 1.

From the process flow of box beam prefabricated, we can draw the conclusion that box beam is composed of formwork construction, reinforcement construction, concrete construction, prestressed construction, duct grouting, blocking, box girder transporting, and loading.

The plant production type is determined by the variety of products and the level of each product, which in turn affects the layout of plant. Based on the given condition, the product variety of beam field to be designed is single, but it has an enormous output of 544 annually, which belongs to the mass production. Therefore, it is suitable for the production to apply the fixed pedestal method.

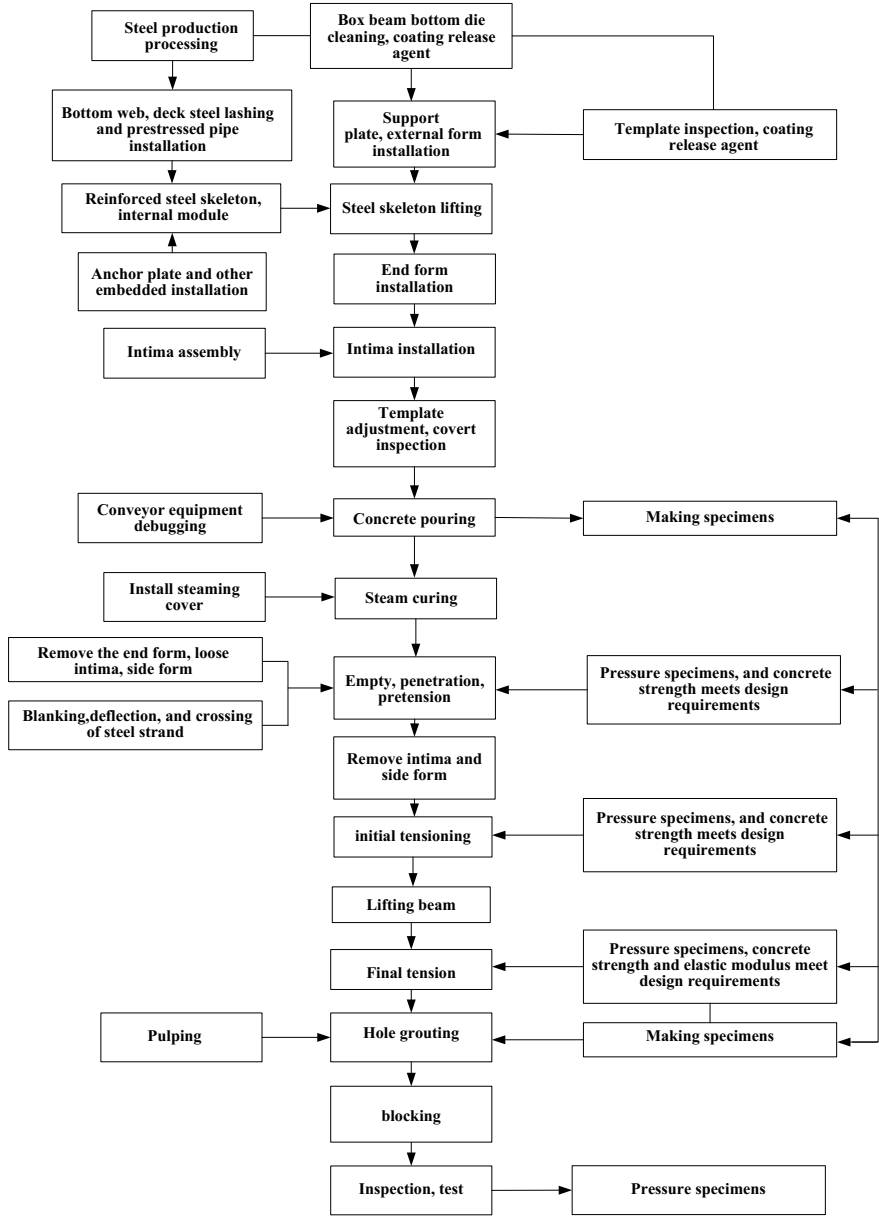


Fig. 1 Box beam process flowchart

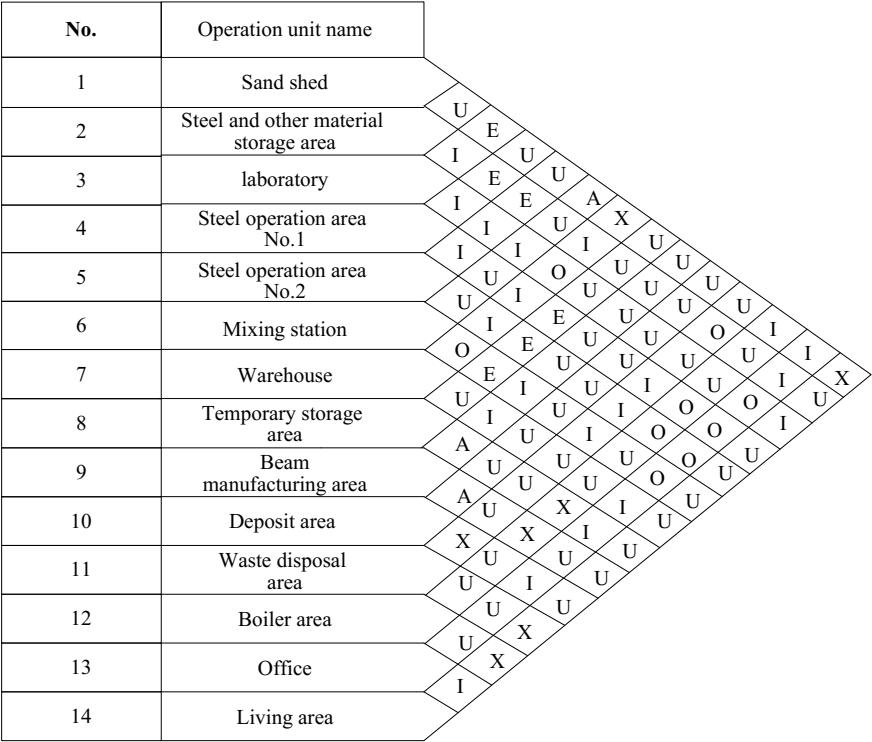


Fig. 4 Comprehensive interrelationships of operation units for box beam prefabricated yard

The selection of weighted value: Weighted value reflects the focus of attention on plant. For prefabricated box beam, the effect of logistics factors (*m*) is greater than the effect of the rest of non-logistics factors (*n*). Therefore, it is concluded that the weighted value is $m:n = 2:1$ [9].

The calculation of comprehensive correlation: Quantify the advanced level of logistics and logistics relations between the various operating units of the plant, and calculate weighted sum.

4.2 Draw the Unit Comprehensive Correlation Table

According to the comprehensive correlation calculation table of operation units in box beam prefabricated factory, the comprehensive correlation table is drawn as shown in Fig. 4.

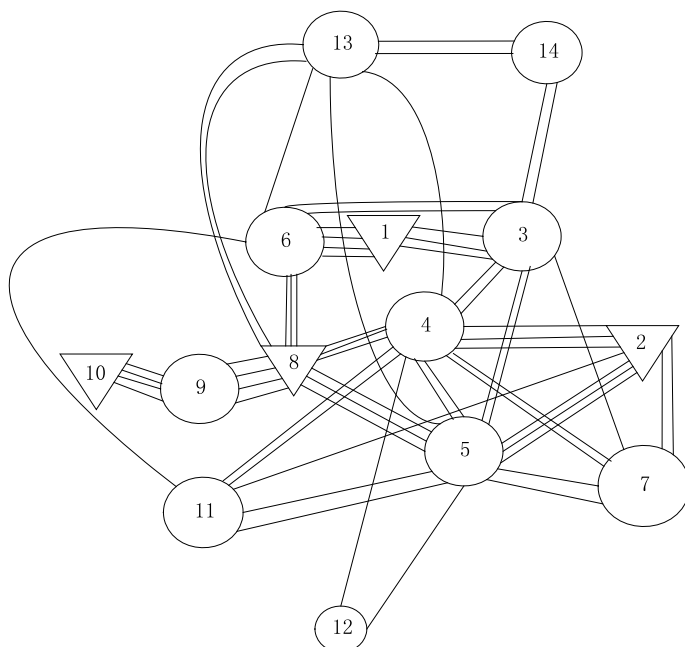


Fig. 5 Related position map of box beam prefabricated field

5 The Drawing of Units Location of Comprehensive Operation

Location correlogram of operation units can be obtained by the comprehensive proximity of operation units and hierarchical representation of it. Then, place in operation department-related diagram as shown in Fig. 5, use the number to represent the operation department, the relationship between operation department with how much solid line connection, line number, the more the closer to each other; Line number less can stay away from.

6 The Layout Design of the Manufacturing Area

We can draw the three layout design schemes as shown in Fig. 6, combing the area of each working unit with spatial geometric shape of their architecture on the relative chart of the working unit: 1-Sand and stone shed, 2-the warehouse for the concrete

iron and some material, 3-laboratory, 4-working unit1 for concrete iron, 5-working unit2 for concrete iron, 6-mixer station, 7-storehouse, 8-temporary storage area, 9-beam making area, 10-beam storage area, 11-waste disposal area, 12-boiler room, 13-office area, and 14-living area.

7 Selection and Optimization of the Scheme

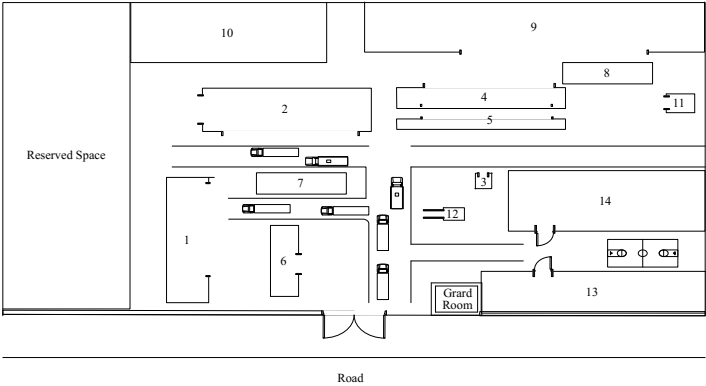
AHP [10, 11] was proposed by American scholar Saaty in 1970s that is a new evaluation method combining qualitative analysis with quantitative analysis. We use this method to decompose the complicated issue into several units, according to the fact, defining each level with quantitative expression, in which the mathematic method would be used to determine numerical value of comparative significance order of all the elements in each level and finally the problem would be solved through all the analysis.

Through the initial understanding and analysis of the evaluation index as shown in Fig. 7, we use the software yaahp to build hierarchical structure model and calculate.

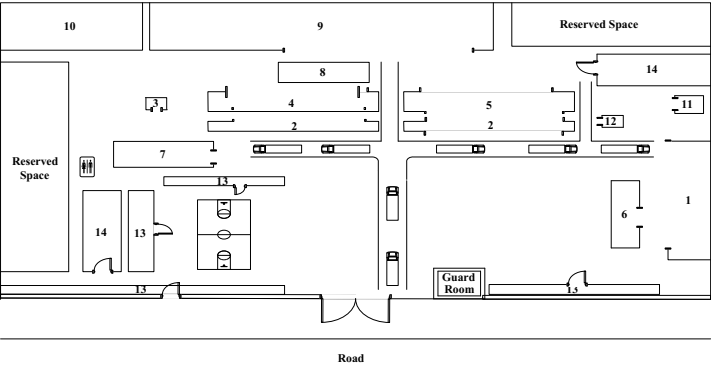
As the result calculated by yaahp, all of the relative imperative degrees are acceptable because of the allowed range of the concordance index. Design scheme 1 is 0.3331, and Design scheme 2 is 0.0974, then Design scheme 3 is 0.5695. Therefore, the best layout scheme of prefabricated beam box is the Design scheme 3 with maximum weight, and Design scheme 1 and Design scheme 2 with the minimum weight.

8 Conclusions

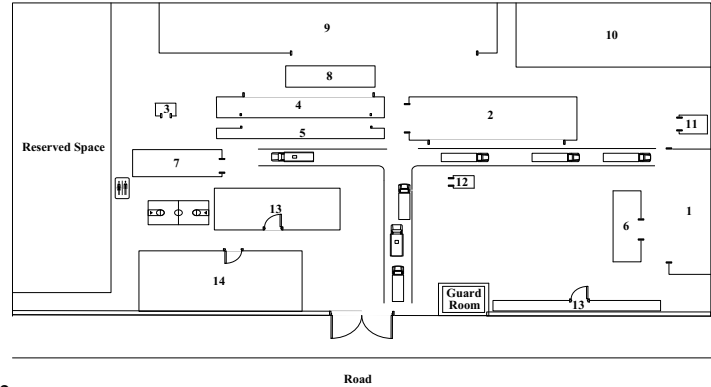
Based on preparation of the process flow of prefabricated box beam production, we utilized SLP method to obtain comprehensive proximity, location correlogram of operation units, area correlogram, and the ultimate layout of beam field. Moreover, we evaluated the layout scheme with AHP, establishing evaluation index system as main elements, such as transportation speed, pollution degree, and investment cost. In addition, the hierarchical structure model was run with yaahp software, in order to build judgment matrix. Finally, the efficiency of prefabricated element production has been improved and the cost of it has been descended.



(a) Optimized layout 1



(b) Optimized layout 2



(c) Optimized layout 3

Fig. 6 Three design schemes

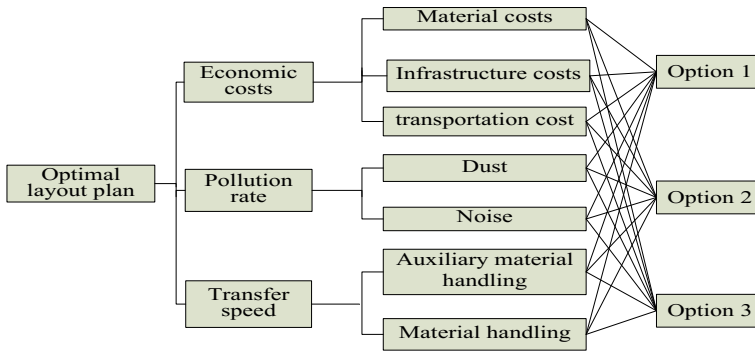


Fig. 7 Analysis of the evaluation index

References

1. Tian, L.D., Li, L.M., Liu, K., Geng, J., Li, M.: Study on detailing design of precast concrete frame structure. In: IOP Conference Series: Earth and Environmental Science, vol. 128 (2018)
2. Fan, Y.: Analysis on the development of concrete prefabricated units in China and abroad. *Concrete World* **2**, 70–75 (2013). (in Chinese)
3. Zhang, J.C.: Benefit evaluation of prefabricated concrete structure and its development in China. *Build. Technol.* **1**, 9–11 (2007). (in Chinese)
4. Zheng, Y., Zhan, B.: SLP-based layout optimization of logistics workshop facilities of Huai'an courier post. In: 3th International Conference on Transportation Information and Safety, pp. 848–851. IEEE, Wuhan, China (2015)
5. Sun, Q., Zheng, Z.H.: Discussion on the preparation of concrete prefabricated components. *Commer. Concrete* **9**, 99–100 (2012). (in Chinese)
6. Li, Z.R., Li, Q., Cao, Z.Q.: Application of SLP method in design of facilities layout in workshop. *Appl. Mech. Mater.* 28–32 (2012)
7. Lei, W., Lu, B.: Systematic layout planning and comprehensive evaluation in manufacture enterprise's logistics facilities. *Int. J. Appl. Decis. Sci. (IJADS)* **8**(4) (2015)
8. Liu, Y.N., Zhao, Q.L.: Research on logistics center layout based on SLP. *Lect. Notes Electr. Eng.* **286**, 17–28 (2015)
9. Fang, J.F., Wang, Y., Zheng, Y.Q., Lei, C.L.: Optimization design of testing center logistics on SLP. In: 2nd International Conference on Mechanic Automation and Control Engineering, pp. 7511–7514 (2011). <https://doi.org/10.1109/mace.2011.5988788>
10. Min, L.: Research on Performance Evaluation Method For Listed Company Based On Fuzzy Analytic Hierarchy Process. In: 5th International Conference on Intelligent Systems Design and Engineering Applications, pp. 837–840. IEEE, Zhengzhou, China (2014)
11. Wiyaratn, W., Watanapa, A.: Improvement plant layout using systematic layout planning (SLP) for increased productivity. *World Acad. Sci. Eng. Technol.* **72**, 373–377 (2010)

Part III

Intelligent Community and Urban Safety

Blockchain in Smart City Development—The Knowledge Governance Framework in Dynamic Alliance



Yi Zhang, Wei Sun and Chenlei Xie

Abstract Smart city, as a developing strategy for a global economy, has been initiated in many countries and presented the booming future. There are various researches suggesting that a blockchain is an indispensable tool for smart city construction, but how to adopt it to smart city strategy is not certainly solved. The study here aims to propose a scheme of employing the blockchain in smart city strategy with the theory of Knowledge Governance and Dynamic Alliance, of which the former is focusing on reducing the organization of dynamic alliance, deciding the information transmission in blockchain, and the latter is organizing the smart services by promoting the knowledge innovation and competition and quickening the smart solution with blockchain in city. The scheme is analyzed and its function is detailed at last in order to help explain how it works for guiding and planning the smart city model.

Keywords Blockchain · Smart city · Knowledge governance · Dynamic alliance

1 Introduction

“Smart city” supported by IT emerges as the latest focus and driving force of global urban development. In terms of the relationship between smart city and IT, it is apparent in what N. Komninos [1, 2] defined as the four aspects: (1) Electronic and digital techniques are applied to communities and cities; (2) IT is employed to improving the basic and important living and working condition; (3) The cities are equipped with information and communication technique; and (4) The ability of innovation, acqui-

Y. Zhang (✉) · C. Xie
Anhui Province Key Laboratory of Intelligent Building and Building Energy Saving, Anhui
Jianzhu University, Hefei, China
e-mail: 83785194@qq.com

W. Sun
School of Economics and Management, Anhui Jianzhu University, Hefei, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_13

sition, knowledge-searching, and problem-solving can be developed by combination of information and communication techniques with human in spatial exploration.

“Blockchain” technique is one of the components for realizing “smart city” which can push social fairness, credit, and communal sharing by distributed database comprising records of digital events, point-to-point protocol, consensus mechanism, and encryption algorithm. It tackles relatively credit deficiency existing on the Internet including unclear property right, and information security, by distributed data storage and protocol to decrease the third parties, trade cost, and risk. I. A. T. Hashem, V. Chand et al. studied the feasibility of “smart city” development supported by big data in which valuable information contributed to the decision-making of the cities [3]. That arouses studies of smart city and its relation to governance, giving birth to smart health care, smart home, smart transportation, smart retirement, smart education, etc. all of which have been involved with “Blockchain”.

2 Background

2.1 On Smart City

The initiation of smart city is attributed to booming urbanization and information technology, while it nurtures new urbanization by more than a decade of literature and case studies.

Conceptual Understanding

Smart city is defined usually at two perspectives: concepts based on purpose and those on means. The former stresses the change resulted from “smart”. One typical definition is given by P. Lombardi, S. Giordano, and H. Farouh who hold the idea that smart city betters the life by creating a fortune-making platform and supporting the economic growth based on modern technology and social capital [4]. The latter level centers on infrastructures construction, IT and communication, and innovation as Komninos emphasized that the smart city contains strong acquisition and creativity embodied in the digital infrastructures [1, 2].

Framework of Smart City

The framework consists of three perspectives: target, motivation, and effect. The first level covers the smart economy, ecosystem, transportation, etc. as introduced in D. Toppeta that smart city is composed of subsystems like information flow, human resources and social capital development, e-governance, etc. [5]. The second includes education, industry, government, as that L. Leydesdorff and M. Deakin relate, it is the deep interaction of education, industry, and government that knot the city net [6]. The last refers to civil service, energy use, and communication which are detailed as civil services like public administration, citizen right, commerce, transportation, water supply, and energy in the report by S. Dirks, M. Keeling, and J. Dencik [7]. To sum up, the three perspectives are all relying on the modern techniques.

Technology in Smart City

Being a new form of city equipped with IT based on IoT, cloud computing, and AI, smart city commits those techniques to smart city grid, intelligent city lamplight, intelligent transport system, and electronic government. D. Menniti, N. Sorrentino, and A. Pinnarelli studied the coordination of small-sized grids as a virtual energy in the future smart city [8]. A. Lavric, V. Popa, and S. Sfichi put forward the intelligent light control system based on the large-scale wireless sensor net [9]. Schlingensiepen, J., Mehmood, R., and Nemtanu, F.C. proposed a design framework of an auto-transmission system to enable the user access to personalized service in the intelligent surrounding [10]. In addition, smart health care, education, and governance are all linking to big data, IOT, and Blockchain in the near future, which will change the city.

Case Study in Smart City

The practice of smart city proved to be creative and to be decisive in city development, which can be found in many surveys. A. Mahizhnan analyzed the “smart island” strategy of Singapore which altered the industrial to an informational city with IT [11]. T. Gea, J. Paradells, and M. Lamarca discussed the experience of Barcelona in adopting the IOT to the smart city [12]. R. P. Dameri and C. Rosenthal-Sabroux focused on the comparison between Amsterdam and Genoa in measures and strategies in intelligent and digital cities [13], while B. Granier and H. Kudo probed into the function of ICT (Information Communications Technology) in citizen participation [14].

The study is helpful to develop smart city, especially in sharing service, just like what has been illustrated in Fig. 1 [15], but it is unable to fulfill its concept because of the limited application of big data and the deficiency in the interdisciplinary study like that of economics, sociology, or architecture. Thus, it is essential that an extensible technique, like “blockchain”, should be introduced to combine the multiple subsystems of smart city at the aspect of comprehensive administration.

2.2 The Blockchain

The blockchain technique turns to be an indispensable tool for social service due to its “sharable, programmable, and trustable” features which have been practiced in Bitcoin as well as some other fields like governance, industry, and The World Economic Forum WEF GAC15 Technological Tipping Points report 2015 [16] predicted that the 10% world economy would be stored with blockchain, as shown in Fig. 2. The Gartner made public in 2016: The annual emerging technology maturity curve (Fig. 3) in which the blockchain was presented and predicted to be mature within 15 years [17].

It is attractive to many countries like the USA and UK, who have published related government reports to guide Blockchain development. In China, the blockchain has interested industries and institutions to make efforts in deep exploitation in many areas, and some can be listed as in Table 1.

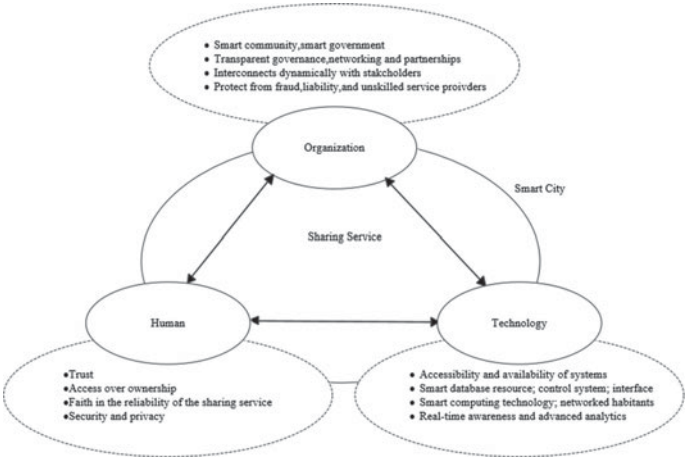


Fig. 1 A conceptual framework of smart city

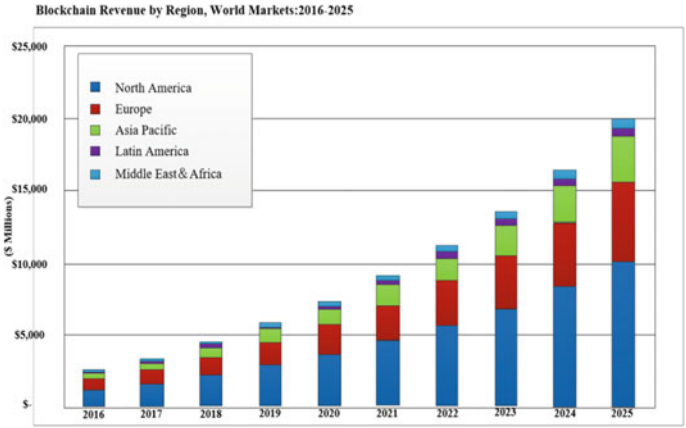


Fig. 2 Blockchain revenue by region, world market: 2016–2025

Technological Concept of Blockchain

Blockchain was named after the step of transactions that groups the distributed ledger into blocks which have to fulfill a proof-of-work condition to prove valid. Lewis Anthony explained the blockchain as the distributed database for both private and public transactions [18]. Swan referred to it as a protocol that can enlarge the IOT and help to complete its system, progress, and operation [19]. Han and Wang defined the blockchain as a technical scheme which links the users to coordinate, record, and maintain the reliable database by the means of the distributed ledger and encryption algorithm [20]. The study of blockchain has blossomed and attracted interests in various systems in which the detailed topics are stressed as follows.

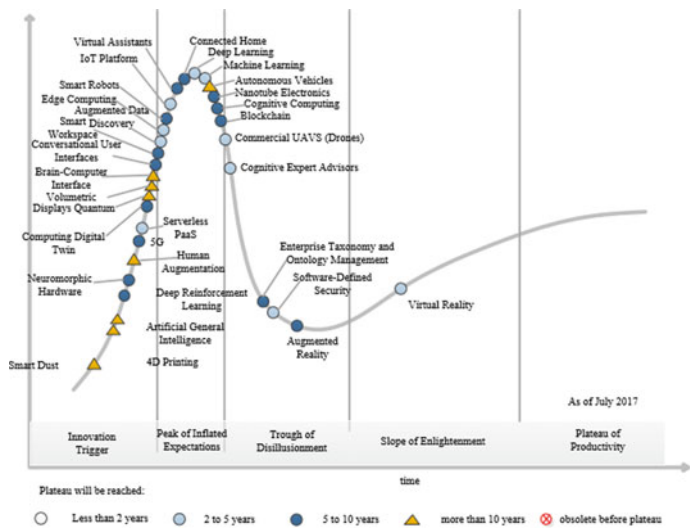


Fig. 3 The annual emerging technology maturity curve (Picture Source Tractica)

Table 1 Industries and institutions setting program with blockchain

Industry	Program with blockchain
Tencent	Public Searching Chain
	Wechat Gold
	BaaS TrustSQL
	TUSISecurity Lab
Baidu	ABS
	BaaS Blockchain Platform
JD.COM	Open Blockchain Supply-chain Quality-tracing Platform
Wal-mart	“LA Packaging” System
Meitu	Facial AI Platform
PHICOMM	NAS Daily Chain N1

Source Tencent News

On Smart Contract

Nicholas J. Szabo proposed first a digitalized contract as a promise under warranty to ensure the implement, which continues running without the third party or broker, and without any stop by personal or institutional interruption [21]. The blockchain is supplemented by smart contract in its application, while smart contract is ensured by the reliable implement surroundings of the blockchain. Jesse Yli-Huumo deepened the significance of smart contract as a solution to constructing contracts [22]. Kishigami created a digital publishing system based on distributed ledger and concept prototype system [23].

On Privacy Protection

The blockchain operates in public condition as a consensus without dependence on any unity, and the transactions are open to the public who has no chance to learn the linkage and the identity information of the transactions, which realizes both the privacy protection and fairness simultaneously. In terms of the privacy, Herrera-Joancomarti et al. introduced the literature and concluded some problems and solutions of the anonymity privacy [24]. Androulaki adopted the invisible signatures and basic protocols for both log manipulation and blockchain adjustment by only public right [25], and Ziegeldorf et al. combined the decryption hybrid net and Threshold signature to assess the prototype for safety [26].

On Network Security

The blockchain as a distributed ledger is relatively immune to attack for its storage of database in sharing copies in blocks which enable to stop the unauthorized modification or malicious tampering and keep the record permanent. Mougayar et al. [27] and Vasek et al. [28] made surveys on system vulnerability and network attack to put forward the countermeasures and their importance in the future use of blockchain.

On Extensibility

The earliest definition of extensibility is given by Swan who mentioned the application program interfaces (APIs) which is the supportive part to other service and application [19]. Now, the blockchain technique is developing toward the application to various business and social services, and the topics related to blockchain are extended to natural science and social study.

2.3 *Dynamic Alliance (DA)*

The concept “Dynamic Alliance” (DA) is a virtualized unity of some industries to seek for the market opportunity advantages together, including manufacturers, suppliers, distributors, consumer, etc. who will disband once the opportunity disappears. It is a strategic alliance for flexibility in market need and for-profit based on an open computer network and information integration.

Features and Foundation of Dynamic Alliance

As a new organization style, DA exists as the dynamic association of part or all of the resources of the members for mutual complementarity and market opportunity [29], and it is characterized by (1) being mutual trust and benefit, (2) being dynamic and temporary, (3) being complicated and interactive, (4) being demand-oriented, and (5) being systematic and coordinate.

The target of DA is to make quick response to the consumer demand and supply high-quality product, which urges the member industries to transfer its traditional external relationship and organization to a brand-new ones on the basis of the following aspects: (1) setting up a stable and reliable Internet and Intranet for information sharing, motivation, assessment, and restraint and coordination; (2) completing supply-chain management for accessible touch with all other members as

an integration, (3) Business Process Reengineering (BPR) for integration optimization of the techniques and management, and (4) advanced scientific technology and management for supporting, like Internet/Intranet and its security, and multi-agent coordination.

3 DA Knowledge Governance Based on Blockchain

The essence of DA is a united system of knowledge including its formation, accumulation, and efficient transmission for strategically supporting the operation of the DA. Thus, the knowledge and IT as resources is the core of the alliance, in which the blockchain as a newly accepted technique is the inevitable option for its shareable, programmable, and trustable features because DA needs the blockchain to utilize the knowledge at utmost and promptly in integrating resources, reacting to the market, and lowering cost.

The DA knowledge governance was initiated in IT field and was introduced to other fields like management, engineering, IA, etc. Grandori classified it as a coordinating mechanism of knowledge knot and later he analyzed the knowledge governance option based on the features of knowledge [30, 31]. Foss tried to introduce it at the level of economics, ethology, and organization and defined it as an individual subject from the industrial study [32–34]. Nickerson and Zenger studied it as a motivation for solving the organizational problems and for enhancing the efficiency and pointed out that it contains the market-, authority-, and consensus-based hierarchy which can tackle dissolvable, comprehensive, and non-dissolvable problem, respectively [35].

3.1 *Strategic Selection of Dynamic Alliance in Smart City for Cooperation*

DA enables resources to be allocated meeting group needs rather than maximum revenue of individual member, selects more advantageous resources to reconstruct the productivity in order to be more specialized, and organizes the information exchange inside to reduce the uncertainty and cost.

Evolutionary Game Model of Dynamic Alliance

The evolutionary game model of DA is used to analyze its mechanism, in which the initiator as “leader” X and the other participators as “members” Y are gaming for incomes and ventures assumed as follows:

Assumption 1 i is expected income when leader x unallied; r_i is his resources for alliance; $\lambda_i x$ is x' income coefficient in alliance, showing the multiple increase of income when allied; v_i is x' venture coefficient when allied; $r_y \lambda_i x$ is expected income of leader for alliance; and $r_x v_x$ is x' initiative cost to construct DA.

Table 2 Payoff matrix of leader and member in dynamic alliance game

Income		Member	
		Allied	Unallied
Leader	Allied	$i_x + r_y \lambda_x - r_x v_x,$ $i_y + r_x \lambda_y - r_y v_y$	$i_x - r_x v_x,$ i_y
	Unallied	$i_x, i_y - r_y v_y$	i_x, i_y

Assumption 2 i_y is expected income when member y unallied; r_y is the resources for alliance; λ_y is y's income coefficient in alliance, showing the multiple increases of income when allied; v_y is y's venture coefficient when allied; $r_x \lambda_y$ is expected income of member; and $r_y v_y$ is x' initiative cost to construct DA.

The payoff matrix is as shown in Table 2.

Assume that alliance begins, the ratio of leaders choosing to ally be $p = p(t)$, and that of members $q = q(t)$, the dynamic evolution is shown as follows.

The incomes of x when allied and unallied, respectively, are

$$EUx_p = q(i_x + r_y \lambda_x - r_x v_x) + (1 - q)(i_x - r_x v_x); EUx_{1-p} = q(i_x) + (1 - q)(i_x)$$

And the average income of x' mixed strategy is

$$EUx = pEUx_p + (1 - p)EUx_{1-p} = p[q(i_x + r_y \lambda_x - r_x v_x) + (1 - q)(i_x - r_x v_x)] + (1 - p)i_x$$

Due to mutual imitation and elimination between leaders, $p = p(t)$ evolves by mimicking dynamic equation, as shown in the dynamic differential equation;

$$\frac{dp}{dt} = p(EUx_p - EUx) = p(1 - p)(qr_y \lambda_x - r_x v_x) \quad (1)$$

Thus, that of members is

$$\frac{dq}{dt} = q(EUy_q - EUy) = q(1 - q)(pr_x \lambda_y - r_y v_y) \quad (2)$$

In Eq. (1), when $\frac{dp}{dt} = 0$, $p = 0$, $q = 1$ or $q = \frac{r_x v_x}{r_y \lambda_x}$, and the ratio of leaders choosing to ally is stable.

In (2), when $\frac{dq}{dt} = 0$, $p = 1$, $q = 0$ or $p = \frac{r_y v_y}{r_x \lambda_y}$, and the ratio of y choosing to ally is stable.

Thus, the balanced point of the evolutionary game when both x and y select alliance are (0, 0), (1, 0), (0, 1), (1, 1), (p^*, q^*) , where $p^* = \frac{r_y v_y}{r_x \lambda_y}$, $q^* = \frac{r_x v_x}{r_y \lambda_x}$.

Evolution of Dynamic Alliance

Based on analysis Friedman tool, the evolutionary stability of dynamic alliance of x and y can be analyzed by the Jacobian matrix, where

Table 3 Partial stability analyzed with Jacobian matrix

Balanced points	$ J $	$tr(J)$
$p = 0, q = 0$	$r_x v_x \cdot r_y v_y$	$-r_x v_x - r_y v_y$
$p = 1, q = 0$	$r_x v_x (r_x \lambda_y - r_y v_y)$	$r_x v_x + r_x \lambda_y - r_y v_y$
$p = 0, q = 1$	$r_y v_y (r_y \lambda_x - r_x v_x)$	$r_y v_y + r_y \lambda_x - r_x v_x$
$p = 1, q = 1$	$(r_y \lambda_x - r_x v_x)(r_x \lambda_y - r_y v_y)$	$r_x v_x - r_y \lambda_x + r_y v_y - r_x \lambda_y$
$p = \frac{r_y v_y}{r_x \lambda_y}, q = \frac{r_x v_x}{r_y \lambda_x}$	$(r_y \lambda_x - r_x v_x)(r_x \lambda_y - r_y v_y) \cdot \frac{v_x v_y}{\lambda_x \lambda_y}$	0

Jacobian matrix of Eqs. (1) and (2) is

$$J = \begin{pmatrix} (1-2p)(qr_y \lambda_x - r_x v_x) & pr_y \lambda_x (1-p) \\ qr_x \lambda_y (1-q) & (1-2q)(pr_x \lambda_y - r_y v_y) \end{pmatrix}$$

then $tr(J) = (1-2p)(qr_y \lambda_x - r_x v_x) + (1-2q)(pr_x \lambda_y - r_y v_y)$.

The five balanced points mentioned above are defined in Table 3.

Then, four propositions are concluded:

Proposition 1 When $r_y \lambda_x > r_x v_x$, $r_x \lambda_y > r_y v_y$, $(0, 0)$ and $(1, 1)$ are evolutionary stable strategy (ESS).

Proof When $r_y \lambda_x > r_x v_x$, $r_x \lambda_y > r_y v_y$, $|J| > 0$ at $(0, 0)$ and $(1, 1)$, and $tr(J) < 0$, thus, both points are ESS. At (p^*, q^*) , $|J| > 0$, and $tr(J) = 0$ is saddle point, while at $(1, 0)$ and $(0, 1)$, $|J| > 0$, and $tr(J) > 0$ is unstable balanced points.

Proposition 1 tells if the expected revenues of x and y in alliance are, respectively, higher than their costs, $(0, 0)$ and $(1, 1)$ are stable points, as shown in Fig. 4a.

Proposition 2 When $r_y \lambda_x > r_x v_x$, $r_x \lambda_y > r_y v_y$, the stable strategy is $(0, 0)$.

Proof When $r_y \lambda_x < r_x v_x$, $r_x \lambda_y > r_y v_y$, $|J| > 0$ at $(0, 0)$, and $tr(J) < 0$, which is ESS.

And $|J| < 0$ at $(1, 0)$ and $(1, 1)$, with uncertain $tr(J)$, the two are saddle points; while $|J| > 0$ at $(0, 1)$ as unstable balanced one with $tr(J) > 0$.

Proposition 2 defines x with lower expected revenue than cost and y with a higher one than cost will select (unallied, unallied) as ESS shown in Fig. 4b. Both sides' unallied strategy leads to convergence to $(0, 0)$

Proposition 3 When $r_y \lambda_x > r_x v_x$, $r_x \lambda_y < r_y v_y$, evolution stability point is $(0, 0)$.

Proof When $r_y \lambda_x > r_x v_x$, $r_x \lambda_y < r_y v_y$, $|J| > 0$ at $(0, 0)$, and $tr(J) < 0$, and $(0, 0)$ is an ESS. $|J| < 0$ at both saddle points $(0, 1)$ and $(1, 1)$ with uncertain $tr(J)$, while $|J| > 0$ at $(1, 0)$, and $|J| > 0$ is an unstable balanced point.

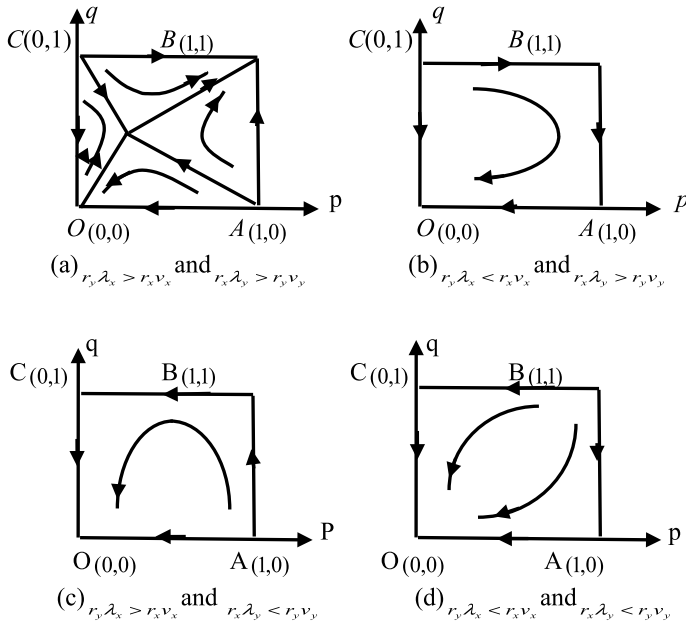


Fig. 4 Evolution trajectory of game of leader and member

Proposition 3 indicates x with higher expected revenue than cost and y with lower one chosen (unallied, unallied) as ESS shown in Fig. 4c, where the system converges to $(0, 0)$.

Proposition 4 When $r_y \lambda_x < r_x v_x$, $r_x \lambda_y < r_y v_y$, ESS is (unallied, unallied).

Proof When $r_y \lambda_x < r_x v_x$, $r_x \lambda_y < r_y v_y$, $|J| > 0$ at $(0, 0)$, and $tr(J) < 0$; thus, $(0, 0)$ is an ESS. At saddle points $(0, 1)$ and $(1, 0)$, $|J| < 0$ with uncertain $tr(J)$, but $|J| > 0$ at $(1, 1)$ and $tr(J) > 0$ is unstable balanced point.

Proposition 4 presents x and y both predicting lower revenue than cost, respectively, will refuse an alliance for stability, as shown in Fig. 4d that it converges to $(0, 0)$.

The analysis of strategic selection and evolution shows that the long-term equilibrium is influenced by payoff parameter, and the strategy is to be evolved to (allied, allied) by parameter adjustments when increasing revenue coefficient, lowering risk, promoting resources sharing, etc.

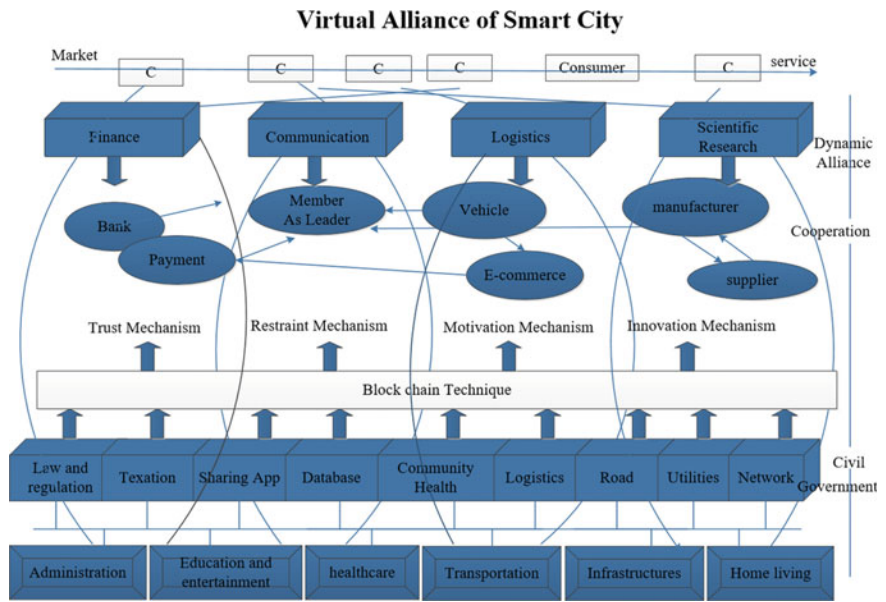


Fig. 5 Knowledge governance in dynamic alliance with blockchain

3.2 The Knowledge Governance in Dynamic Alliance Based on Blockchain

The production of knowledge is a must of DA in smart city and the governance of knowledge is a supporting system in shaping and affecting the motives and behaviors of the knowledge users and beneficiary in order to promote the behavior and avoid risk. Asan institutional guarantee, knowledge governance can guide, motivate, restrain the behavior, and balance the participators' benefit.

Governance of knowledge of DA covers five steps: knowledge seeking, sharing, integrating, innovating, and assessing. Knowledge seeking refers to the course that the leader and members try to find other member and its related knowledge which is affected by openness and supplements. Sharing is driven by credit risk and incentives, in which the higher the shareability is, the less the risk is. Integrating is to collect knowledge and abstract it. Innovating is to rebuild core competitiveness, and assessing is to improve the management and quality. Their relation is illustrated in Fig. 5.

In Fig. 5, the blockchain technique supports the five-step procedure by its features like “sharable, trustable, extensible, secure, and quick access” with the supply of trust, sharing, incentive, supervision, or extensive mechanism as depicted in Fig. 5. The knowledge governance in DA nurtures the blockchain with the innovation and better service in smart city as interpreted in the next part.

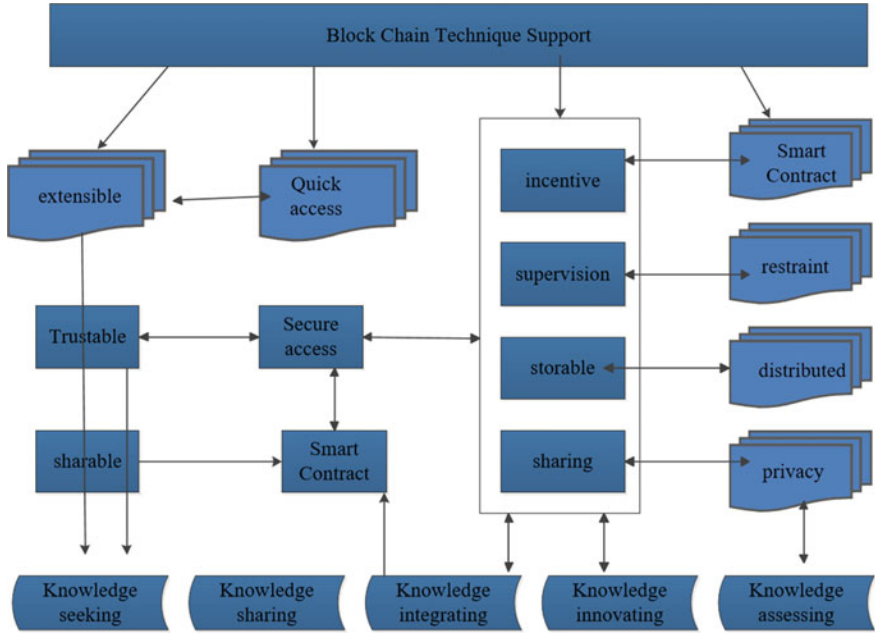


Fig. 6 Organization of dynamic alliance of smart city-based blockchain

3.3 The Organization of DA of Smart City Based on Blockchain

The operation of DA has to be adjusted to the market and the development which at present is symbolized by IT, especially the blockchain technique, enabling the smart city function newer and better, and enabling the dynamic cooperation with multiple units accessible. In order to better understand their working flow, an organizational scheme of integrated resources is presented as shown in Fig. 6.

Smart city is based on advanced information technology and open innovation ecology, which can bridge the gap between supplier and consumer of information, between innovation and social activities to quicken management and operation of city and industry activity.

In Fig. 6, social activities are analyzed at two levels, of which the lower is the government and the upper is the market, bridged by blockchain mutually.

At the level of city government, smart civil service covers administration, education, and others as shown in Fig. 6. Departments subordinate to the government like that of law-and-regulation-making, taxation, sharing-application-operation, or others are carrying out the activities by horizontal dynamic alliance with the help of knowledge governance in blockchain functioning by coordinating the different activities as follows:

- (1) Offering policy instruction and optimizing the service structure. With the support of blockchain, the government makes full use of the platform to share and extend the information safely and quickly, and organizes the dynamic alliance to allow the smart city service broader and better.
- (2) Specializing the service and flattening the organization. The dynamic alliance can help to quick access to information contributes to lessening the participation of more organization and labor force, focusing the service as the core of a single organization with the help of blockchain.
- (3) In civil service, the dynamic alliance concentrates the service on the punctual service and development which can lower the service cost according to the theory of "experience curve effect."
- (4) Effectively coordinating the visible and invisible capital. The alliance can help to share the tacit knowledge which should be protected by the smart contract to allocate rationally the resource both visible and invisible.
- (5) Mitigating the risk inside the alliance by sharing the knowledge which increases the ability to survive the risk. With the supervision and guidance of government and alliance, the trial and error can be lessened and the city core competitiveness can be strengthened.

At the level of the market, the enterprises in the alliance follow the market mechanism aiming at the maximizing the profit while cooperating with other members guided by shared knowledge. In Fig. 6, only some enterprises are presented as examples including finance, communication, logistics, and research industries in which the communication is supposed as the leader while others as members. They form the dynamic alliance in the order of seeking, sharing, integrating, creating, and assessing the information knowledge, and make production decision based on the blockchain storage and based on the civil service alliance. The result is to be the following.

- (1) The blockchain helps to lower the risk in allying the enterprises. It is the sharable, reliable, and symmetrical system in blockchain that decreases the risk of enterprises in allying who may suffer inevitable risks of asymmetric information and unstable cooperation when unallied.
- (2) The core competitiveness of enterprises is reinforced based on innovation under the smart contract. The enterprises have to innovate, raise the core competitiveness, and complement each other's advantages, share the risk and benefit following the contract, in order to avoid the malicious competition between members at the same level.
- (3) Allocating the resources and distributing the profit reasonably. The dynamic alliance is a typical cooperative game, and the total profit of alliance should be higher than the sum of that of individual enterprise. The fairness in distribution is important, and it can promote the alliance development.
- (4) Saving the transaction cost. The shareable information and technique save the transaction cost by avoiding the market opportunism which leads to the increase in cost to make a contract because of the rent-seeking or cheating.

Between the market and the government, there are knowledge governance mechanisms of the trust, restraint, motivation, innovation implemented by blockchain

bridging the upper and the base to offer the necessary service, save the cost, and unite the opponents of the smart city. The blockchain is important in bridging as well as supervision, motivation, restraint, assessment, and penalty as detailed below:

- (1) The distributed ledger offers public and traceable knowledge which can avoid the cheating and forgery in finance, insurance, ID certificate, Internet education, or secondary market.
- (2) The broker and the third party can be stripped away from the transaction in order to decrease the cost and avoid the “rent-seeking,” which can help government purchase, education, entertainments, and sales to lessen the intermediate cost and benefit the consumers.
- (3) The information can be recorded by the distributed ledger to be kept permanently and safely. The security of privacy, especially the intellectual property right, is protected.
- (4) The smart contract contributed to blockchain security by offering contractive storage, transmission, usage, and integration of knowledge to emphasize the responsibility and duty as an effective tool.

4 Conclusions

Smart city is the inevitable trend of the city development, which has been on the way to success with the help of scientific technology. Blockchain as a new technique presents the accessibility, creditability, extensibility, and security in the realization of smart city, especially when the dynamic alliance is formed based on the blockchain. Blockchain is the medium of different smart services and it is a collector and editor of different information and knowledge which are filtered by the processing of knowledge seeking, sharing, integrating, innovating, and assessing, all of which are covered by knowledge governance helping the DA leader and members to find, share, integrate other members and related knowledge, and then innovate all to improve the management and quality in smart services.

Focuses on the proposing a framework of knowledge governance in the dynamic alliance of smart city based on blockchain and scheming the knowledge governance course, the purpose of the study is helping the government and industries to make full use of the IT and other technology in developing the civil service related. However, as a basic protocol, computation and verification with the computers are energy-consuming, but we cannot reject innovation because of its flaw, and more researchers have been stimulated to make efforts to lower the resources consumption as progress has been made for future solution to the problem because the development of blockchain needs introduction and application like all the other new techniques that the newer is, the brighter its future is.

References

1. Komninos, N.: *Intelligent Cities: Innovation, Knowledge Systems and Digital Spaces*, pp. 1–15. Spon Press, London (2002)
2. Komninos, N.: The architecture of intelligent cities: integrating human, collective, and artificial intelligence to enhance knowledge and innovation. In: *Proceedings of the 2nd International Conference on Intelligent Environments*, pp.13–20 (2006)
3. Hashem, I.A.T., Chand, V., Anuar, N.B., et al.: The role of big data in smart city. *Int. J. Inf. Manage.* **36**(5), 748–758 (2016)
4. Lombardi, P., Giordano, S., Farouh, H., et al.: Modeling the smart city performance. *Innov.-Eur. J. Soc. Sci. Res.* **25**(2), 137–149 (2012)
5. Toppeta, D.: *The Smart City Vision: How Innovation and ICT Can Build Smart, “Livable”, Sustainable Cities*. The Innovation Knowledge Foundation (2011)
6. Leydesdorff, L., Deskin, M.: The triple-helix model of smart cities: a neo-evolutionary perspective. *J. Urban Technol.* **18**(2), 53–63 (2011)
7. Dirks, S., Keeling, M., Dencik, J.: How smart is your city?: helping cities measure progress [20160806]. http://www.ibm.com/smarterplanet/global/files/uk_en_uk_cities_ibm_sp_pov_smartcity.pdf (2018)
8. Menniti, D., Sorrentino, N., Pinnarelli, A., et al.: In the future smart cities: coordination of micro smart grids in a virtual energy district. In: *Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion*, pp. 676–681(2014)
9. Lavric, A., Poppa, V., Sfichi, S.: Street lighting control system based on large-scale WSN: a step towards a smart city. In: *Proceedings of 8th International Conference and Exposition on Electrical and Power Engineering (EPE)*, pp. 673–676(2014)
10. Schlingensiepen, J., Mehmood, R., Nentanu, F.C.: Framework for an autonomic transport system in smart cities. *Cybern. Inf. Technol.* **15**(5), 50–62 (2015)
11. Mahizhnan: Smart cities: the Singapore case. *Cities* **16**(1), 13–18 (1999)
12. Ges, T., Paradells, J., Lamarca, M., et al.: Smart cities as an application of internet of things: experiences and lessons learnt in Barcelona. In: *Proceedings of the IEEE Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, pp. 552–557 (2013)
13. Dameri, R.P., Rosenthal-Sabroux, C.: Governing smart cities: an empirical analysis. *Soc. Sci. Comput. Rev.* **34**(6), 693–707 (2016). [sagepub.com/journalsPermissions.nav. https://doi.org/10.1177/0894439315611093](https://doi.org/10.1177/0894439315611093)
14. Granier, B., Kudo, H.: How are citizens involved in smart cities? analysing citizen participation in Japanese “Smart Communities”. *Inf. Policy* **21**(1), 61–76 (2016)
15. Sun, J.J., Yan, J.Q., Zhang, K.Z.K.: Blockchain-based sharing services: What blockchain technology can contribute to smart cities. *J. Financ. Innov.* **2**, 26 (2016)
16. The World Economic Forum WEF GAC15 Technological Tipping Points Report 2015. <https://www.weforum.org/reports/deep-shift-technology-tippingpoints-and-societal-impact>
17. Furlonger, D., Valdes, R.: Practical Blockchain: A Gartner Trend Insight Report 2017. <https://www.gartner.com/doc/3628617?ref=SiteSearch&sthkw=&fnl=search&srcId=1-3478922254>
18. Lewis, A.: A gentle introduction to blockchain technology. Bits on blocks <https://bitsonblocks.net/2015/09/09/a-gentle-introduction-to-blockchain-technology>. Accessed 10 Apr 2018
19. Swan, M.: Blockchain thinking: the brain as a DAC (Decentralized Autonomous Corporation). *IEEE Technol. Soc.* **34**(4), 41–52 (2015)
20. Han, Q.M., Wang, G.A.: Review of foreign research of blockchain technology. *Sci. Technol. Prog. Policy* **35**(1), 154–160 (2018)
21. Szabo, N.J.: Smartcontracts. <http://wuh.com/download/WECSmartContracts.pdf> (2017)
22. Yli-Hummo, J., Deokyo, K.O., et al.: Where is current research on blockchain technology—a systematic review. *PLoS ONE* **11**(10), e0163477 (2018). <https://doi.org/10.1371/journal.pone>
23. Kishigami, J., Fujimure, S., et al.: The blockchain based digital content distribution system. In: *IEEE Fifth International Conference Big Data and Cloud Computing 2015*, pp. 187–190 (2015)

24. Herrera-Joancomartí, J.: Research and challenges on Bitcoin anonymity. In: Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance. LNCS, vol. 8872, pp. 3–16. Springer, Cham (2018)
25. Androulaki, E., Karame, G.: Hiding Transaction Amounts and Balances in Bitcoin (2018). <https://PDFS.Semanticscholar.org/da9a/85e5d7b7bd9f43255656558d40989ac23a43>
26. Ziegeldorf, J.H., Grossmann, F., et al.: Secure multiparty mixing of bitcoins. <https://www.martinhenze.de/wpcontent/papercitedata/pdf/zgh+15.pdf> (2014)
27. Mougayar, W.: Why fragmentation threatens the promise of the blockchain. <http://www.coindesk.com/fragment-blockchain-identity-market/>. Accessed 2017
28. Vasek, M., Moore, T.: There's no free lunch, even using bitcoin: tracking the popularity and profits of virtual currency scams. In: International Conference on Financial Cryptography and Data Security. FC 2015: Financial Cryptography and Data Security, pp. 44–61. https://link.springer.com/chapter/10.1007/978-3-662-47854-7_4. Accessed 10 Apr 2018
29. Roger, N.N.: The 21th century manufacturing enterprise strategy report. AD-A257167.3: 1–4(1991)
30. Grandori, A.: Governance structures, coordination mechanisms and cognitive models. *J. Manage. Gov.* **1**, 29–47 (1997)
31. Grandori, A.: Neither hierarchy nor identity: knowledge-governance mechanisms and the theory of the firm. *J. Manage. Gov.* **5**, 381–399 (2001)
32. Foss, N.J.: The knowledge governance approach. SMG Working Paper, No. 1/2005. Center for Strategic Management and Globalization, Copenhagen Business School (2005)
33. Foss, N.J., Husted, K., Michailova, S., Pedersen, T.: Governing knowledge processes: theoretical foundations and research opportunities. CKG Working Paper, No. 1/2003. Center for Knowledge Governance, Copenhagen Business School (2003)
34. Foss, N.J.: The emerging knowledge governance approach: challenges and characteristics. Druid Working Paper, No. 06-10. Danish Research Unit for Industrial Dynamics (2006)
35. Nickerson, J.A., Zenger, T.R.: A knowledge-based theory of the firm: the problem-solving perspective. *Organ. Sci.* **15**(6), 617–632 (2004)

Attendance and Security System Based on Building Video Surveillance



Kailai Sun, Qianchuan Zhao, Jianhong Zou and Xiaoteng Ma

Abstract The attendance system plays a very important role in the modern enterprise's operation, and the security of the building has always been a matter of concern to the people. Based on networked surveillance video, this paper integrates the attendance and security functions and fuses video image processing, deep learning, and face recognition to design an intelligent attendance and security system. We propose a sliding average method to identify persons' identities. The experimental results verify the effectiveness of our method. The false reject rate (FRR) in our system reaches 0.51%, the false accept rate (FAR) reaches 2.52%, and the correct identification rate reaches 98.85%. The system is applied to some video surveillance areas, with advantages of nonintrusive, passive attendance and multiple persons' attendance at the same time.

Keywords Networked video surveillance · Deep learning · Face recognition · Convolutional neural network · Attendance and security system

1 Introduction

The attendance system is an indispensable system for the modern office units and companies. The most frequently used attendance system is the IC card system, the fingerprint attendance system [1], and the face attendance system. IC card system is the most backward mechanical attendance. The phenomenon of loss of cards, forgetting and substituting for others can be easily caused. Fingerprint system and face system need to identify specific employees' fingerprints or faces for attendance. Although they have eliminated the phenomenon of substituting for others, they need

K. Sun · Q. Zhao (✉) · J. Zou · X. Ma

Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Center for Intelligent and Networked Systems (CFINS), Tsinghua University, Beijing 100084, China
e-mail: zhaoqc@tsinghua.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_14

153

employees to come to designated locations to perform specified actions. This is not natural, and it is inefficient and unacceptable in many situations.

In addition, security issues in corporate companies and residential areas are very crucial. Since 911 events, international demand for security precautions in public areas has skyrocketed. For security problems such as burglaries and terrorist attacks, effective solutions have not been proposed. With the continuous advancement of electronic information technology, artificial caveat will gradually be replaced by intelligent video surveillance system [2].

Networked video surveillance technology uses an IP camera to capture surveillance areas and transmit data over the Internet, enabling remote surveillance. However, the IP camera is only a passive monitoring and evidence. People need to view the monitoring screen to recognize the occurrence of an incident. Only after a security incident has occurred, surveillance video verification will be investigated, and post-event analysis will be performed, but it will not be used to warn early.

In the area of face recognition, Facebook proposed DeepFace [3], a face recognition network based on deep learning, in 2014. Face recognition was performed by using deep convolutional neural networks and large-scale face image datasets. It achieved 97.35% on LFW dataset. The performance is comparable to artificial recognition. The VGGNet [4] network achieves 98.95% accuracy with a deep network structure and a large data of images. In 2015, Google proposed FaceNet [5], which achieved an accuracy of 99.63%. It uses the triple loss function and adds an embedded layer extraction feature in the network. Due to the low image quality, different angles, various poses, and distances between face and camera in actual scenes, we need a large amount of data with the same identity to train and learn, and we also need to adjust the model to obtain better results.

2 Approach

In this section, we will introduce the structure of our system and the principle of the algorithm, including face detection, image preprocessing, face recognition, and multi-frame identification based on sliding average (Fig. 1).



Fig. 1 System structure

2.1 Face Detection

Face detection is very challenging in an unconstrained environment due to various poses, illuminations, and occlusions. Universal object detection based on deep learning has developed rapidly, such as Faster R-CNN [6], SSD [7], and YOLO V3 [8]. This paper uses the MTCNN [9] algorithm which is widely used for human face detection. Like YOLO, it treats the detection problem as a regression problem and uses the convolution network to train optimize classification and object locations. He first built the image pyramid, and then designed three CNN models: P-Net, R-Net, and O-Net to perform cascading prediction to improve the accuracy, defined the label of the image, and designed cross-entropy loss and Euclidean distance loss function. The position of the face and the five key points of the face are obtained by training (Fig. 2).

2.2 Image Preprocessing

Image processing is the intermediate link between face detection and face recognition. Therefore, the appropriate processing methods need to be selected according to the actual scene. We cut out the detected face image and normalize it to get the same size. We use the bilinear interpolation method to zoom in and out of the image. Then, we normalize the image pixels, so that the pixel values of the image are scaled linearly from 0 to 1 in proportion.



Fig. 2 The results by MTCNN

2.3 Face Recognition

Typical CNN models include AlexNet, VGGNet, and GoogleNet. They all achieved outstanding results in the ILSVRC competition. The AlexNet structure has five convolutional layers and three fully connected layers. The performance of the network have been improved with the Inception of GoogleNet [10] and the improvements of V2, V3, and V4. For the face classification and identification problem, based on AlexNet, GoogleNet, and FaceNet, we have comprehensively improved and designed a convolutional network structure and achieved excellent experimental results on our own dataset.

As shown in Table 1, for the classification layer, we use the softmax classification, but softmax is an exponential mapping, which leads to explosive growth of the output results for the introduction of larger parameters, and the overflow of the values leads to uneven probability distributions and results in poor experimental results. We add ZM norm (zero-max normalize) on the result of the full-connection layer output and achieve better results. The ZM norm layer is shown in Eq. (1), and the softmax layer is shown in Eq. (2).

$$a_i = \begin{cases} 0 & , x_i \leq 0 \\ \beta \times \frac{x_i}{\max(x_i)} & , x_i > 0 \end{cases} , \forall i \in 1 \dots C \tag{1}$$

Table 1 CNN structure

Layer	Size-in	Size-out	Kernel
conv1	100×100×1	100 × 100 × 16	1 × 1 × 16, 1
maxpool1	100 × 100 × 16	50 × 50 × 16	3 × 3 × 16, 2
conv2	50 × 50 × 16	50 × 50 × 16	3 × 3 × 16, 1
maxpool2	50 × 50 × 16	25 × 25 × 32	3 × 3 × 32, 2
Inception3	25 × 25 × 32	25×25×128	6
Inception4	25×25×128	25×25×256	8
Inception5	25×25×256	12×12×512	5
Inception6	12×12×512	6 × 6 × 640	6
fc1	6 × 6 × 640	2048	
fc2	2048	256	
fc3	256	14	
ZM Norm	14	14	

$$y_i = \frac{e^{a_i}}{\sum_{k=1}^C e^{a_k}}, \forall i \in 1 \dots C \quad (2)$$

where x_i is the output of the FC layer, where C represents the number of output categories and β is the normalized coefficient (we use $\beta = 10$), and y_i represents the output of softmax.

2.4 Multi-frame Identification Based on Sliding Average

For most recognition system, the method of improving accuracy is mostly to improve convolutional networks. However, for the surveillance video, if we only consider the recognition results of a single frame, set the corresponding threshold (above 80% confidence level, we think its identification is accurate), it is very likely that visitors will be mistaken for registered persons, resulting in errors. For this reason, we analyzed in detail the probability distribution of the confidence after recognition. For visitors, their average recognition accuracy in a video clip is very low, and its variance is large, which means that the fluctuation is obvious, but the registered persons have a more stable confidence, and the average confidence is also very high. Therefore, we design a method based on the sliding average to analyze the recognition results within the video. Our algorithm is as follows:

- (1) In the first step, an $M \times C$ matrix queue is created for each surveillance video. $(C - 1)$ is the number of the registered persons. We use $M = 10$.
- (2) In the second step, during the processing of each frame, for each face image, softmax will output a confidence vector of the category recognition $Y = [y_1, y_2 \dots y_C]$, and we will binarize it. In the process, the maximum confidence category result is set to 1 and the remaining categories are 0, so we have a binary vector:

$$\begin{cases} a = [a_1, a_2 \dots a_C], a_i \in \{0, 1\} \\ \sum_{i=1}^C a_i = 1 \end{cases} \quad (3)$$

For the case where multiple faces appear at the same time, we will sum the a vectors of each face to obtain a binary vector A .

- (3) The third step is to set the existing matrix queue as $S = [S_1, S_2 \dots S_C]$. We use A to update S , as shown in Eq. (4).

$$S = \eta \times A + (1 - \eta) \times S \quad (4)$$

while η is the update rate, which means that A weighs the final decision result, $0 \leq \eta \leq 1$, and $\eta = 0.2$ is set here.

- (4) In the fourth step, we think that there are no pedestrians in the frame where the face area is not detected. We also need an average S . Let A be a zero vector and set $\eta = 0.1$ to update S .
- (5) In the fifth step, based on this, we can set the threshold value of S to control the result. If $S_i > 0.4$ here, then the category i can be considered as one of the final results.

3 Experiment

3.1 Dataset

Considering that the attendance security system in this paper needs to be applied to the actual scene, we must collect the monitoring images and videos under the actual scene. This paper uses the IPC (IP Camera) in the Information Science and Technology Building to collect video data during a week (8:00–17:00 every day). Thirteen pedestrians in the lab are selected to pass the IP Camera and stay for many minutes.

Because the pedestrians in the building under the actual monitoring scene pay more attention to the front figures and labs, it is difficult for the camera to collect standard frontal images; thus, we collect some top faces and side faces. Besides, the quality of the surveillance camera is not high, the resolution is 1280×720 , and the face image is smaller, less than 100×100 . This paper performs face detection of the collected video by frame, get a total of 15,419 photos of 100×100 gray face images, including 13 registered persons, about 1000 photos per person, and the rest are negative samples.

3.2 Platform

This experiment is based on python3.6; the operating environment is Win10 $\times$ 64, Intel (R) Core (TM) i5-6300HQ@2.30GHZ, NVIDIA GTX 960 M; and computer memory is 8 GB.

3.3 Experimental Results

As shown in Fig. 3, our system is able to produce some boxes to locate persons' faces. Besides, through the face recognition and sliding average method, the text of every face classification and confidence will be drawn in the face boxes.



Fig. 3 The experimental results

3.4 Performance and Comparison

FRR and FAR. The false reject rate (FRR) is the probability of misidentifying registered persons as unregistered persons, while the false accept rate (FAR) is the probability of misidentifying unregistered persons as registered persons as shown in (5) and (6).

$$FRR = \frac{N_{reg}^{non}}{N_{reg}} \tag{5}$$

$$FAR = \frac{N_{non}^{reg}}{N_{non}} \tag{6}$$

where N_{reg} indicates the total number of sampled registered person images, N_{reg}^{non} indicates that the number of the registered person images is erroneously detected as unregistered, N_{non} indicates the total number of unregistered person images, and N_{non}^{reg} indicates that the number of unregistered person images is erroneously detected as registered (Table 2).

The correct identification rate. The above indicators are only applied to the security situation. As for attendance, we need to classify and identify each of the registered persons.

$$CIR = \frac{N_{CIR}}{N_{reg}} \tag{7}$$

Table 2 Evaluation on FRR and FAR

Sliding average	FRR (%)	FAR (%)
No	2.66	6.59
Yes	0.51	2.52

Table 3 Evaluation on *CIR*

Sliding average	CIR (%)
No	91.77
Yes	98.85

Table 4 Runtime on our system

Modules	No one (ms)	One people (ms)	Two people (ms)
MTCNN	157.26	165.29	189.45
Image processing	0.00	3.01	5.01
CNN Recognition	0.00	10.04	14.04
Total time	157.26	178.34	208.50

In (7), *CIR* represents the identification accuracy rate of registered person, N_{reg} represents the total number of samples of registered persons, and N_{CIR} represents the total number of samples correctly identified. Obviously, we expect our system has the higher accuracy of recognition. The results are shown in Table 3.

Runtime efficiency. Runtime efficiency is an important indicator. Only the security and attendance system with good real-time performance can be used in practice. Therefore, the response time of each module is counted in the video, as shown in Table 4.

In Table 4, it can be seen that with the increase in the number of detections, the time-consuming of the system also increases, while the method of multi-frame identification based on sliding average need no extra runtime in our experiments. Actually, each detection and recognition of the system in this paper is between 0.2 and 0.3 s, basically meeting the real-time response of the system.

4 Discussion

For the improvement of recognition accuracy, this paper believes that this is mainly due to the design of the sliding average method. For the recognition results of each frame in the video, if we just set the similarity threshold, then we only identify based on a single frame. Moreover, setting a high similarity threshold will lead to a long waiting time for people to sign in, while setting a similarity threshold to a lower value will lead to incorrect identification. This is unfair to attendance, so the setting of a threshold value brings a difficult challenge. This paper designs the $M \times C$ queue matrix, stores the discriminant results, updates the queue matrix in real time, considers each frame’s results over a period of time, and finally obtains the similarities of multiple frames in real time, and then compares and discriminates them. The advantage of this approach is to consider the recognition results of the

previous frames and to filter out the results with lower frequencies and recognition probabilities, thereby enhancing the results with higher frequencies and recognition probabilities and obtaining more accurate identification. Besides, this sliding average method also could apply to other face recognition methods based on video.

This paper also investigates the performance indicators of various attendance systems, and we compare and analyze them with the system (Table 5).

In terms of applications, the system is noncontact attendance, more hygienic, and better in real time, because it can perform multiple attendances at the same time to ensure the timeliness, avoiding the phenomenon of queued attendance. The system is monitored under the condition of being able to sign attendance without notice. Because this is a passive type of attendance, compared to active attendance such as fingerprints, it is very friendly to users and does not require to worry about forgetting attendance problems; and the system has a security function that can effectively prevent illegal entry of outsiders, reducing the number of human caretakers.

5 Conclusions

With the implementation of China's "Shiny Project", surveillance video cameras have been widely used. Attendance and security are one of the typical applications of surveillance video. This paper has designed a video surveillance system with security and attendance functions. It mainly includes summarizing and analyzing the development of face recognition technology at the present stage, implementing MCTNN algorithm for face detection. This paper combines the network structure design of AlexNet and Inception to construct a convolutional neural network for face recognition and classification, proposes a method based on sliding average to improve the accuracy of recognition, and builds face dataset in actual scenes, verifying the effectiveness of the algorithm through experiments.

Table 5 The performance of various attendance systems

Systems	Runtime(s)	FRR (%)	FAR (%)	CIR (%)
Fingerprint	1–2	0.6	0.4	99.6
IC card	<0.1	0.01	0.0001	100
NFC card	–	0.1	0	99.89
Face verification	1–2	2.04	1.94	97.96
Monitor	1.04	10.64	12.20	89.36
Our system	0.2–0.3	0.51	2.52	98.85

Acknowledgements, Ethic Approval, and Informed Consent This work was supported by National Key Research and Development Project of China (No. 2016YFB0901900 and 2017YFC0704100 entitled new generation intelligent building platform techniques) and the National Natural Science Foundation of China (No. 61425027). This work was also supported by the National 111 International Collaboration Project (B06002, BP2018006) and Special Fund of Suzhou-Tsinghua Innovation Leading Action (Project Number: 2016SZ0202).

We appreciate constructive comments from the anonymous reviewers.

This study has been conducted with ethics approval obtained from Institutional Review Board of Tsinghua University, Beijing, China.

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

References

1. Kamaraju, M., Kumar, P.A.: Wireless fingerprint attendance management system. In: IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), pp. 1–6. IEEE, Coimbatore, India (2015)
2. Banu, V.C., Cosea, I.M., Nemtanu, F.C., et al.: Intelligent video surveillance system. In: IEEE 23rd International Symposium for Design and Technology in Electronic Packaging (SIITME), pp. 208–212. IEEE, Constanta, Romania (2017)
3. Taigman, Y., Yang, M., Ranzato, M.A., et al.: DeepFace: closing the gap to human-level performance in face verification. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1701–1708. IEEE, Columbus, USA (2014)
4. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *Computer Science* (2015). [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
5. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: a unified embedding for face recognition and clustering. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815–823. IEEE, Boston, MA, USA (2015)
6. Ren, S., He, K., Girshick, R., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149 (2017)
7. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: single shot multibox detector. In: European Conference on Computer Vision, pp. 21–37. Springer International Publishing (2016)
8. Joseph, R., Ali, F., et al.: YOLOv3: an incremental improvement. *Computer Science* (2018). [arXiv:1804.02767](https://arxiv.org/abs/1804.02767)
9. Zhang, K., Zhang, Z., Li, Z., et al.: Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett. (SPL)* **23**(10), 1499–1503 (2016)
10. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9. IEEE, Boston, MA, USA (2015)

Simulation Study on Collaborative Evacuation Among Stairs and Elevators in High-Rise Building



Xiaodong Liu, Hui Zhang and Ping Zhang

Abstract In order to improve the efficiency of occupant evacuation in high-rise and intelligent buildings, this paper analyzes the disadvantages of traditional emergency evacuation in high-rise buildings and conducts the feasibility analysis of using an elevator as the evacuation tool. With the help of evacuation software, Pathfinder, simulation experiments are given to study the collaborative strategies among stairs and elevators under partial and total building evacuation. Results show that with the guarantee of safety and reliability, an elevator is a very efficient tool to evacuate occupants in high-rise buildings, especially those mobility impaired occupants. Combination of stairs and elevator can not only reduce the evacuation time but also shorten the occupants' travel distance.

Keywords High-rise building · Intelligent evacuation · Elevator and stairs

1 Introduction

A growing number of high-rise buildings are appearing in the world, especially in China. By 2017, there are 347,000 high-rise buildings and more than 6,000 high-rise buildings above 100 meters in China, which ranks first in the world. High-rise buildings own the characteristics of huge volume, complicated functions, dense, many dangerous sources, and large fire loads, which face enormous safety challenges. In the past 10 years, there have been 31,000 high-rise building fires in China, with the death of 474 people and 1.56 billion in direct property losses. Under emergencies,

X. Liu · H. Zhang (✉) · P. Zhang

Department of Engineering Physics, Institute of Public Safety Research, Tsinghua University, Beijing, China

e-mail: zhhui@tsinghua.edu.cn

X. Liu

e-mail: lx88@tsinghua.edu.cn

P. Zhang

e-mail: zhangp14@mails.tsinghua.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_15

occupants' evacuation in high-rise buildings will be greatly challenged. The traditional method of building evacuation is the evacuation through stairs. In accordance with the relevant regulations for building fire prevention, the public is instructed to not use elevators in case of fire. However, the traditional stair evacuation is faced with many disadvantages [1]. First, it is difficult for mobility impaired occupants such as the elderly and the handicapped in the building to evacuate through stairs. Second, the occupants in a high floor have to travel a long distance to reach the safe area. It may be resulting in long evacuation time, physical fatigue, and mental panic, which can even lead to crowd stampede in stairs with a high density of people. Therefore, other methods should be encouraged, such as the evacuation of elevators [2, 3]. Using elevator as the evacuation tool has several advantages, for instance, it can reduce evacuation time and shorten travel distance, which is convenient for occupants with mobility impairments and lead to less panic and confusion [4]. However, this method also encounters great challenges, that is, the safety and reliability of an elevator under emergencies, especially in case of fire.

In this work, we study the feasibility of using an elevator as the evacuation tool in high-rise buildings and collaborative evacuation strategies among stairs and elevators. The arrangement of this paper is as follows: Sect. 2 discusses the feasibility of elevators used for evacuation in combination with the safety of elevators and the evacuation cases using elevators at home and abroad. Section 3 introduces the simulation scenarios and the collaborative evacuation strategies of elevators and stairs. Section 4 analyzes the results and gives recommendations. Section 5 gives a summary.

2 Feasibility Analysis of Using Elevator as the Evacuation Tool

It has been a long time that elevators are prohibited for use in case of fire in a high-rise building. The main reason is due to the fact that the safety performance of traditional elevators would be affected by fire accidents [5]. For example, elevators are vulnerable to smoke. Since an ordinary elevator mostly has no fireproof front room, smoke can easily pass through the elevator shaft and spread upward quickly, causing damage to the occupants in the elevator. Meanwhile, the smoke can transfer to other floors through the chimney effect. The safety of the elevator itself is affected in the event of a fire, and an ordinary elevator usually does not have the waterproof function. The use of water may cause the function failure of some elevator's electrical components and the power lines. The ordinary elevator's normal power supply cannot be guaranteed in the event of a fire, resulting in the elevator being out of service and occupants being trapped in the elevator. Another reason is that ordinary elevators cannot guarantee reliable and effective personnel operations in the event of an emergency. Using an elevator without an emergency evacuation plan and a reasonable and scientific elevator evacuation rule may cause counterproductive results. However, with the development of intelligent building technology and the require-

ment for rapid evacuation of high-rise buildings, the demand for elevator evacuation in high-rise buildings has become more urgent, and the feasibility and reliability have also been continuously improved.

(1) Technical support measures for elevators

Similar to firefighting elevators, when ordinary elevators are used for the evacuation of high-rise buildings, safeguard measures should be taken in aspects such as fire prevention, smoke prevention, waterproof, and power security. At present, the development of elevator technology can solve these problems, such as setting up a separate fireproof front room. Positive pressure air supply measures are taken to provide elevator car protection performance, which can prevent smoke, fire, and water. Elevator doors and steel wire ropes and other important components meet the design of fire protection and high-temperature resistance. The power supply of the elevator adopts fire-resistant cables, and standby power is supplied to ensure the reliability of the elevator's power.

(2) Elevator evacuation criteria and recommendations

In the early 1990s, the National Fire Protection Association (NFPA), the American Institute of Standardization and Technology (NIST), the American Society of Mechanical Engineers (ASME), and other international organizations organized a series of seminars on the application of elevators in fires, and some standards were successively issued, such as ASME A17.1/CSA B44-Safety Code for Elevators and Escalators, NFPA 101-Life Safety Code, NFPA 72-National Fire Alarm and Signaling Code, International Building Code (IBC), etc. [6]. These standards or recommendations allow elevators to be used for occupant evacuation under fire conditions, especially for those who are mobility impaired in super high-rise buildings. At the end of 2005, the Shanghai Fire Protection Institute of the Ministry of Public Security cooperated with the Special Equipment Supervision and Inspection Technology Institute to study the feasibility of elevator emergency evacuation in the case of high-rise building fires and obtained the conditional use of elevators for emergency evacuation under fire conditions. These standards and researches provide important references for the use of elevators in the evacuation of high-rise buildings under fire and other emergencies.

(3) Influence of intelligent building technology on elevator evacuation

The development of intelligent building technology has developed rapidly in the last decade. Artificial intelligence, Internet of Things, big data, and cloud computing technologies have been used in building control. The level of automation and intelligence in buildings has been continuously improved, such as the distributed or non-centered research in recent years. In the intelligent building system, the building system has the ability to efficiently share and process related information. When a fire happens, the heat-sensitive smoke detectors in the building can accurately detect the occurrence of fire and predict the spread of fire based on the intelligent fire protection system. And the spread of smoke, while intelligent node computer process node (CPN) in the intelligent building subspace, can collect occupants' distribution

and movement data from smart sensors such as video surveillance, radio frequency identification (RFID), infrared sensors, Bluetooth, and WIFI positioning devices. Based on the intelligent evacuation indication system, path planning for the safety and rapid evacuation of occupants can be provided, including the rational use of elevators and stairs, such as the reasonable dispatch of elevators in case of a fire.

(4) Successful cases of elevator evacuation

In high-rise building fire accidents, elevators played an important role in some evacuation cases. For example, in the 2001, 9.11 terrorist attacks in the United States, many people in the South Tower of the World Trade Center successfully evacuated through elevators. In 1974, four elevators successfully evacuated 300 people in a fire in Sao Paulo, Brazil. In the fire accident at the Motomachi apartment building in Hiroshima, Japan, many elderly people used elevators to evacuate. At present, some emergency evacuation plans for high-rise buildings in the world also consider elevator-assisted evacuation methods, such as the Shanghai World Financial Center, Taipei 101 Building, Burj Khalifa, Eureka Tower, and other high-rise buildings. These evacuation cases showed that elevators are efficient evacuation tools when emergencies happen in high-rise buildings [7], which can greatly shorten the evacuation time and effectively improve the evacuation efficiency, especially for people with mobility impairment.

3 Simulation Experiment Setting

In this section, we use a specific high-rise building to explore the effectiveness of using elevators for occupant evacuation in high-rise buildings through a series of simulation scenarios. The building has a height of 75.8 m, a total of 20 floors, and a building area of about 25,000 m². There are two elevators and one staircase on each side of the building. We used the evacuation simulation software Pathfinder to establish a simulation scenario. Each floor was randomly distributed with 100 people and a total of 2,000 people were in the building. The speed of pedestrians was set to 0.5 m/s. The rated capacity of the elevator is 20 occupants, and the acceleration is 1.2 m/s², the rated speed is 2.5 m/s, and the opening and closing time is 3.2 s. The scene graph is shown in Fig. 1.

After establishing the architectural scenario, a series of evacuation plans, including different coordinated strategies for elevators and stairs, are simulated to investigate the effectiveness of elevator-assisted evacuation. When an emergency such as a fire occurs on a certain floor, if the accident has little effect on other floors, then only the affected floors or adjacent floors need to be evacuated. This is called partial building evacuation. If the event poses a threat to the safety of the residents of the entire building, it is necessary to evacuate all people in the building, which is called total building evacuation. In this work, these two kinds of building evacuation are simulated and analyzed.

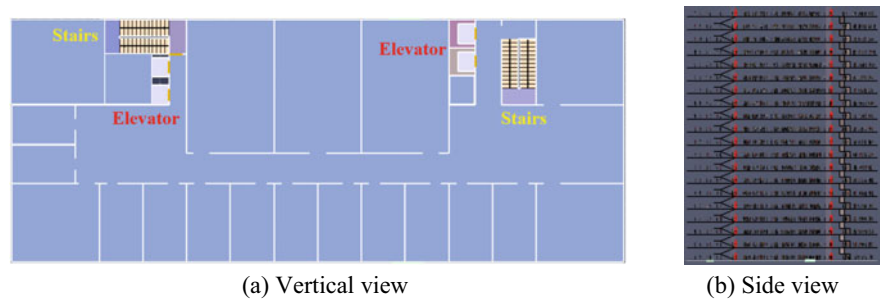


Fig. 1 The scene graph of a high-rise building

4 Results and Analysis

Through Pathfinder’s simulation of various scenarios, we focused on the overall evacuation time, evacuation process, evacuation congestion, etc. We discussed the impact of evacuation strategies on partial and total building evacuation.

4.1 Partial Building Evacuation

First, simulation of partial building evacuation was conducted, in which three different fire floor scenarios were set, that is, 10th, 15th, and 20th floors. In the scenario, when a fire occurs on a certain floor, the adjacent upper and lower floors need to evacuate people. The simulation results are shown in Table 1. The floor 10 stairs scenario represents that the fire happened on the 10th floor, and the stairs represent that occupant only uses stairs to evacuate. The elevator represents that occupant only uses elevators to evacuate. The stairs + elevator stands for use of elevators and stairs for coordinated evacuation (50% of the people on each floor use stairs or elevators to evacuate), and evacuation time represents the total time for 300 people on three floors to evacuate the building.

In these simulation scenarios, the evacuation time and the evacuation distance using the stairs alone are relatively the largest. When elevators were used for assisted evacuation, the evacuation time is significantly reduced and the average travel distance of occupants is greatly reduced. With the increase of evacuation floors, the advantages of elevator evacuation are more obvious. For example, if the fire happens on the 20th floor, the time for evacuation using stairs is 1080.4 s, and the time for using elevators to evacuate is 863 s. The evacuation efficiency is improved, and what is more important is that it reduces the average evacuation distance, which has an important impact on the mobility impaired occupants in high-rise buildings, because long-distance evacuation will cause physical fatigue and psychological panic on people.

Table 1 Simulation results in different scenarios under partial building evacuation

Scenario	Evacuated floor	Evacuation time (s)	Average travel distance (m)
Floor 10 stairs	9, 10, 11	698.5	177.7
Floor 10 elevator	9, 10, 11	674.6	43.4
Floor 10 stairs + elevator	9, 10, 11	608	104.7
Floor 15 stairs	14, 15, 16	921.6	248.1
Floor 15 elevator	14, 15, 16	787.7	44.2
Floor 15 stairs + elevator	14, 15, 16	857.6	136.0
Floor 20 stairs	18, 19, 20	1080.4	301.7
Floor 20 elevator	18, 19, 20	863.9	44.4
Floor 20 stairs + elevator	18, 19, 20	1009.8	160.6

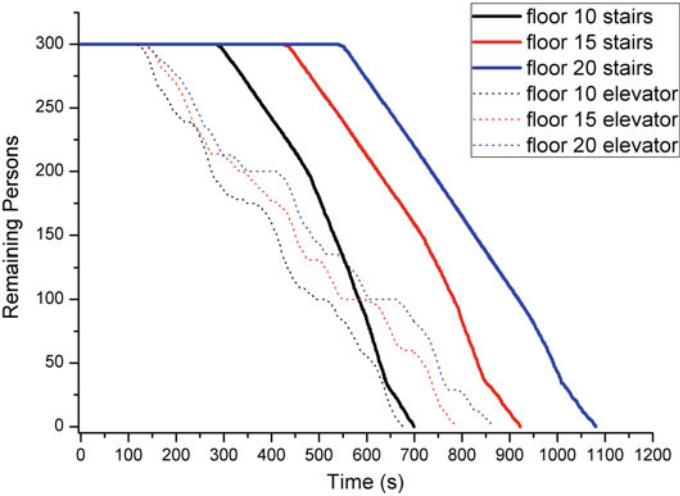


Fig. 2 Evacuation process under different scenarios

Figure 2 shows the evacuation process under different evacuation scenarios. It can also be seen that the use of an elevator can evacuate more occupants, especially in the early stage of evacuation. In the early stages of an emergency such as a fire, the accident may not be very strong and the safety of the elevator can be guaranteed. It is of great significance to evacuate as many people as possible during this time.

In the scenario where elevators cannot be used, firefighters need to go through the stairs for search and rescue work. This may lead to conflicts between upstream firefighters and downstream evacuees, slowing down the overall evacuation process.

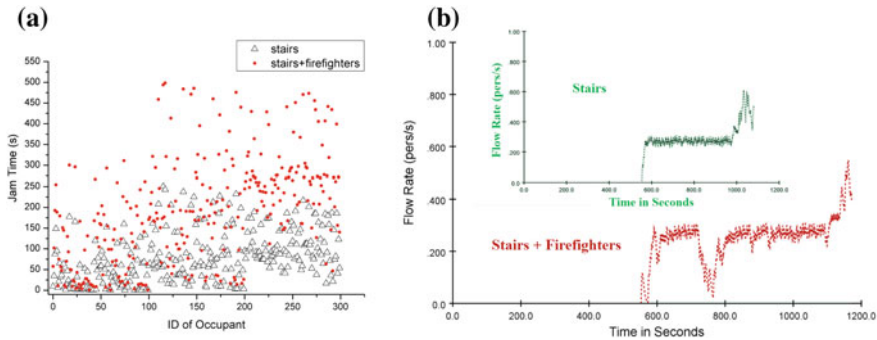


Fig. 3 Comparison of Jam time (a) and flow rate (b) between scenarios with and without converse firefighters

Furthermore, the opposite pedestrian flow can easily cause high density and may lead to crowd stampede accidents. We set up a simulation scene to study the impact of firefighters retrograde on the overall evacuation. In the floor 20 stairs scenario, 30 firefighters need to go to 18th, 19th, and 20th floor to conduct rescue work on the three floors.

When there are retrograde firefighters, the time required for occupants to evacuate through stairs increases from 1080 to 1172.1 s. What is more important is that due to the presence of retrograde firefighters, occupants are more obstructed during evacuation. As can be seen in Fig. 3, the occupants' jam time is significantly increased, and the flow at the evacuation exit will also have significant fluctuations due to the congestion on the stairs, as shown in Fig. 3b.

4.2 Total Building Evacuation

Then, we assume that fire and other emergencies affect the safety of occupants in the entire building and require total building evacuation. We focus on the impact of the use of elevators or stair evacuation and coordinated evacuation strategies on the evacuation time of the entire building. There are 100 people randomly distributed on each floor and a total of 2,000 people need to be evacuated. Specific simulation scenarios are set as in Table 2. Total elevator represents that all occupants use elevators to evacuate. Total stairs represent all occupants use stairs to evacuate. Stairs (14) + elevator (6) represents that occupants at the top 6 floors are evacuated by an elevator, and other floors are evacuated by stairs.

By setting up different elevator stair coordination evacuation strategies, it can be seen from Fig. 4 that there exists an optimal evacuation scheme balance point for a specific scenario. For example, in our simulation scenario, stairs (16) + elevator (4) is the scenario with the least evacuation time. This result is obtained in a specific building and it may differ under other tall buildings with various sizes, numbers, and

Table 2 Simulation results in different scenarios under total building evacuation

Scenario	Floor with elevator	Evacuation time (s)	Average travel distance (m)
Total elevator	2–20	3598	52.3
Stairs (14) + elevator (6)	15–20	1715	117.5
Stairs (15) + elevator (5)	16–20	1502	128.8
Stairs (16) + elevator (4)	17–20	1327	140.4
Stairs (17) + elevator (3)	18–20	1413.3	153.1
Stairs (18) + elevator (2)	19, 20	1493.6	166.5
Stairs (19) + elevator (1)	20	1581.9	180.7
Total stairs	–	1656.9	195.6

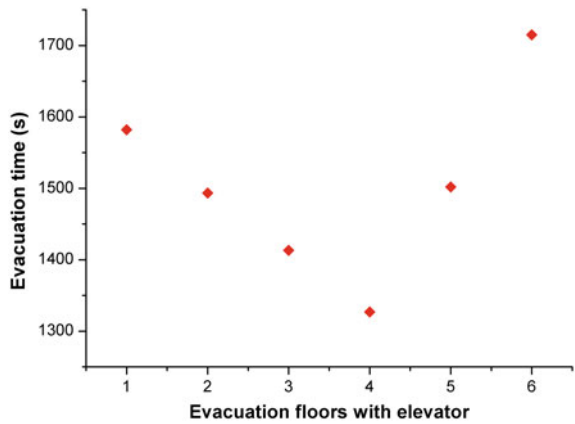


Fig. 4 Evacuation time under different scenarios with elevator evacuation

occupant loads. However, for tall buildings, combination of stairs and elevators can greatly shorten evacuation times and travel distance. The effect is more evident in the taller floor of the buildings, especially when there are inconvenient occupants in some floors, the shortest evacuation time is not the only goal, and the convenience of evacuation is also very important [8].

On the basis of ensuring the safety of the elevator, the use of elevator-assisted evacuation can improve the evacuation efficiency. However, when a fire and other emergencies occur, what kind of strategies should be adopted should be considered in combination with the building structure, the development of accidents, the composition and spatiotemporal distribution of the occupants, etc. New technologies

for intelligent buildings can improve the safety and reliability of high-rise building elevators, and can improve the monitoring and forecasting of emergencies such as fires. Building managers should formulate emergency evacuation plans in advance and improve the emergency management capabilities through daily exercises. To prepare in advance, respond reasonably when emergencies happen, and summarize experience after emergencies, and can continuously improve the efficiency and safety of elevator-assisted evacuation in intelligent buildings. Ultimately, it can provide a reference for the popularization of elevator evacuation in smart buildings in the future.

5 Conclusions

This paper analyzes the disadvantages of traditional building evacuation. Combining with the latest technology of intelligent building, we analyze the feasibility analysis of elevator evacuation of high-rise buildings. With the help of an evacuation software, Pathfinder, simulation experiments are given to study the collaborative strategies among stairs and elevators under partial and total building evacuation. Through simulations and comparison of results, it was concluded that elevator-assisted evacuation can greatly increase the efficiency of building evacuation. The higher the floor, the more obvious the effect. It can not only reduce the evacuation time but also can greatly reduce the travel distance, especially for mobility impaired occupants. The impact of retrograde firefighters on building evacuation in stair evacuation was studied. Collisions between occupants and firefighters increased the jam time. Simulation results in this work provide a reference for the popularization of elevator evacuation in intelligent buildings in the future.

Acknowledgements This work is supported by National Key Research and Development Project of China No. 2017YFC0704100 (entitled New generation intelligent building platform techniques), National Natural Science Foundation of China (Grant No. 71704091), and the China Postdoctoral Science Foundation (Grant No. 2017M610917)

References

1. Peacock, R.: Cooperative Research on the Use of Elevators During Fire Emergencies, in National Institute of Standards and Technology Special Publication 1620 (2009)
2. Luh, P.B., Xiong, B., Chang, S.C.: Group elevator scheduling with advance information for normal and emergency modes. *IEEE Trans. Autom. Sci. Eng.* **5**(2), 245–258 (2008)
3. Ma, J., Lo, S.M., Song, W.G.: Cellular automaton modeling approach for optimum ultra high-rise building evacuation design. *Fire Saf. J.* **54**, 57–66 (2012)
4. Kinatader, M.T., Omori, H., Kuligowski, E.D.: The use of elevators for evacuation in fire emergencies in international buildings-NIST (2014)
5. Bukowski, R.W.: Addressing the needs of people using elevators for emergency evacuation. *Fire Technol.* **48**(1), 127–136 (2012)
6. Ronchi, E., Nilsson, D.: Modelling total evacuation strategies for high-rise buildings. *Build. Simul.* **7**(1), 73–87 (2014)

7. Chen, J., Ma, J., Lo, S.M.: Event-driven modeling of elevator assisted evacuation in ultra high-rise buildings. *Simul. Modell. Pract. Theory* **74**, 99–116 (2017)
8. Butler, K., Kuligowski, E.: Perspectives of occupants with mobility impairments on evacuation methods for use during fire emergencies. *Fire Safety J.* **91**, 955–963 (2017)

Dynamic Emergency Evacuation System for Large Public Building



Yongming Zhang, Zhe Yan, Xueli Zhu and Wenjie Piao

Abstract Nowadays, it has been an important issue to evacuate mass occupants for safety from large public buildings under emergency conditions. However, evacuation design in public buildings still remains in a static plan, which is difficult to deal with diverse types of hazards and problems of uneven personnel distribution. Thus, a set of guidelines for safety design of large public buildings, or even a set of dynamical emergency evacuation systems, is necessary. In this paper, an emergency evacuation system for large public buildings is proposed based on intelligent building by means of two commercial simulation platforms—PyroSim and Pathfinder.

Keywords Large public building · Emergency evacuation · Simulation · PyroSim · Pathfinder

1 Introduction

In recent years, with the rapid development of the economy and the accelerating process of the cities in China, commercial complexes, multifunctional office buildings, stadiums, city subways, high-speed rail stations, and other buildings burgeon all around the cities have gradually become an integral part of people's daily life.

The large public buildings have the character of large-scale, complicated structure, single space, complex function, and high population density. In order to improve the land usage rate, the floor area ratio of the high-rise building is so large that it is easy to occur emergency and cause casualties. In underground buildings, due to poor ventilation and smoke extraction, it is very easy to cause significant casualties and property losses in case of fire and toxic gas leakage.

Y. Zhang (✉) · Z. Yan · W. Piao
Tongji University, Shanghai, China
e-mail: zhangyongming@tongji.edu.cn

X. Zhu
Silicon Lake College of Vocational Technology, Suzhou, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_16

According to the incomplete statistics of the United Nations World Fire Statistics Center, the annual global fire hazard was reported about 6–7 million times, causing the death of 65–75 thousand. According to the statistics from the Ministry of Public Security Fire Department of China, in 2015, there were a total of 338 thousand cases of fire, resulting in 1742 people dead and 1112 people injured, and the direct property loss of 3.95 billion RMB.

The study of personnel evacuation system is based on the research of fire science and fire protection analysis. In the 1970s, there was a study of the relationship between the average travel speed and population density based on statistical methods. In the 1980s, with the aid of modern computer technology, the scientific workers have established a model to describe the characteristics of building evacuation. This model first appeared in 1982, and Stahl established the first evacuation model BFIRES [1]. With constant development, there are dozens of models about the evacuation in the world. Among them, it is well-known commercial software such as University of Greenwich's building EXODUS [2], University of Edinburgh's SIMULEX [3], etc. There are a number of large-scale international consulting agencies which developed network software of personnel evacuation system, such as Pathfinder by the United States RJA and EGRESS by the British company AEA, etc.

The formation and development of large public buildings is the result of comprehensive effect of city society, economy, technology, etc. With providing a convenient and comfortable environment for our life, it takes the new task for the security personnel evacuation in various emergencies. It is a huge challenge for fire evacuation. The intelligent fire evacuation system is a good solution to this problem.

2 Frame of Emergency Evacuation System

2.1 *Frame of Information Input System*

The information input system (Fig. 1) is mainly divided into two aspects: the one is the existing static building information, including the building structure, function partition, and other basic information; the other one is dynamic information within the building, including hazard information and personnel distribution information in the building (personnel density, location, congestion area). The input information needs judging by different types to select methods: static information collection can use Building Information Modeling (BIM); in dynamic information acquisition, hazard information collection can be based on fire detector system, and personnel distribution information collection can be based on wireless sensor technology.

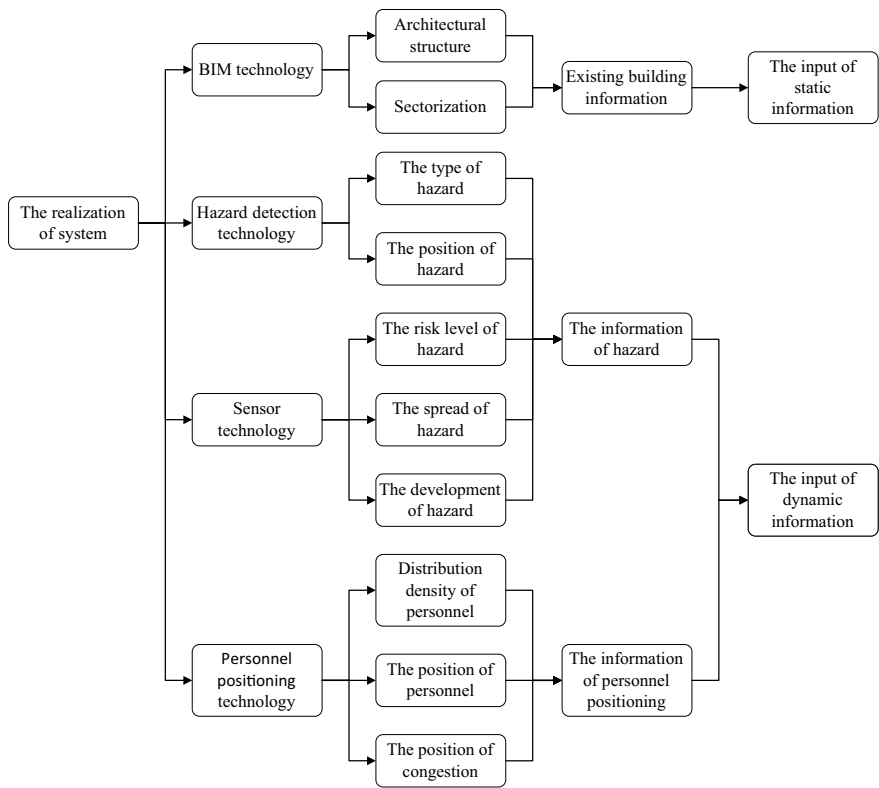


Fig. 1 The input system of dynamic emergency evacuation

2.2 Frame of Main Control Module System

The information input system (Fig. 2) sends the hazard information and personnel distribution information to the hazard prediction model, and then in the first time it calculates limit evacuation time and dangerous zone and sends the information to the evacuation route calculation model. The evacuation model on the basis of the actual personnel distribution, the limit time of evacuation, and dangerous area calculates the best escape plan, issued guidelines for evacuation instructions. The real-time information of hazard, personnel real-time distribution, and congestion as the feed-back value are sent into the hazard prediction model and evacuation route calculation model, for real-time calculation and dynamic optimization of the evacuation plan.

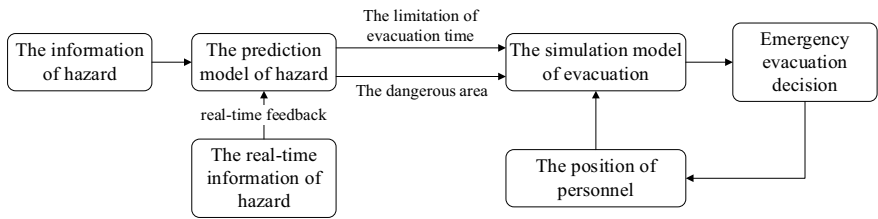


Fig. 2 Main control module system of dynamic emergency evacuation

2.3 Frame of Implementation System

The execution module of an emergency evacuation system (Fig. 3) mainly consists of two parts: the one is the emergency lighting and the other one is the emergency evacuation route guidance. They are the display devices of a dynamic emergency evacuation system.

The traditional fire evacuation guidance includes fire alarms, fire broadcast, and emergency exit lights. In the optimized system, the power supply of emergency lighting and other guidance devices can be used in the fire alarm and firefighting linkage system, such as the Uninterrupted Power Supply (UPS). The evacuation guidelines can be improved from two aspects: the one is the light guidance like

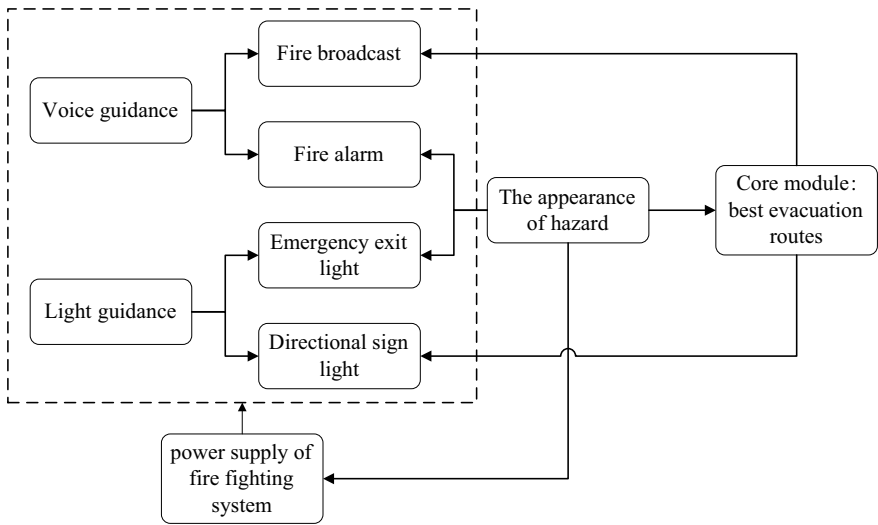


Fig. 3 The execution system of dynamic emergency evacuation

two-way adjustable fire evacuation signs with special effects such as flicker, visual continuity; the other one is the voice guidance, which can be based on the fire broadcast. It can notice people the best evacuation routes from the calculation of core module.

3 The Realization of the Emergency Evacuation System

3.1 The Realization of Information Input System

The input information system mainly includes two kinds of information: static and dynamic information. In order to realize the data collection, the following three kinds of methods can be adopted:

- To extract building information from BIM;
- To use various types of sensors for monitoring the occurrence of hazards, the development of hazard; and
- To use infrared count technology and image recognition technology for the distribution of personnel.

3.2 The Realization of Main Control Module System

In the second chapter of the frame of main control module system, there are two submodels: hazard prediction model and evacuation route calculation model.

Whether the prediction of the hazard or the calculation of evacuation route, they are required that building models and specific hazard types are provided. So, in this chapter, a comprehensive laboratory building with the fire hazard is chosen as the background.

PyroSim is used for hazard prediction model, and Pathfinder is used for evacuation route calculation model.

Hazard Prediction Model

PyroSim [4] software is fire dynamic simulation (FDS) software. On the basis of FDS, it provides a graphical user interface for the dynamic simulation of fire. It is used to create a fire simulation to accurately predict the fire smoke flow, the fire temperature, and the concentration distribution of toxic gases. The software is based on computational fluid dynamics, and the simulation model predicts the development information of gases, such as smoke, carbon monoxide, and other toxic gases. PyroSim provides a model for the establishment of grid, the definition of material, create surface, and structures and air ventilations or the definition of human attributes under the condition of evacuation, and finally get a numerical model for solving the complete fire operation.

Evacuation Route Calculation Model

Pathfinder [5] software is based on the personnel movement, which provides a graphical user interface to results of three-dimensional visualization tools. It provides a complete 3D triangular mesh design, in order to match the actual construction mode. It can calculate the independent movement of each member and give everyone a unique set of parameters. There is society of fire protection engineers (SFPE) model and steering model which are used in the simulation of human behavior. In SFPE mode, the movement of people is just a straight line from its current position to the evacuation routes, and people will not affect each other. The steering model is based on reverse thinking, and when the distance between people and the nearest point of the path reaches a certain threshold, it will regenerate new path to adapt to the new situation. When people meet the obstacles and congestion, they may move in the direction as shown in Fig. 4. The model uses the steering algorithm to determine the movement and the speed, and therefore forms a curve path.

Because the steering mode is more complex and uses the method of path planning, guiding mechanism, and collision treatment to control the movement of people, the final simulation shows more realistic. However, due to the lack of export flow constraints, the results of the calculation have a certain difference with the calculation results of SFPE manual. In summary, in order to ensure the calculation of the authenticity and accuracy, the steering mode is used for simulation of evacuation.

Analysis

Through the simulation results and its analysis, it can be found that the two software can realize the system main control module to the hazard development forecast and evacuation guidance, but it also has some problems.

First, even if the information input end can send the various types of information to the main control module, required simulation time is still too long. In order to

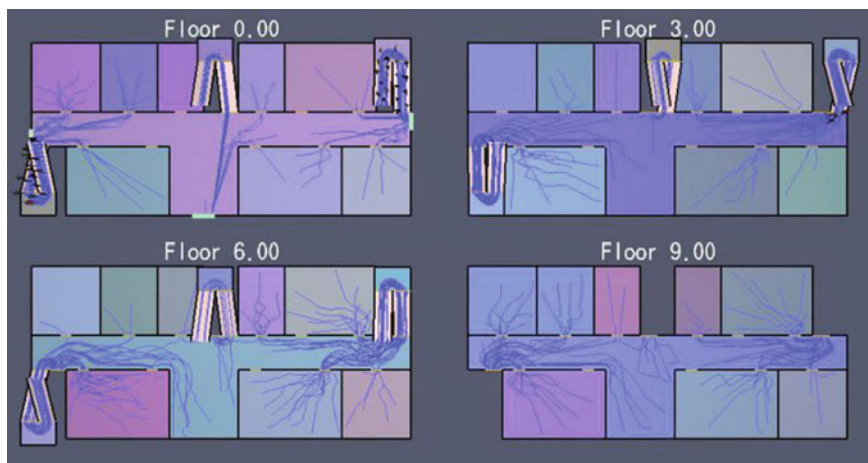


Fig. 4 Escape route in steering mode

solve this problem, in the process of building evacuation design, it can in advance simulate a set of reference scheme. When an actual hazard occurs, it can be detected by the sensor location, and the system can choose one reference scheme for achieving “semi-dynamic” evacuation guidance.

Second, the evacuation process of all path can be shown: in this fire scene, the pressure of the central evacuation exit is significantly more than the eastern side evacuation exit. People will follow the principles of the nearest in the steering mode, so it will cause a congestion trouble.

On this issue, the relevant researchers have done a research and put forward the “exit attraction” principle of safe export selection:

$$A_i = e^{F_1} e^{F_2} e^{S_i} \quad (1)$$

$$F_1 = k_1(D_{\max} - D_i) \quad (2)$$

$$F_2 = k_2(N_{\max} - N_i) \quad (3)$$

where D_i is the distance to No. i exit, k_1 is the attractive coefficient, N_i is the number of people to No. i exit, k_2 is the resistance coefficient, $D_{\max}$ is the distance to the farthest exit, S_i is the speed to No. i exit, and A_i is how the No. i exit attracts to the people.

In practical application, according to the simulation results with the actual distribution, people should be evacuated before congestion in order to avoid adverse consequences caused by congestion.

3.3 The Realization of Implement System

The implement module improves mainly from the light guidance, the sound guidance, and the independent power supply. The improvement of broadcast and power supply can be realized by using the fire broadcast and UPS.

3.4 Expectation

The system depends on BIM, detection of fire detectors and other dangerous source sensors, personnel counting technology, and video monitoring technology, and relies on the bus connection of the building automation system (BAS).

In the design of evacuation, the basic information of building needs to be integrated, and it also needs to be simulated for every hazard-prone position (such as the use of PyroSim to simulate the fire hazard). After obtaining the simulation results, it needs to change the improper design and input database.

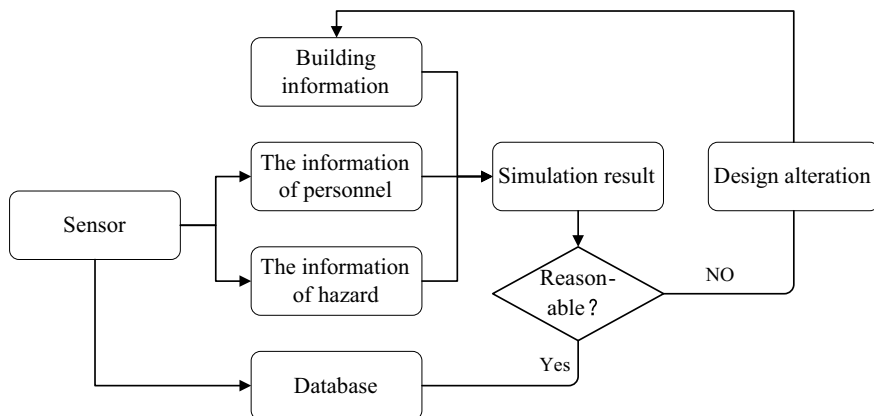


Fig. 5 The improvement of the design of evacuation and guidance

According to hazard simulation database, the managers formulate schemes of evacuation and make plan database for every hazard-prone position. In practical application, according to the type and position of the detected hazard, the system will choose one from plan database and execute instructions for evacuation.

Based on the completion of this research part, there are several aspects worth further improvement and exploration:

- Optimization of personnel behavior model: The model can be further optimized by the principle of “exit attraction”, and even the further development of the software.
- To reduce the simulation time: The simulation software in this passage is more advanced, but it still needs a long time. It is expected to achieve in the real sense dynamic evacuation guidance in emergencies.
- Technology of Building Information Modeling (BIM): Although BIM technology is used as the source of building information, in the actual simulation, only traditional ichnography is used. It is expected to achieve the direct transformation from the construction to the information after in-depth study of BIM (Fig. 5).

4 Case Study

This case is based on 28-floor building in Shanghai which occurred fire accident on November 15, 2010. The building was a high-rise building with 17,965 m² area.

With the simulation in SFPE model, the total evacuation time of the building is 1057 s (Fig. 6).

If the evacuation route of the building personnel is designed, a more explicit guidance can be given. As shown in Fig. 7, the evacuation time is 784.8 s under the steering model. In other words, if the crowd is guided in the guidance mode, it

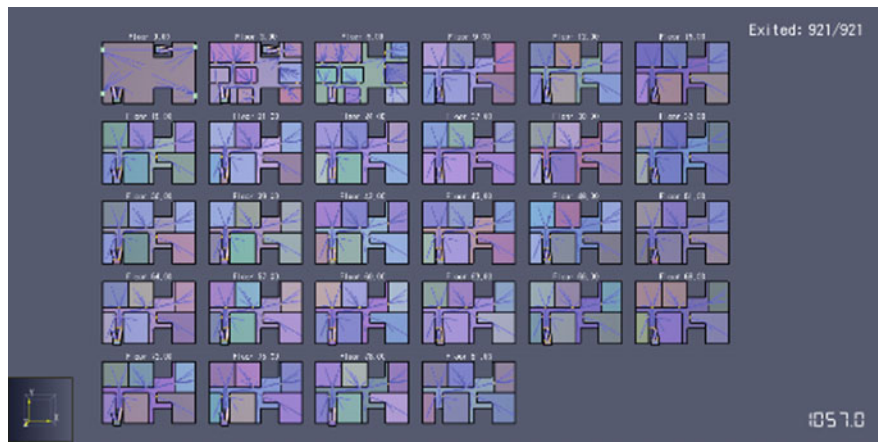


Fig. 6 The evacuation time under SFPE model

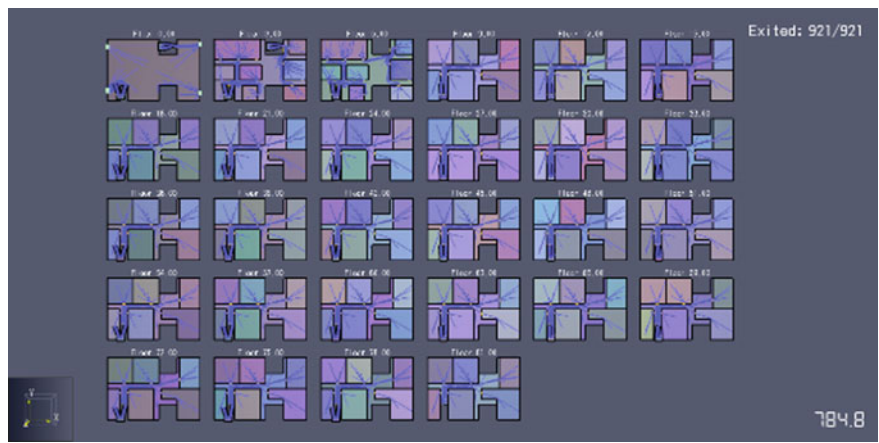


Fig. 7 The evacuation time under steering mode

can save 272.2 s for the whole evacuation operation, which means more time for evacuation.

5 Conclusion

The research topic of this paper is based on the issue of building intelligent engineering, fire protection engineering, and safety engineering in recent years, how to realize in short time with high efficiency and safety evacuation in building. Especially in the condition of uneven distribution and partition of various functions, the personnel

evacuation once meets problems, which is extremely easy to cause a serious casualty. After the simulation of the actual building model, the design of evacuation and practical guidance is improved.

Acknowledgements This paper is supported by Science and Technology Support Plan Project from Ministry of Housing and Urban-Rural Development of China (Grant No. 2016-K4-024).

References

1. Stahl, F.I.: BFIREs-II: a behavior based computer simulation of emergency egress during fires. *Fire Technol.* **18**(1), 49–65 (1982)
2. Gwynne, S., Galea, E.R., Lawrence, P.J., Owen, M., Filippidis, L.: A systematic comparison of model predictions produced by the building Exodus evacuation model and the Tsukuba pavilion evacuation data. *J. Appl. Fire Sci.* **7**(3), 235–266 (1998)
3. Thompson, P.A., Marchant, E.W.: A computer model for the evacuation of large building populations. *Fire Saf. J.* **24**(95), 131–148 (1995)
4. Glasa, J.: Use of PyroSim for simulation of cinema fire. *Int. J. Recent. Trends Eng. & Technol.* (2012)
5. Wang, H.R., Chen, Q.G., Yan, J.B., Yuan, Z., Liang, D.: Emergency guidance evacuation in fire scene based on pathfinder. In: *International Conference on Intelligent Computation Technology and Automation* (pp. 226–230). IEEE (2015)

Status of Intelligent Building Development of China—Questionnaire Analysis



Huai Li, Zhen Yu and Wei Liu

Abstract A questionnaire survey on the status of intelligent building development was carried out. Based on actual survey questionnaire data, the development status of technology and applications was described in this paper, mainly discussing the existing problems and the proper future development. It is aimed at providing valuable findings on intelligent building development, exploring an innovative way to build a practical, reliable, and easy-used intelligent system to meet the requirements of users. Among all the designing aspects, the configuration of both hardware and software is the most tedious and error-prone part when building an intelligent system, as a huge workload is inevitable when solving the difficulties on system transformation. Unsuitable operation logic of the system and equipment is another problem, as the distance between the original expectation of the intelligent system and the real operation effect is still large. Though some building systems realize integration between the subsystems, they work independently, thus enabling different subsystems to work as a whole is another problem which needs to be emphasized. A much simpler configuration process or structure of the software and hardware system is significantly required in this field. If the controller could adapt its own operation logic automatically according to preset equipment parameters, this problem might be solved fundamentally.

Keywords Intelligent building · Building automation system · Questionnaire

H. Li (✉) · Z. Yu
China Academy of Building Research, Beijing, China
e-mail: lihuai027@qq.com

W. Liu
Nanjing Tech University, Nanjing, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_17

1 Introduction

It is shown that human beings spend almost 80% of their daily time in buildings. Hence, to create a comfortable indoor environment which can provide a high-quality living condition that is extremely important, as it is beneficial for people's physical and mental health, increasing both the quality and efficiency of daily works. Building intelligence combines computer software, hardware, and communication technologies together, which can facilitate the interaction between building equipment, facilities, and personnel, creating a safe, comfortable, convenient, and efficient indoor environment for residents living and working in the buildings.

During the past 30 years, building intelligent system development has achieved a huge progress, being widely used in different kinds of public buildings. Some of the systems have achieved some breakthrough, but it is worth noticing that many productions have also revealed flaws. Ref. [1] and [2] pointed out that intelligent building shows its advantage in aspects such as "improve indoor environment, safer, easier organize", but only 17% of people agree that the current intelligent building performance meets their expected requirements.

Specifically speaking, the rapid development of the Internet of Things technology (IoT) has provides a solid basis for the upgrading of building intelligence. However, it is required that there should be a better performance on building management, energy saving, and indoor environmental quality improvement.

The China national key R&D program "New generation intelligent building platform techniques" conducted a questionnaire survey on the current status of intelligent building development. The research questions are designed to explore the following aspects: (1) what kind of problems exists during the whole application process of intelligent systems; (2) what are the practical requirements of the users, owners and operation managers of intelligent building; and (3) what kind of system-designing or technology can help to achieve the ideal version of intelligent buildings.

Based on the questionnaire data, the development statuses including technology and applications were analyzed, and the existing problems were discussed as well in the paper. It is inferred from the survey that (1) integrated management and operation of individual subsystem is required; (2) automatic detection and diagnosis is necessary when system or equipment error happens; (3) intelligent integration process should be simple and reliable, and it is better that the system can operate when plugging; (4) control strategy can adjust different indoor environment; and (5) a user-friendly operation system should be optimized.

Since the main target is to identify the current development of intelligent buildings, people who are mainly involved in this work, including the integrators, system operators, and users, are interviewed. Three different questionnaires are compiled according to the features of different sample groups, and feedback from the building intelligent integrators is mainly analyzed.

2 Questionnaire Construction

2.1 Section Ration

The questionnaire involves main aspects of the building intelligent system and Ref. [1–5] also provide valuable reference. The questionnaire is conducted from the following six aspects: (1) general information of the buildings; (2) the development and application status of the intelligent systems; (3) cost components; (4) user requirements; (5) most commonly problems; and (6) technologies development and trends. Among all the questions, 65% is single choice questions, 30% is multiple choice questions, and 5% is blank-filling questions. The general information of the interview is shown in Table 1, and the detailed percentage of each item is listed.

2.2 Structure of the Questionnaire

The general information of the intelligent system building is collected in Sect. 1, from the following six aspects: (1) company size; (2) basic condition of sub-intelligent systems; (3) importance of different subsystems; (4) operational stability; (5) failure rate; and (6) subsystem integration rate.

The second section is designed to find out the current application status of intelligent systems. They are: usage rate of BAS system, refresh cycle of sensors and controllers, mainstream program languages, communication protocol, and application of APPs.

Cost section focuses on the hardware, software, installation cost, and labor cost. Requirement section is set to explore the real requirement of the users from aspects of information interactions, energy saving, BIM operation, satisfaction, and data analyze.

Table 1 Components of the questionnaire

	Content	Question number	Percentage
1	Building information	4	0.12
2	Current status	6	0.18
3	Cost	7	0.21
4	Requirement	3	0.09
5	Existing problem	8	0.24
6	Technology development	2	0.06
7	Response information	3	0.09
	Total	33	

Existing problem section gathers information on the limitations of successful and large-scale application of the intelligent systems from the aspects of system construction, daily operation, energy conservation, and comfort needs.

The questionnaire was carried out based on the 3rd China International Intelligent Building Exhibition 2018. A total of 158 valid survey questionnaires were collected on-site, and the following analysis was based on the questionnaire statistics.

3 Results Analysis

3.1 General Information of the Interview

Figures 1 and 2 show the detailed career information of the interviewees. Overall, 39% of the respondents are engineers either majored in hardware/software or common engineer, 15% are people involved in this industry, and 36% of the respondents did not leave job-related information (Fig. 1). Among the valid data group, 50% of the respondents have been worked in this industry for more than 10 years, 15% have work experience of about 5–10 years, and 30% of them left this blank empty.

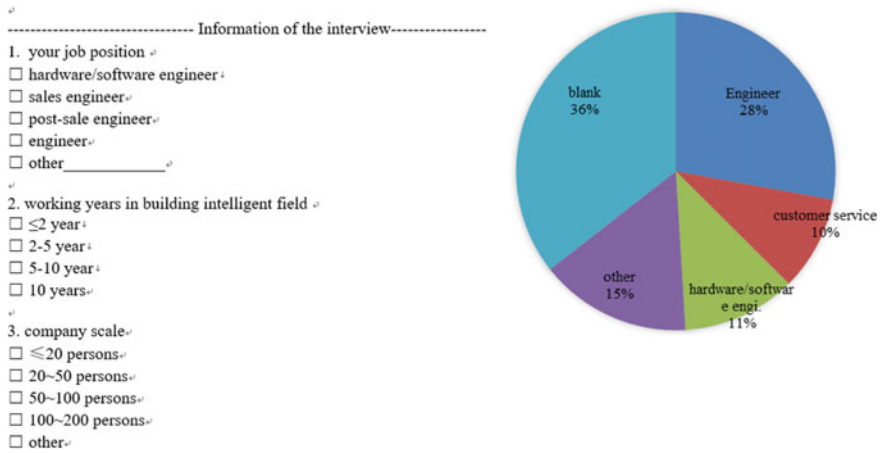


Fig. 1 Career of the interviewer

3.2 Application Status of the Building Intelligent

Application: It is shown in Fig. 3 that the intelligent system has been widely used in almost all kinds of buildings, such as public buildings, commercial buildings, hospitals, schools, and government office buildings.

Scope: The feedback reveals that the basic components included in intelligent building systems are “Building Automation System” (BAS), “Building Energy Man-

Fig. 2 Working years

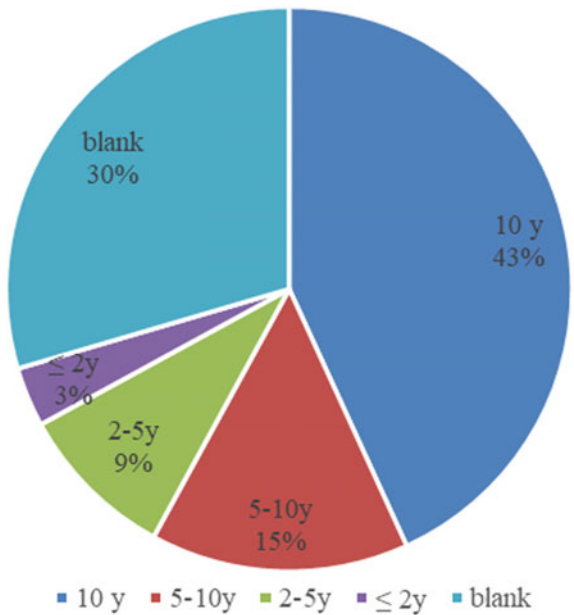
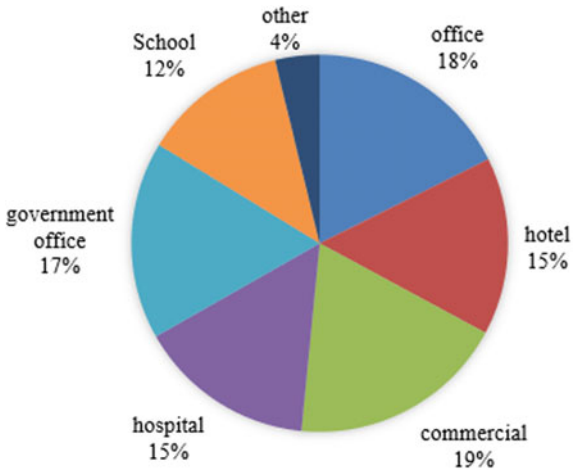


Fig. 3 Building types



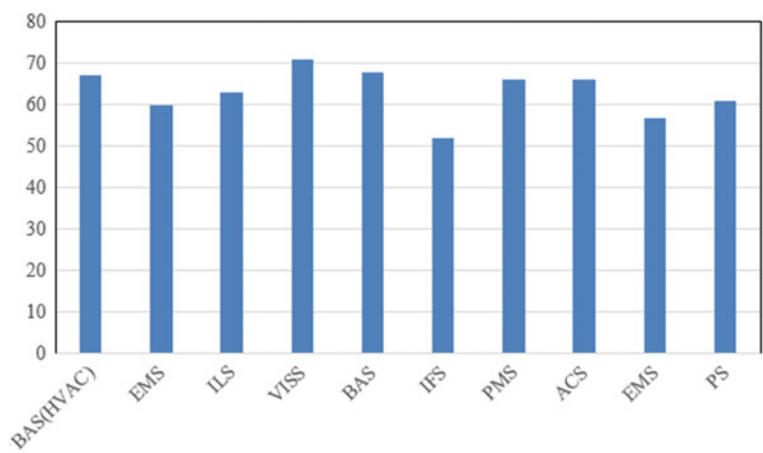


Fig. 4 Subsystems included in an intelligent building

agement System” (BEMS), “Intelligent Lighting system” (ILS), “Video Surveillance System” (VSIS), “Burglar Alarm Systems (BAS)”, “Intelligent firefighting system” (IFFS), “Parking Management System (PMS)”, “Access Control System” (ACS), “Elevator Monitoring System (EMS)”, “Electronic Patrol System (EPS)”, etc. (Fig. 4).

Brand: Among the intelligent subsystems, IFFS, BAS (HVAC), (EMS), and (ILS) are dominated by foreign brands, while ACS, VSIS, and PMS are gradually controlled by local brands. This infers that currently the building intelligent systems rely on foreign technology to some extent, which suggests that local technologies and systems should develop further on efficiency and stability.

Failure rate: Failure frequency of the intelligent systems was also investigated. ACS, BAS (HVAC), and BAS are among the top three, which have a higher failure rate. It is understandable that the most frequently used system would have the highest failure rate (Fig. 5). Integration: About 58% of the sub-system realize big system integration.

3.3 Application Status of the BAS (HVAC)

The BAS (HVAC) market is dominated by mainstream brands, such as Siemens, Honeywell, and Johnson Controls, while the market share of the local brands is relatively small. According to the survey, 59% of the respondents consider that mainstream brands dominate the market for a long time, and thus their system is relatively steady, and will maintaining the leading position for a while. Besides, 29% of the respondents believe the lack of disruptive subversive BAS system/products in the commissioning and installation process is the problem of low market share of local brand.

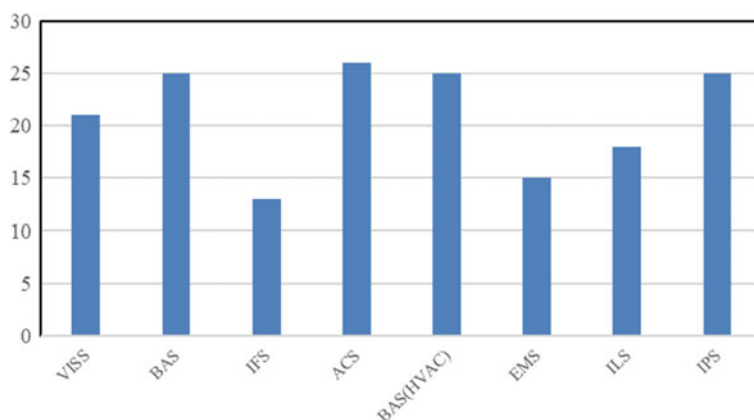


Fig. 5 Failure rate of different subsystems

About 38% of HVAC system could realize remote on/off control, of which 64% works normally; 38% of HVAC system could work under a certain logic, such as time table, special schedule.

Among the normal operated HVAC systems, 50% of the BA platform could get accurate parameter feedback value.

Intelligent control logic of the chiller station are always compiled by chiller producers and then integrated to BAS. From the questionnaire, the control logic was seldom utilized and rarely well executed, the operator pointed out that: 1) the system require professionally maintenance once the system failure; 2) no obvious energy saving or comfortable indoor environment after the usage of the logic.

Besides, 48% of the respondents believe that the upgrading period of the building intelligent system is about 5–10 years, which infers that the maintains cost is an important factor for the system development.

For the BAS (HVAC), 33% of users pay more attention to the stability and reliability of the system, 23% of users are more concerned about the improvement of the entire management level brought by BAS (HVAC) usage, 16% believe that the BAS (HVAC) shows its advantage in equipment optimization control, and 16% believe that the intelligent systems contribute to labor costs reduction.

50% of the respondents believe that the update period of actuators and sensors in the BAS is 3–8 years. 52% of the respondents think that the controller update cycle is 5–10 years (Figs. 6 and 7).

As important parts of BAS, quantities of actuators and sensors are guarantee of the steady operation. Their actions are decided by control logic and utilization frequencies. Imagine if an actuator moves 100 times a day, and it would soon be worn out. So, it is essential to have a proper control logic for a system. Generally speaking, under the current system structure, system or equipment control logic is written by software engineers, and only a small proportion of them have sufficient knowledge of HVAC system, which may cause problems.

Fig. 6 Update period of actuators and sensors

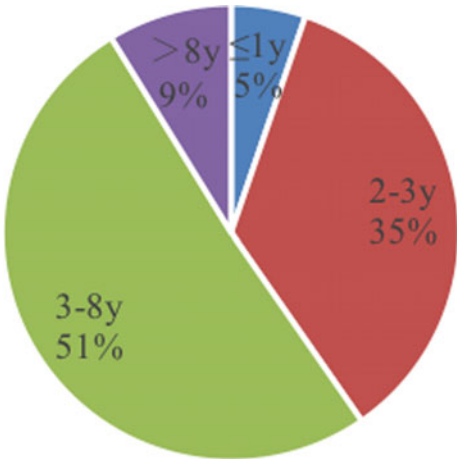
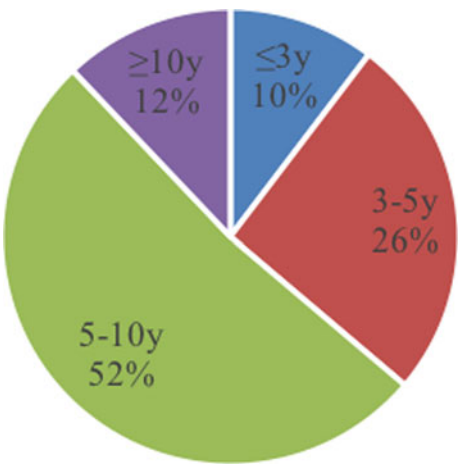


Fig. 7 Update period of the controller

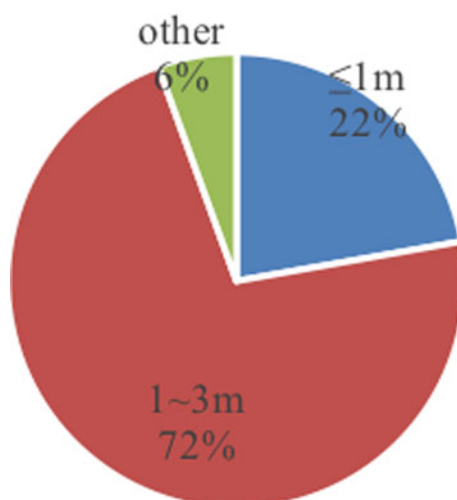


66% of integrators believe that developing various types of smart building APP is a trend of technical innovation on intelligent building in the future. Currently, access control system and BEMS are mostly used.

3.4 Cost

Cost is one of the most important factors which influences the end-use affection and popularization of the intelligent building. Being influenced by many different factors, the price level presents a multipolar phenomenon, especially in BAS. It is

Fig. 8 Debugging period of the on-site hardware



relatively sensitive to discuss as the detailed data is often confidential, and this survey investigates a few general aspects of cost components.

Take BAS (HVAC) system, for example. The investigation shows that cost of system integration is about 500–1000 yuan/point; the normal price of BAS software in the market is between 100,000 and 20,000 yuan; and the average debugging time of hardware and software is about 1–3 months (Fig. 8). And about two staff are involved in the daily system operation and maintenance work. In a comprehensive way, cost is an important matter for the user to choose the system.

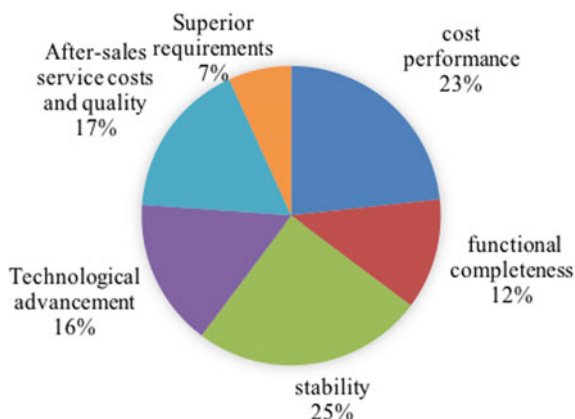
3.5 Client Requirement

It is shown that the stability and security of the intelligent system are the most important factors for the user to consider when choosing intelligent products, and the technical advancement and integrity are important as well. The combination of intelligent building and BIM to realize 3D operation and display are the current development trend, and there are many project applications coming out currently. Therefore, the quality and performance of intelligent products and system are the basic requirements of the market (Fig. 9).

3.6 Problems

The problem of building intelligence in building safety and security is one of the issues that are mostly concerned by users.

Fig. 9 Client requirement of the intelligent system



From the viewpoint of security, 20% of the respondents believe that the detection point is incomplete in many underused intelligent systems, while 19% believe that the electrical safety monitoring system is incomplete, and potential risks still exist. From the aspect of green energy saving, the main problems in intelligence system are the imperfectness of monitoring and measurement systems, and the energy-saving analysis function is insufficient as well. From the convenience aspect, the intelligent system is not intelligent as it is expected, and the control strategy of the equipment and system do not fit the real system.

Regarding the indoor environment, some intelligent buildings which have installed indoor environment sensor, but most of the equipment have not performed fit table control strategy.

After system integration, the integration or linkage between one subsystem and the third party is relatively unstable and prone to failure.

Regarding the steady auto-operation of the energy station, the feedback shows that

- The incomplete understanding of the equipment and system by the operation and maintenance person is the main reason why the HVAC system cannot be fully automated;
- The investment on the chillers or pumps is relatively high, and there are few successful and satisfying auto-operation applications which have satisfying automatic management stop the step to widely use the auto system;

The failure of the completely automated monitoring of the energy station is also one of the factors affecting the improvement of system intelligence.

In the process of intelligent system implementation, the most difficult things are

- a. Integration with third parties;
- b. System transformation; and
- c. Point configuration and proofreading.

4 Conclusions

From the questionnaire of current intelligent building development, it is found that:

- The domestic products in the field of access control, video surveillance, and other systems are becoming mainstream brands, while the market of building control, lighting, and elevator control are still firmly controlled by foreign brands. High damage rate of sensors and actuators is one factor which caused system instability. A much easier configuration process or structure of the software and hardware system is eagerly required in the field.
- The unsuitability of operation logic of the system and equipment is another main question that is far beyond the expectation of the intelligent system and the real operation result. The lack of HVAC knowledge of the programmers might be a problem, but not a key issue. If the controller could adapt its own operation logic which performs automatically according to equipment parameters, this problem might be solved fundamentally. Take HVAC system as an example, most of the system could realize remote auto on/off control, automatically increase or decrease equipments has realized, the automatic, reasonable control of the equipment is still far from the expectation. The current structure of intelligent system is complicated and the configuration process is relative complicated, moreover, the long configuration period of hardware and software made the debug cost and requirement relatively high.
- Though some building systems realize positive integration between subsystems, they work totally independently. Enabling different subsystems to work as a whole is essential. In the process of intelligent system implementation, the most difficult things are integration with third parties, system transformation, point configuration, and proofreading. The current performance of building intelligence is still problematic when considering aspects like building energy efficiency, comfort, and security, and thus the smart systems are not sufficiently “smart”.
- The design, construction and configuration cost must decreased then a further application will be realized; A system which could be self-identifying, easy configuration, will be the future trend of intelligent system.

Acknowledgements This work is supported by National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques)

References

1. Chen, Z., Wang, F., Ma, R.: Questionnaire survey and analysis on intelligent system of public buildings. *Intell. Build.* **3**(175), 13–18 (2015). (in Chinese)
2. Chen, Z., Wang, F., Ma, R.: Investigation and analysis on the current situation of intelligent system for public buildings. *Intell. Build.* **3**(175), 23–26 (2015). (in Chinese)
3. Wang, F.: Discussion on technical measures for realizing energy saving effect of intelligent buildings. *Intell. Build.* **11**, 54–58 (2012). (in Chinese)

4. China smart building research white paper (2015). Honeywell.com.cn
5. Alibaba intelligent building white paper (2017). http://www.199it.com/archieves/575702.html?url_type=39

Part IV
Intelligentialization of Heating Ventilation
Air Conditioning System

Regression Model of Wet-Bulb Temperature in an HVAC System



Luping Zhuang, Xi Chen and Xiaohong Guan

Abstract It can result in substantial energy saving in heating, ventilation, and air-conditioning (HVAC) system by improving the control strategy of heating, ventilation, and air-conditioning system. However, it is challenging to obtain the optimal control strategy of an HVAC system due to its model's complexity. In this paper, a regression model is proposed for the wet-bulb temperature which is a key variable in cooling tower and fan coil unit. The proposed model avoids the iterative computing process of obtaining the value of the wet-bulb temperature and reduces the complexity of an HVAC system's model. Numerical results show that the proposed model takes less than 7% computing time to get the value of wet-bulb temperature, and the relative deviations are less than 0.4%, compared to the original model.

Keywords Wet-bulb temperature · Regression model · HVAC system

1 Introduction

Building consumes approximately 20% of the entire energy consumption in China [1]. Among different energy consumption devices in the building, an HVAC system consumes nearly 40–50% of the entire building energy consumption [2]. A small increase in operating efficiency of an HVAC system can result in substantial energy

L. Zhuang · X. Chen (✉) · X. Guan

Center for Intelligent and Networked System, Department of Automation, Tsinghua University, Beijing, China

e-mail: bjchenxi@mail.tsinghua.edu.cn

L. Zhuang

e-mail: zlp12@tsinghua.org.cn

X. Guan

e-mail: xhguan@tsinghua.edu.cn

X. Guan

MOE KLINNS Lab, Xi'an Jiaotong University, Xi'an, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_18

savings. Therefore, it is of great practical interest to formulate the model of an HVAC system and obtain its optimal control strategy.

However, the model of an HVAC system is complex, since it contains several coupled variables and a considerable amount of control variables. In practical application, it is hard to get the optimal control strategy in limited time because of the high complexity of the HVAC system's model. For example, when using the interior point method to obtain the hourly optimal control strategy of the HVAC system, the calculation time always exceeds 1 or even 1.5 h [3]. Thus, it is hard to get hourly control strategy of an HVAC system. To overcome this problem, we propose the regression model of the wet-bulb temperature to avoid the iterative calculation process of obtaining the wet-bulb temperature and reduce the complexity of the HVAC system's model.

The wet-bulb temperature is one of the air's attributes. It is also used as an intermediate variable to compute the other characters of air. Many efforts have been made to use the wet-bulb temperature to evaluate and improve the performance [4–6]. In HVAC system, the wet-bulb temperature also has attracted researchers' attention, since it is an indispensable intermediate variable to compute the heat exchange process between air and water in cooling tower and fan coil unit. Hason analyzes the model of the sensible heat exchanger and obtains the value of wet-bulb temperature [7]. Many different methods have been used to obtain the value of the wet-bulb temperature. First, using a wet-bulb thermometer to measure the wet-bulb temperature is a traditional physical method. However, the cost of measuring equipment is high and the measuring process has the strict request to measurement environment. Then, some researchers propose the methods to get the value of the wet-bulb temperature to save the measuring cost and simplify the computing process. Thorpe proposes the Newton–Raphson method to compute the wet-bulb temperature based on the temperature and humidity [8], and Brent develops the Brent's method to obtain the value of the wet-bulb temperature [9]. However, the iterative computing process enhances the complexity and reduces the computing efficiency. Moreover, some researchers propose simplified computing methods to overcome these difficulties. Zheng et al. use statistical and nonlinear regression analysis to get the relation between the wet-bulb temperature and the dry-bulb temperature [10]. However, the deviation between the values obtained by simplified methods and original model is considerably large.

In this paper, we introduce the original model of the wet-bulb temperature in practical application. Then we propose the regression model of wet-bulb temperature based on the working principle of the wet-bulb thermometer. Compared with the original model, the regression model avoids the iterative computing process and obtains the values of wet-bulb temperature with small deviations.

The rest of this paper is organized as follows. The original model of the wet-bulb temperature is presented in Sect. 2, and the regression model of the wet-bulb temperature is discussed in Sect. 3. Numerical results are illustrated in Sect. 4, and the conclusions drawn are presented in Sect. 5.

2 Original Model of the Wet-Bulb Temperature

2.1 Model Formulation

In practical application, the following variables are used to calculate the wet-bulb temperature.

- (1) Saturated vapor pressure P_{qb}^*

$$P_{qb}^* = \exp\left(\sum_{i=1}^5 c_i (T_{awb})^{i-2} + c_6 \ln(T_{awb})\right) \quad (1)$$

where T_{awb} is the wet-bulb temperature, $c_i (i = 1, 2, \dots, 6)$ are the coefficients.

- (2) Vapor pressure P_q

$$P_q = P_{qb}^* - A(T_{air} - T_{awb})B \quad (2)$$

where T_{air} , B and A represent the air temperature, the atmospheric pressure and the empirical coefficient, respectively.

- (3) Humidity H_{air}

$$H_{air} = 622 P_q / (B - P_q) \quad (3)$$

- (4) Dew-point temperature T_{al}

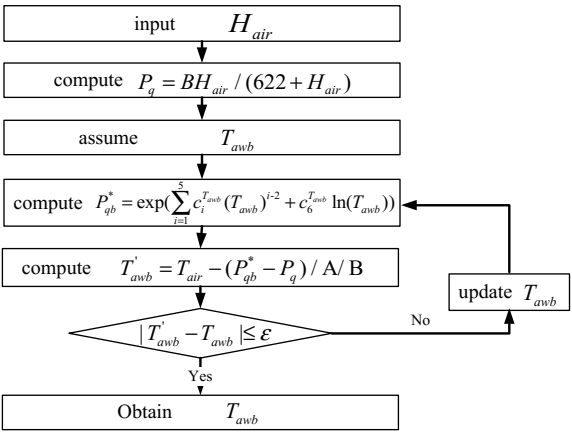
$$T_{al} = \sum_{i=7}^{10} c_i [\ln(T_{awb})]^{i-7} + c_{11} (P_q)^{0.1984} \quad (4)$$

where $c_i (i = 7, 8, \dots, 11)$ are the coefficients.

2.2 Calculation Process

In practical application, the vapor pressure and dew-point temperature are obtained based on the temperature and humidity of air. The supposed value of wet-bulb temperature is obtained based on temperature and dew-point temperature. Then the saturated vapor pressure is obtained with the supposed wet-bulb temperature, and the calculated value is obtained based on vapor pressure and saturated vapor pressure. If the deviation between the supposed value and calculated value of wet-bulb temperature is less than a small enough value, output the supposed value as the result of wet-bulb temperature. Otherwise, update the value of wet-bulb temperature based on the supposed value and calculated value. The calculation process is shown in Fig. 1.

Fig. 1 The calculation process of wet-bulb temperature



3 Regression Model of the Wet-Bulb Temperature

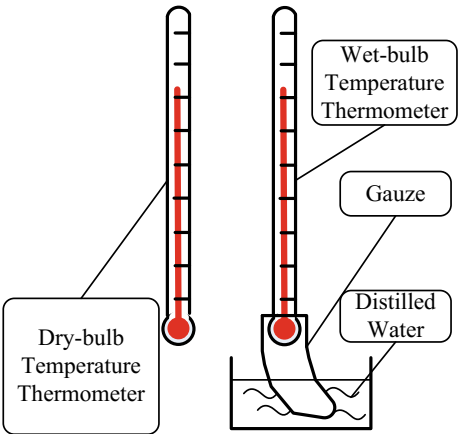
3.1 Generalized Model of Wet-Bulb Temperature

The working principle of the wet-bulb thermometer is shown in Fig. 2.

The wet-bulb thermometer is composed of the dry-bulb thermometer and gauze. The bulb of the dry-bulb thermometer is wrapped up in gauze and the bottom of the gauze immerses in distilled water which keeps the gauze long-term wet.

While ignoring the effect of radiation heat transfer between the surface of the bulb and the surrounding object, keeping the air flow around the surface of the bulb, and maintaining heat and moisture exchange, the heat that the air transfers to the surface of the bulb is equal to the heat which is required to evaporate the water on the surface

Fig. 2 Wet-bulb temperature thermometer



of bulb. Thus, the value of temperature shown on the thermometer is the temperature of water on the surface of bulb, and it is also the wet-bulb temperature of the air.

Hence, the wet-bulb temperature depends on the temperature and humidity of air. We obtain the generalized model of wet-bulb temperature.

$$T_{awb} = \sum_{j=J_{awb}^1}^{J_{awb}^2} \sum_{i=I_{awb}^1}^{I_{awb}^2} b_{ij}^{T_{awb}} (T_{air})^i (H_{air})^j \quad (5)$$

where $b_{ij}^{T_{awb}}$ represents the coefficients, I_{awb}^1 and J_{awb}^1 represent the least degree of T_{air} and H_{air} , respectively, and I_{awb}^2 and J_{awb}^2 represent the highest degree of T_{air} and H_{air} , respectively.

3.2 Regression Model of Wet-Bulb Temperature

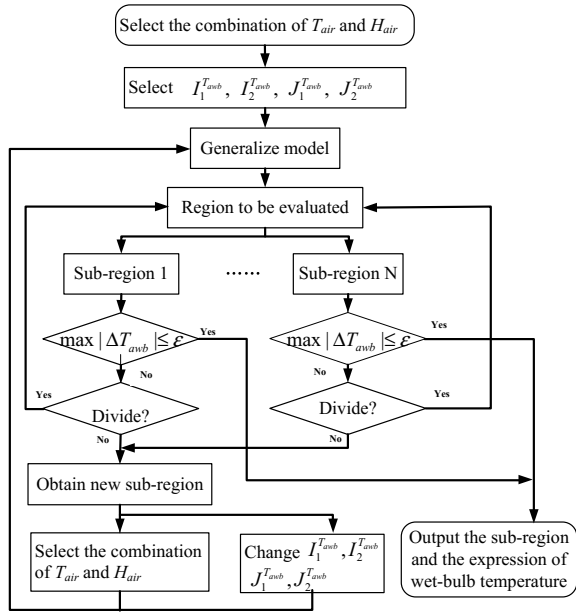
In practical application, we can hardly get the measured data of wet-bulb temperature under all kinds of combinations of temperature and humidity. Therefore, we use the original model in Sect. 2 to predict the wet-bulb temperature under various combinations of temperature and humidity. We use the data obtained by the original model to get the coefficients of regression model and compare the results obtained by the two models.

For the entire region, the deviations obtained by only one regression model exceed the allowable tolerance (<0.04). Thus, we combine zoning method and regression analysis to obtain the regression model of wet-bulb temperature. Figure 3 shows the flowchart of obtaining the regression model, where the region to be evaluated can represent the entire region or subregion.

First, for the regression model, select the least degree, highest degree, and the combination of temperature and humidity, and get the expression of wet-bulb temperature. Obtain the coefficients of regression model and compute the deviations between the results of original model and regression model. Based on the deviations, the entire region is divided into several subregions, and each subregion updates the coefficients of regression model and the deviations. For each subregion, if all the updated deviations are less than a small enough number, output the expression of the wet-bulb temperature as the regression model of the wet-bulb temperature. Otherwise, if the subregion is large enough, divide the subregion into several subregions and update the expression and deviations for further analysis. If the subregion is too small, integrate it with its neighboring subregions which do not obtain their regression model of wet-bulb temperature, update the expression, coefficient, and corresponding deviations. Repeat this process, until obtaining the regression model of wet-bulb temperature of the entire region.

We obtain the regression model of the wet-bulb temperature as follows:

Fig. 3 The flowchart of obtaining the regression model



$$T_{awb} = \sum_{i=1}^7 \left[a_{1,i}^{T_{awb}} (T_{air})^{i-4} + a_{2,i}^{T_{awb}} (T_{air} H_{air})^{i-4} + a_{3,i}^{T_{awb}} (H_{air})^{i-4} + a_{4,i}^{T_{awb}} (T_{air} / H_{air})^{i-4} \right] \quad (6)$$

where $a_{j,i}^{T_{awb}}$ ($j = 1, 2, 3, 4$) is the coefficients, obtained by regression analysis.

The regression model in Eq. (6) is an approximation of the original model in Sect. 2. With different subregions or deviation tolerance, we may get different regression models.

4 Case Study

To show the performance of the model of the wet-bulb temperature in Eq. (6), we take the dry-bulb temperature, ranging from 4 to 43 °C, and humidity, ranging from 5 to 30 g/kg as an example, and use the data obtained by the original model as reference data. To improve the accuracy, we divide the entire domain of the example into nine subregions, based on the temperature and humidity of the air, as shown in Table 1, in which T_{air} and H_{air} are the dry-bulb temperature and humidity of the air, respectively, and $S_i^{T_{awb}}$ ($i = 1, 2, \dots, 9$) is the label of each subregion.

In each subregion, we choose one-third of the entire data in this subregion as the observation data. And other data are used to measure the accuracy of the regression

Table 1 Subregions of the wet-bulb temperature

H_{air} (g/kg)	T_{air} (°C)		
	(4, 17]	(17, 30]	(30, 43]
(5, 14]	$S_1^{T_{awb}}$	$S_2^{T_{awb}}$	$S_3^{T_{awb}}$
(14, 22]	$S_4^{T_{awb}}$	$S_5^{T_{awb}}$	$S_6^{T_{awb}}$
(22, 30]	$S_7^{T_{awb}}$	$S_8^{T_{awb}}$	$S_9^{T_{awb}}$

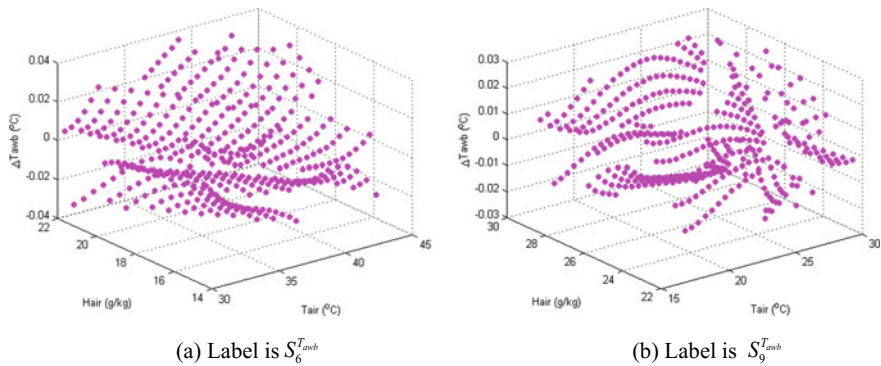


Fig. 4 Results of the wet-bulb temperature

model. With the least square method, we use regression analysis to get the coefficients.

The deviations of wet-bulb temperatures of subregions $S_6^{T_{awb}}$ and $S_9^{T_{awb}}$ obtained by regression model and original model are illustrated in Fig. 4a and b, respectively, where ΔT_{awb} represents deviation of the wet-bulb temperature of the air. In Fig. 4, all the deviations fall in the range $[-0.04, 0.04]$. Thus, the results obtained by the two models are close enough.

To illustrate the difference between the results obtained by regression model and original model in a direct way, we compute the maximal deviation and relative deviation of each subregion and make brief statistical analysis. We use $T_{awb,s}$ and T_{awb} representing the wet-bulb temperature computed by the regression model and original model, respectively. In Table 2, $R_{T_{awb}}^2$ represents the correlation coefficient, $D_{T_{awb}}$ and $RD_{T_{awb}}$ are the deviation and relative deviation between the values obtained by the two models, which are computed by $D_{T_{awb}} = |T_{awb,s} - T_{awb}|$ and $RD_{T_{awb}} = |T_{awb,s} - T_{awb}| / |T_{awb}|$, respectively. $P_{D_{T_{awb}} > 0.03}$ represents the proportion of the number of data whose deviation is larger than 0.03°C to the total number of data, while $P_{RD_{T_{awb}} > 0.0015}$ represents the proportion of the number of data whose relative deviation is larger than 0.15% to the total number of data. From Table 2, we notice that the correlation coefficients of all subregions are more than 0.9994 . So the regression performs well. The deviations of all subregions are less than 0.04°C and for each subregion, less than 5% of the total data have the deviations larger than 0.03°C . The relative deviations

Table 2 Deviations between regression model and original model of wet-bulb temperature

Subregion	$R_{T_{awb}}^2$	$\max D_{T_{awb}}$	$P_{D_{T_{awb}} > 0.03} (\%)$	$\max RD_{T_{awb}} (\%)$	$P_{RD_{T_{awb}} > 0.0015} (\%)$
$S_1^{T_{awb}}$	1.0000	0.0266	0.0000	0.3723	3.8461
$S_2^{T_{awb}}$	1.0000	0.0118	0.0000	0.0867	0.0000
$S_3^{T_{awb}}$	1.0000	0.0146	0.0000	0.0581	0.0000
$S_4^{T_{awb}}$	1.0000	0.0139	0.0000	0.0588	0.0000
$S_5^{T_{awb}}$	1.0000	0.0259	0.0000	0.0983	0.0000
$S_6^{T_{awb}}$	1.0000	0.0388	4.3269	0.1356	0.0000
$S_7^{T_{awb}}$	1.0000	0.0264	0.0000	0.1031	0.0000
$S_8^{T_{awb}}$	0.9998	0.0251	0.0000	0.0823	0.0000
$S_9^{T_{awb}}$	0.9995	0.0295	0.0000	0.0891	0.0000

Table 3 Average calculation time of wet-bulb temperature

Method	Average calculation time(s)
Regression model	0.000237
Original model	0.003538

of all subregions are less than 0.4%, and less than 4% of the data have the relative deviations larger than 0.15%. Therefore, the regression model is good enough.

Now we compare the computing time spent by the regression model and original model. The Matlab (2012a) in the laptop (Intel CoreTM 2 P8600@2.4 GHz) is used to implement the operation of the two models. And the average computing time is shown in Table 3. The regression model uses about 0.00024 s to get the result, while the original model needs almost 0.004 s, which is more than 16 times of regression model. Therefore, from the point of computing time, the regression model is good enough.

In similar way, we can get the regression model of the efficiency of the cooling tower and fan coil unit. The usage of these three regression models in optimal control problem of the HVAC system can avoid the iterative computing process, reduce the complexity of the HVAC system’s model, and decrease the calculation time, compared to regular control [11, 12]. Compared with the regular control strategy obtained by the original model, the improved control strategy obtained by the regression model can save about 34–40% energy consumption [11, 12].

5 Conclusions

In this paper, we introduce the working principle of the wet-bulb thermometer and propose the regression model of the wet-bulb temperature. Compared to the original model, the regression model obtained the accurate enough value of the wet-bulb temperature with less calculation time. In practical application, the regression model decreases the calculation time, reduces the energy consumption, and provides a reasonable way to optimize the operation of the HVAC system.

Acknowledgements This work was supported in part by the National Key Research and Development Program of China (2016YFB0901900 and 2017YFC0704100).

References

1. Building energy research center of Tsinghua University, 2017 Annual Report on China Building Energy Efficiency (in Chinese), China Architecture & Building Press, Beijing(2017)
2. Building energy research center of Tsinghua University, 2010 Annual Report on China Building Energy Efficiency (in Chinese), China Architecture & Building Press, Beijing(2010)
3. Zhuang, L., Chen, X., Guan, X.: A decentralized ordinal optimization for energy saving of an HVAC system. In: American Control Conference (ACC), pp. 611–616. Boston, IEEE (2016)
4. Wilson, S.G., Desmarchelier, J.M.: Aeration according to seed wet-bulb temperature. *J. Stored Product Res.* **30**, 45–60 (1994)
5. Noyes, R., Navarro, S., The mechanics and physics of modern grain aeration management. In: Navarro, S., Noyes, R. (Eds.) *Operating Aeration Systems*. Boca Raton, CRC Press
6. Ranjbaran, M., Emadi, B.: A mathematical model of commodity wet-bulb temperature (CWBT) for grain storage applications. *Biosys. Eng.* **139**, 128–135 (2015)
7. Hasan, A.: Going below the wet-bulb temperature by indirect evaporative cooling: analysis using a modified ε -NTU method. *Appl. Energy* **89**, 237–245 (2012)
8. Thorpe, G.R.: Heat and mass transfer in ventilated bulks of respiring grains—a theoretical analysis and technological application. Melbourne, Melbourne, Australia, Department of Civil and Building Engineering, Victoria University of Technology. Report 1/94 (1994)
9. Brent, R.P.: Table errata: algorithms for minimization without derivatives (Prentice-Hall, Englewood Cliffs, N.J. 1973). *Math. Comput.* **29**(132), 1166 (1975)
10. Zheng, K., Watt, J., Wang, C., Cho, T.K.: Development and implementation of a virtual outside air wet-bulb temperature sensor for improving water-cooled chiller plant energy efficiency. *Sustain. Cities Soc.* **23**, 11–15 (2016)
11. Zhuang, L.P., Chen, X., Guan, X.H.: A bi-level optimization for an HVAC system. *Cluster Comput.* **20**, 3237–3249 (2017)
12. Zhuang, L.P., Yang, Z., Chen, X., Guan, X.H.: Decentralized optimization for energy conservation of HVAC system (in Chinese). *Smart Power* **45**(7), 31–36 (2017)

Research on Optimal Control Algorithm of Ice Thermal-Storage Air-Conditioning System



Junqi Yu, Xiong Yang, Anjun Zhao, Meng Zhou and Yanhuan Ren

Abstract The constraint-based nonlinear multivariate function optimization algorithm was used to optimize the distribution of cooling load between chillers and ice-storage tanks. The goal is to minimize the cooling load and system running costs of the air-conditioning system. Based on the peak-valley price principle of the power grid system, the most economical running of the ice-storage air-conditioning system is achieved. The results show that compared with the traditional ice-storage air-conditioning system control algorithm, the proposed method can reduce the power consumption of the system by 10.32% and reduce the system operating cost by 12.07% under the premise of satisfying the demand for terminal cooling capacity.

Keywords Optimal control · Ice-storage air-conditioning · Building energy conservation

J. Yu (✉) · X. Yang · A. Zhao · M. Zhou · Y. Ren
School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi, China
e-mail: junqiyu@126.com

X. Yang
e-mail: 18691811913@163.com

A. Zhao
e-mail: 928888015@qq.com

M. Zhou
e-mail: zhou_meng0113@163.com

Y. Ren
e-mail: 1564359894@qq.com

J. Yu
Smart City Research Center, Xi'an, Shaanxi, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_19

1 Introduction

With the rapid development of social economy, energy shortage and environmental degradation have become increasingly prominent. The heavy energy burden has become an important obstacle to the sustainable development in China. Building energy consumption accounts for about 30% of the total energy consumption of the entire society [1]. As a powerful means to alleviate this tension, the ice-storage air-conditioning system uses the potential of water/ice phase change potential to store energy, saved the cold at night, and melted ice cold during the daytime, to meet the cooling load demand of the building, not only has the effect of shifting peaks and valleys to the power load of the grid [2–4] but also reduces the running costs of air-conditioning system.

There are four different modes of operation, such as ice-storage mode, chiller single cooling mode, chiller and ice tank combined cooling mode, and ice tank single cooling mode when different electric valves opened, closed, or adjusted. In the ice-storage mode, the cooling water circulation system cuts out the cooling system by adjusting the on–off of the valve, and the refrigeration capacity of the chiller is all used for ice making. The ice inlet and outlet temperature is about -2 to 5 °C during ice-storage conditions. The refrigerant is exchanged and cooled at the evaporator, and then it circulates between the chiller and ice-storage tank to freeze the water to ice. In the chiller single cooling mode, the chillers provide the building with the required cooling load. The ice-storage tank is completely separated from the system by the on–off of the regulating valve. The cooling ethylene glycol solution of the chillers provides low-temperature frozen water to the terminal air-conditioning room through the plate heat exchanger. In the chiller and ice tank combined cooling mode, the chillers are in the cooling mode. The refrigerant that absorbs the ice in the ice tank is mixed with the refrigerant solution cooled by the refrigeration unit, and the heat exchange with the chilled water is completed in the plate heat exchanger. In the ice tank single cooling mode, ice-storage tanks provide cooling to the building. When the electricity price policy is not conducive to the use of chillers cooling, the building's required cooling load is not too high, in order to avoid opening up too many chillers during the peak time of the electricity price or let the chillers in an inefficient operating condition, with excessive air-conditioning operating costs.

Reasonably distributing the cold load between the chillers and the ice-storage tank in the peak-valley price period is the core issue of the optimal control of the ice-storage air-conditioning system. Minimized system operating costs are based on meeting the cooling load requirements [5, 6]. There are three commonly used control methods: chiller priority mode, melting ice priority mode, and optimization control [7]. The chiller priority mode (Fig. 1a) is that the terminal cooling load of the system is preferentially borne by the chiller, and the insufficient part is supplemented by the ice-storage tank. However, this mode cannot effectively play the function of “shifting peak and filling valley” and reducing operation cost of ice-storage air-conditioning system. The chiller priority mode (Fig. 1b) is that the terminal cooling load of the system is preferentially borne by the ice-storage tank and the insufficient

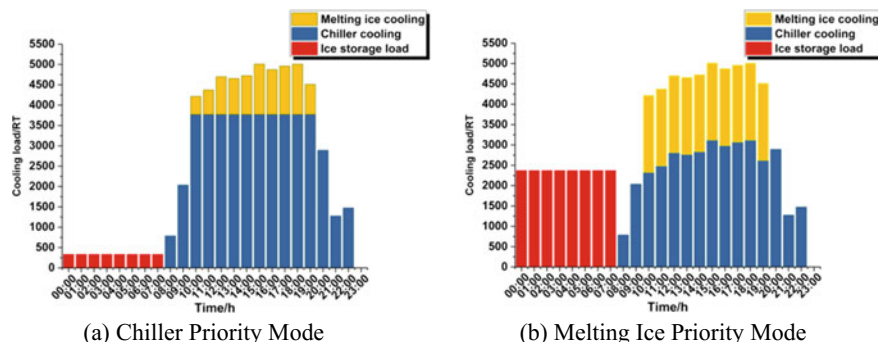


Fig. 1 Ice-storage air-conditioning system operation mode

part is supplemented by the chiller. The main shortcomings of this model are as follows: (1) When the air-conditioning load is large, it is easy to appear more ice melting at the early stage of cooling, and the remaining ice amount is too little in the later period of the system, and the chillers cannot meet the requirement of the load. (2) As the chillers are in a low load state for a long time, the system runs poorly economically. The optimized control strategy refers to using the local electricity price policy to meet the requirements of air-conditioning users and maximize the function of the ice-storage device “shifting peaks and valleys” to minimize the running cost of the ice-storage system [8]. Scholars have used Genetic Algorithm (GA) [9–12] and Particle Swarm Optimization (PSO) [13] to optimize the running mode of ice-storage air-conditioning systems. The algorithm has the following deficiencies: the structural design of the algorithm itself is easy to fall into a local optimum. Second, the output of the genetic algorithm has instability, and the same algorithm parameters and algorithm structure output have large errors, which seriously hinder the intelligent algorithm applied in the actual project.

Aiming at the shortage of the existing control mode, this paper proposes a non-linear multivariate function optimization algorithm based on constraints to achieve optimal control of the ice-storage air-conditioning system. The optimization is based on the minimum cooling load demand and the minimum running cost of the system. The goal is to achieve the most economical operation of ice-storage air-conditioning systems.

2 Ice-Storage Air-Conditioning System Energy Consumption Modeling

The structure of the ice-storage air-conditioning system is complex, and it is difficult to consider the total energy consumption of each component, and the energy consumption of components such as the board swap is very small and can be ignored. So

only the chillers, cooling towers, and pumps that consume large amounts of energy are considered in the paper.

2.1 Notations

$COP(k)$	the coefficient of performance of the k chiller
$PLR(k)$	the partial load rate of the k chiller
$Q_p(i)$	the total cooling load provided by the chiller during the i period
$Q_n(k)$	the rated cooling capacity of the k chiller
$x(k)$	the ratio of the cooling load provided by the k chiller to the total chiller cooling load during this period
$P_{ch}(i)$	the chiller power during the i period
$W_{ch}(i)$	the chiller power consumption during the i period
$W_{ch.day}$	the total chiller power consumption throughout the day
$p_{ct}(i)$	the cooling tower load during the i period
$W_{ct}(i)$	the cooling tower power consumption during i time
$W_{ct.day}$	the total cooling tower power consumption throughout the day
$W_{pump.day}$	the total pump power consumption throughout the day
$W_{ice.day}$	the total power consumption of ice-storage air-conditioning system throughout the day
$P_{n.pump.cw}$	the rated power of cooling water pump
$P_{n.pump.eg}$	the rated power of ethylene glycol pump
f_{un}	the optimal objective function
lb, ub	the upper and lower bound vectors of linear inequality constraints
A, b	the linear and inequality constrained coefficient matrix
$e(i)$	the electricity price during the i period
$Q_{k.ch.max}(i)$	the maximum cooling capacity of the k chiller
$Q_{ice.p.max}(i)$	the maximum cooling provided by ice-storage tanks during i time
$Q_{ice.p}(i)$	the cooling provided by ice-storage tanks during i time
$Q_r(i)$	the hourly cooling requirement at the end of the system during i time
Q_{ice}	the cooling provided by ice-storage tanks throughout the day
$Q_{ice.st.day}$	the storage capacity of ice-storage tank
$Q_{ice.max}$	the maximum capacity of ice-storage tank.

2.2 Chiller Energy Consumption Model

In a system with multiple chillers operating in conjunction, the chillers run under partial load conditions for most of the time. Energy efficiency ratio COP (Coefficient Of Performance) is an index to evaluate partial load performance of chillers [14]. The COP of a chiller is accompanied by the change of partial load rate (PLR). Therefore, how to more effectively optimize the distribution of the load assumed by each chiller

in a joint operation makes the overall COP of the system maximize is very important. The COP of the chiller and its PLR can be expressed as follows:

$$COP(k) = b_1 + b_2 PLR(k) + b_3 PLR(k)^2 \quad (1)$$

$$PLR(k) = \frac{Q_r(i) * x(k)}{Q_n(k)} \quad (2)$$

According to (1) and (2), the power consumption per hour of chiller can be expressed by (3). The total power consumption of all-day can be expressed by (4).

$$W_{ch}(i) = \sum_{k=1}^{23} \frac{Q_r(i) * x(k)}{COP(k)} * 1 \quad (3)$$

$$W_{ch.day} = \sum_{i=0}^{23} W_{ch}(i) \quad (4)$$

2.3 Cooling Tower Energy Consumption Model

Cooling tower power consumption is proportional to the cooling tower load [15]. Therefore, the power consumption per hour can be approximated by (5). Therefore, the total power consumption of the all-day can be expressed by (6).

$$\begin{aligned} p_{ct}(i) &= P_{ch}(i) + Q_r(i) \\ W_{ct}(i) &= 0.025 * p_{ct}(i) * 1 \end{aligned} \quad (5)$$

$$W_{ct.day} = \sum_{i=0}^n W_{ct}(i) \quad (6)$$

2.4 Pump Energy Consumption Model

The study only focused on the primary side of ice-storage air-conditioning, so only cooling water pumps and glycol pumps were studied. The cooling pump is only operated during the chiller's cooling conditions (the working duration is t_1), and the glycol pump is operated in the ice-storage conditions and cooling conditions of the chiller (the working duration is t_2). The pumps are constant power and constant speed pumps, so the energy consumption of all pumps can be expressed by (7).

$$W_{pump.day} = W_{n.pump.ct} * t_1 + W_{n.pump.eg} * t_2 \quad (7)$$

To sum up, the total power consumption of the ice-storage air-conditioning system throughout the day can be expressed by (8).

$$W_{ice.s.day} = W_{ch.day} + W_{ct.day} + W_{pump.day} \quad (8)$$

3 Constraint-Based Nonlinear Multivariate Function Optimization Algorithm

In this paper, a constraint-based fmincon (Find minimum of constrained nonlinear multivariable function) function is used. Its form can be expressed by (9).

$$\min f(x) = fmincon(fun, x0, A, b, Aeq, beq, lb, ub) \quad \begin{cases} A * x \leq b \\ Aeq * x = beq \\ lb \leq x \leq ub \end{cases} \quad (9)$$

Ice-storage air-conditioning system, maximizing the use of night electricity ice storage, can greatly reduce the operating costs, the text to the minimum operating system for the whole day, and solve each time step (hourly) cooling strategy. As shown in (10),

$$f = \text{Min } F_{ice.s.day} = \sum_{i=0}^{23} \{ [W_{ch.day}(i) + W_{ct.day}(i) + W_{pump.day}(i)] * e(i) \} \quad (10)$$

In (10), the peak-valley electricity price is represented by the peak electricity price, the flat-rate electricity price, and the valley electricity price. The peak-valley electricity price in Xi'an is shown in Table 1.

The optimal control constraints for the ice-storage air-conditioning system are as follows:

- (1) The cooling capacity $Q_s(i) * x(k)$ of the k chiller at i hour should be lower than the maximum cooling capacity $Q_{k.ch \max}(i)$ of each chiller.

$$0 \leq Q_p(i) * x(k) \leq Q_{k.ch \max}(i) \quad (11)$$

Table 1 The peak-valley electricity price on Xi'an

Period	Time	Price (RMB/kw h)
Valley period	00:00–8:00	0.45
Flat period	12:00–20:00	0.79
Peak period	8:00–12:00, 20:00–24:00	1.16

- (2) The amount of melting ice $Q_{ice.sup}(i)$ in the i hour of the ice tank shall not exceed the remaining ice capacity of the ice-storage tank $Q_{ice.sup.max}(i)$ and the maximum melting ice capacity of the ice-storage tank at the i hour.

$$Q_{ice.p}(i) \leq Q_{ice.p.max}(i) = b_1 + b_2 \left(\sum_{j=0}^7 Q_p(j) - \sum_{j=8}^{i-1} Q_{ice.p}(j) \right) \quad (12)$$

- (3) The required cooling capacity $Q_r(i)$ of the building is provided by the chillers and the ice-storage tank.

$$Q_r(i) = Q_p(i) + Q_{ice.p}(i) \quad i = 8, 9, \dots, 23 \quad (13)$$

- (4) The total storage capacity of the ice-storage tank is equal to the night storage capacity of the chillers and cannot exceed the maximum storage capacity of the ice-storage tank; In order to prevent the existence of “years of ice,” the ice-storage tank is based on the maximum cooling efficiency of the U-type coil. The amount of ice cooling required for the whole day must be more than 98% of the amount of cold storage.

$$Q_{ice.st.day} = \sum_{i=0}^7 Q_p(i) \leq Q_{ice.max} \quad (14)$$

$$0.98 * Q_{ice.st.day} \leq \sum_{i=8}^{23} Q_{ice}(i) \leq Q_{ice.st.day} \quad (15)$$

4 Case Study

The case study is a large-scale commercial complex. The building has 10 floors, 8 floors above the ground, and 2 floors underground. The height is 40.6 meters, and the area is 200,000 square meters. The ice-storage air-conditioning system of the mall uses a mass storage cold storage operation. The capacity of the ice-storage tank and the chiller is selected according to the all-day cooling load (66537 RT) on a typical design day. It can provide 30% of the daily cooling load for the design. At present, the mall's ice-storage air-conditioning adopts a proportional control method that the chillers and the ice-storage tank, respectively, bear a certain proportion of cooling load at various times. However, in the actual operation process, the cooling load of the air conditioner may fluctuate irregularly under the influence of various factors, and the matching scheme of the chillers and ice-storage tank currently installed in the mall depends on the experience of the technicians' control.

5 Experimental Results

Using the above constraint-based nonlinear multivariate function optimization algorithm to meet the cooling requirement of the air-conditioning system and minimize the operating cost of the system, the data of chiller ice-storage capacity, chiller cooling capacity, and melting ice cooling capacity are obtained in Table 2.

As shown in Table 2, (1) compared to the melting ice priority mode, the optimization mode proposed in the paper at night storage capacity is smaller (ice priority cooling mode per chiller storage capacity is $789.77 \text{ RT} \cdot \text{h}$). (2) By optimizing the distribution of cooling load between the chillers and the ice-storage tank, the cooling load at high electricity prices (08:00–2:00, 20:00–23:00) is completely met by the ice-storage tank or only by one chiller and ice-storage tank meet together to reduce the running cost. (3) During the peak load period (10:00–19:00), except for the cooling capacity provided by melting ice, the three cooling machines evenly share the cooling load. The partial load ratio PLR of the chiller is 0.82–0.93 (corresponding energy efficiency ratio COP is 5.18–5.24), the chiller has low power consumption and provides high system operation efficiency based on the same cooling load.

Figure 2 shows the hourly system operating cost and power consumption of the melting ice priority mode, chiller priority mode, chiller–ice ratio mode (currently used mode) and the optimized control mode proposed in this paper, which can be obtained from Fig. 2: (1) In melting ice priority mode, since the chillers are fully loaded with ice storage in the valley electricity price period (00:00–07:00), the system running cost is high during this period and the ice-storage process involves multiple layers of state transformation and heat dissipation. Therefore, the power consumption of the system is relatively large. In the peak electricity price period (08:00–12:00) and the later stage of system operation (20:00–23:00), due to the adequate amount of ice storage, priority is given to the use of melting ice to meet the cooling load, the system running costs and power consumption are lower. (2) Chiller priority mode has less ice storage in the valley electricity price period, and almost all of the melting ice is used in the high-load period (10:00–19:00) to meet the cooling load, so in the peak electricity price period (08:00–12:00, 20:00–23:00) the unit's high-load running leads to higher costs. (3) Chiller–ice ratio mode distributes the cooling load by 3:7, so the power consumption and cost are lower in the valley electricity price period relative to the melting ice priority mode, and the power consumption and cost of the peak electricity price period are higher. Compared to the chiller priority mode, the power consumption and cost in the valley electricity price period are higher, and the power consumption and costs are lower in the peak electricity price period. (4) The optimization control mode proposed in the paper is lower than the melting ice priority mode in the valley electricity price period and high-load periods. The cost is relatively high only during the period from 20:00 to 23:00. Compared with the chiller priority mode, the power consumption is higher only in the valley electricity price period, and the rest of the operating time is lower.

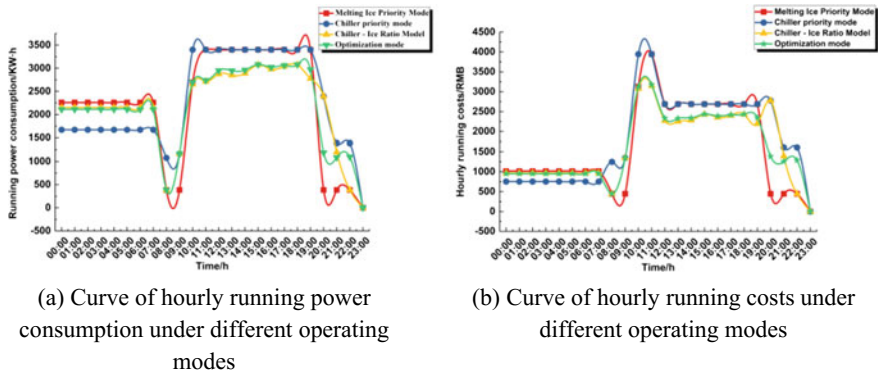


Fig. 2 Curve of hourly power consumption—costs under different operating modes

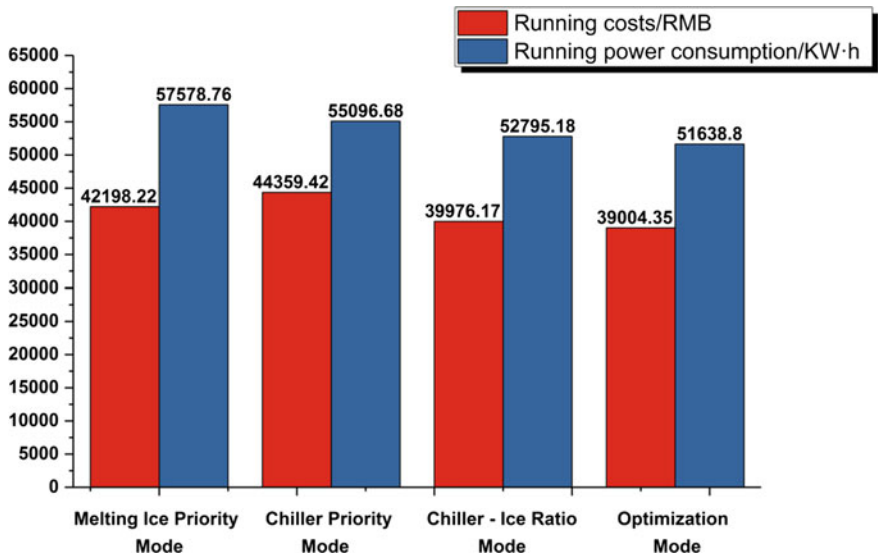


Fig. 3 Power consumption—cost histograms in different operating modes throughout the day

The total electricity consumption and running costs of the above four modes are shown in Fig. 3. Compared to the traditional melting ice priority mode system, the power consumption of optimization control mode proposed in the paper is reduced by 10.32%. Compared to the traditional chiller priority mode, the running cost for the whole day is reduced by 12.07%, which realizes the system’s economic operation throughout the day.

6 Conclusions

Ice-storage air-conditioning is an advanced technology on power demand-side management. Reasonably distributing the cold load between chillers and the ice-storage tank in the peak-valley price period is the core issue of the optimal control of the ice thermal-storage air-conditioning system. In this paper, a nonlinear multivariate function optimization algorithm which is based on constraints is used to implement optimal control of the ice-storage air-conditioning system. Based on the peak-valley price principle of the power grid system, the distribution of cooling load is optimized between the chillers and the ice-storage tank to achieve the most economical operation of the system.

Acknowledgements This work is supported by the National Key Research and Development Project of China with the grant number: 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques) and Xi'an Beilin District Science and Technology Plan Project with the grant number: GX1603 (entitled An Energy-Saving Optimized Operation Strategy for an Ice Storage Air Conditioning System in Xi'an, China).

References

1. Building energy conservation research center, Tsinghua University, Annual Report on China Building Energy Efficiency. Beijing, China Architecture & Building Press (2018)
2. Arteconi, A., Hewitt, N.J., Polonara, F.: State of the art of thermal storage for demand-side management. *Appl. Energy* **93**(5), 371–389 (2012)
3. Sehar, F., Rahman, S., Pipattanasomporn, M.: Impacts of ice storage on electrical energy consumptions in office buildings. *Energy Build.* **51**(51), 255–262 (2012)
4. Lin, H., Lia, X., Cheng, P.: Study on chilled energy storage of air conditioning system with energy saving. *Energy Build.* **79**(4), 41–46 (2015)
5. Yau, Y.H., Hasbi, S.: A Comprehensive case study of climate change impacts on the cooling load in an air-conditioned office building in Malaysia. *Energy Procedia* **143**, 295–300 (2017)
6. Rawlings, L.K.: Ice storage system optimization and control strategies. *ASHRAE Trans.* **91**(2), 12–23 (1985)
7. Arcuri, B., Spataru, C., Barrett, M.: Evaluation of ice thermal energy storage (ITES) for commercial buildings in cities in Brazil. *Sustain. Cities Soc.* **2**(29), 178–192 (2017)
8. Kang, Z.Q., Wang, R.T.: Research status of ice-storage air conditioning system. *Procedia Eng.* **205**, 1741–1747 (2017)
9. Sanaye, S., Shirazi, A.: Four E analysis and multi-objective optimization of an ice thermal energy storage for air conditioning applications. *Int. J. Refrig.* **36**(3), 828–841 (2013)
10. Sanaye, S., Hekmatian, M.: Ice thermal energy storage (ITES) for air conditioning application in full and partial load operating modes. *Int. J. Refrig.* **66**, 181–197 (2016)
11. Sanaye, S., Shirazi, A.: Thermo-economic optimization of an ice thermal energy storage system for air conditioning applications. *Energy Build.* **60**(1), 100–109 (2013)
12. Cai, Z.Y.: Analysis of the principle and operating characteristics of the ice storage air conditioning system. *J. Green Sci. Technol.* **2**, 164–165 (2016)
13. Lee, W.S., Chen, Y., Wu, T.H.: Optimization for ice-storage air conditioning system using particle Swarm algorithm. *Appl. Energy* **86**(9), 1589–1595 (2009)

14. Henze, G.P., Krarti, M.: Guidelines for improved performance of ice storage systems. *Energy Build.* **35**(2), 111–127 (2003)
15. King Dion, J., Potter, Robert A.: Description of a steady-state cooling plant model developed for use in evaluating optimal control of ice thermal energy storage systems. *ASHRAE Trans.* **104**(1), 42–53 (1998)

Decentralized Optimization Algorithm for Parallel Pumps in HVAC Based on Log-Linear Model



Junqi Yu, Xuegen Qian, Anjun Zhao, Shiqiang Wang and Qite Liu

Abstract In order to deal with the issue of lacking universality in the centralized optimization algorithm for parallel pumps, a fully decentralized optimization algorithm based on log-linear model is proposed to complete pumps group operation, consisting of many same type pumps, under the least total power consumption. Through analyzing the characteristics of distributed control system of parallel pumps and the characteristic model of pumps, the difference between centralized optimization model and distributed optimization model of pump system is explained. Introduce the probability model based on log-linear model to calculate the probability distribution on speed ratio space of each pump at each iteration and determine the operating strategy of each pump based on probability distribution at the last iteration. Finally, take a chilled water circulation system in practical project as a study to research and certify the effectiveness of the mentioned algorithm by simulate experiment. The simulation experiment shows that the algorithm can optimize the number and speed of the same type of parallel pumps.

Keywords Decentralized optimization algorithm · Parallel pumps · Log-linear model · HVAC system

J. Yu (✉) · X. Qian · A. Zhao · Q. Liu
School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi, China
e-mail: junqiyu@126.com

X. Qian
e-mail: 2523409905@qq.com

A. Zhao
e-mail: zhao_anjun@163.com

Q. Liu
e-mail: 807973812@qq.com

S. Wang
School of Defense Engineering, Army Engineering University of PLA, Nanjing, Jiangsu, China
e-mail: swiptwsq@163.com

1 Introduction

In recent researches on optimal operation for pump system, most researchers focused on the study of pump scheduling optimization. Koor [1] proposed a new algorithm based on Levenberg–Marquardt algorithm only to optimize the number of pumping stations to reduce the energy consumption of pump operation. Pawel Olszewski [2] studied the importance of different target functions, concerning minimization of power consumption, and minimization of excess flow rate and maximization of overall pumping system efficiency, utilizing the genetic algorithm to optimize four pumps in parallel mode by simulation experiment. The experiment shows that minimization of power consumption, selected as the objective function of optimization, could improve the accuracy of the algorithm. Through monitoring real pump station system, an optimization model for four different sewage pumps based on the method of data mining is presented and its experiment result showed that the method of optimization algorithm saves energy 6–14% [3]. Viholainen [4] proposed a dynamic algorithm based on programmable logic controller to improve the efficiency of pump in the automatic control system of parallel pumps. However, the mentioned algorithms are developed on the basis of centralized control system architecture, and applied to special pump system.

A new decentralized intelligent control system for buildings is proposed by many researchers in building energy conservation research center of Tsinghua University [5]. Its algorithm has the characteristic of no center, self-organization, and extendibility. Therefore, there are many scholars making a great contribution to this field. Dai [6] put forward a kind of decentralized optimization algorithm of HVAC system, which provides a new idea to optimize the energy consumption of the chilled station in HVAC system. Wang [7] proposed a decentralized algorithm of sensor fault diagnosis by introducing the energy function in Hop-field network.

In this paper, the optimal allocation problem of the number and speed of the pump is solved by using a decentralized optimization algorithm with the least total power consumption under load demand and system constraint. First, the characteristic of decentralized control system of parallel pump is introduced simply. Second, the decentralized optimization control problem of parallel pump is defined, which mainly includes model of parallel pump and fitness function of optimization algorithm. Third, a parallel pump decentralized optimization algorithm based on log-linear model is elaborated in detail. Finally, take a chilled water circulation system in practical project, including four pumps in parallel mode, as a study to research and certify the effectiveness of the mentioned algorithm by simulating experiment compared to the result of literature [9].

2 System Model and Problem Definition

2.1 Notations

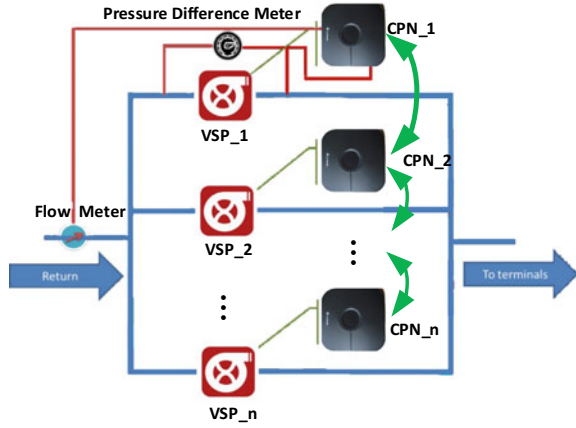
S	the resistance of pipe network;
Q_{met}	the total flow required for the system;
n_o	the rated speed of water pump;
n	the actual running speed of the pump;
w_i	the speed ratio of the i th pump;
H_i	the working pressure head of the i th pump;
η_i	the working efficiency of the i th pump;
W_i	the output power of the i th pump;
Q_i	the workflow of i th pump;
W	the output power of parallel pump system;
$a_i, b_i, c_i, j_i, k_i, l_i$	the performance constant of the pump;
Q_i^{min}, Q_i^{max}	the upper and lower limits of the working flow rate of the pump;
N	the number of parallel pumps;
Ob_i	local fitness function;
φ_i	the local penalty function;
∂	the penalty function;
$p(y x)$	the conditional probability of output y when input x ;
$p_i(k)$	the probability of the i th pump when the algorithm is iterated at k ;
μ	the learning parameters of the algorithm;
S_i	the strategy set of the i th pump operation;
S_{best}	point of best running.

2.2 Pumps Model and Its Decentralized Control System

Figure 1 is shown as a schematic diagram of a typical parallel pump and its decentralized control system. It is known that the traditional central controller is replaced by many intelligent units which are named as computer process node (CPN). CPN can also be embedded in each device to form an intelligent pump. Intelligent pump connects with a neighbor pump with a flow meter on the main pipe and a sensor of pump working pressure on the two ends of pump by wired or wireless transmission. Each intelligent pump cooperates with adjacent water pump to accomplish the computing task together.

When the differential pressure difference H_o is not equal to H_{set} , by the pressure difference, sensor monitors the pressure difference between the two ends of the pump. In order to maintain the constant pressure difference H_{set} of the system, it is to meet the demand of the end demand flow Q_{met} and adjust the speed of the pump to meet

Fig. 1 Schematic diagram of the structure of decentralized control system for parallel pump



the demand. The relation between the flow and the pressure head in the network system of parallel pump work is given in (1) [8].

$$H_o = S \cdot Q_o^2, H_{set} = S \cdot \left(\sum Q_i \right)^2 = S \cdot Q_{met}^2 \quad (1)$$

The centrifugal pump is the most common type of chilled water circulating pump, and its lift and flow model can be expressed as $H = aQ^2 + bQ + c$. The model of efficiency and flow can be expressed as $\eta = jQ^2 + kQ + l$. Among them, a, b, c, j, k , and l are pump performance parameters, and their specific values are obtained from experimental measurements and polynomial fitting.

Among them, a, b, c, j, k , and l are pump performance parameters, and their specific values are obtained from experimental measurements and polynomial fitting.

According to the similarity rate of the pump [9],

$$\frac{Q(n)}{Q(n_0)} = \frac{n}{n_0}, \frac{H(n)}{H(n_0)} = \left(\frac{n}{n_0} \right)^2, \frac{\eta(n)}{\eta(n_0)} = 1 \quad (2)$$

If the pump speed ratio of the pump is defined as w_i :

$$w_i = \frac{n}{n_0} \quad (3)$$

The model of parallel pumps at any speed can be expressed as

$$H_i = a_i Q_i^2 + b_i w_i Q_i + c_i w_i^2 \quad (4)$$

$$\eta_i = j_i \left(\frac{Q_i}{w_i} \right)^2 + k_i \left(\frac{Q_i}{w_i} \right) + l_i \quad (5)$$

$$W = \sum_{i=1}^n \frac{\rho g H_i Q_i}{\eta_i} \quad (6)$$

The optimization problem of parallel pump can be described as the combination of the minimum running pump speed when the required flow rate Q_{met} is required for the system. The mathematical description of the optimization problem is shown as in (7):

$$\begin{aligned} &\text{Min. } W \\ &s.t. \quad Q_i^{\min} < Q_i < Q_i^{\max} \\ &\sum_{i=1}^N Q_i = Q_{met}, \quad i = 1, 2, \dots, N \end{aligned} \quad (7)$$

The parallel pump optimization problem described by (7) is a centralized optimization model, which requires a central controller to process global information. And each intelligent node is only based on local information (mainly communication with the neighbor nodes) to optimize the task of load distribution without a central controller. It runs in an intelligent pump controller, and each intelligent pump runs the same algorithm.

The above optimization problem can be defined as the decentralized optimization problem on each intelligent pump.

$$\begin{aligned} &\text{Min. } W_i(Q_i) \\ &s.t. \quad Q_i^{\min} < Q_i < Q_i^{\max} \\ &\sum_{i=1}^N Q_i = Q_{met}, \quad i = 1, 2, \dots, N \end{aligned} \quad (8)$$

In fact, the optimization target described by Eq. (8) is only the sub-objective function of Eq. (8), and $W_i(Q_i)$ is obtained only through the adjacent intelligent pump controller communication by the intelligent pump controller. The fitness function of the decentralized optimization control problem of the parallel water pump is set up in (9).

$$\text{Min. } Ob_i = W_i + \varphi_i = W_i + \partial \cdot \left(\sum_{i=1}^n Q_i - Q_{met} \right)^2 = W_i + \partial \cdot (\Delta Q)^2 \quad (9)$$

φ_i is a penalty function and the ∂ is a penalty coefficient. When the value of the ∂ is high, the algorithm is premature, and the low value of ∂ will lead to the decrease of the precision and even the unfeasible solution. Literature [10] points out that ∂ should take a smaller value at the beginning of the algorithm, and ∂ should take a

larger value in the later stage of algorithm search. Therefore, according to the Gauss function [11], it can establish Eq. (10) in the iterative process of algorithm.

$$\partial(t) = \partial_0 \cdot e^{(-t^2/\sigma^2)} + 1 \quad (10)$$

Equation (9) shows that, for each intelligent pump in its feasible operation strategy space, a set of optimal rotational speeds can be determined, and the energy consumption of parallel pumps is minimized when the terminal flow demand is met. Therefore, the key issue is how to design a global optimization function only through local information interaction. In this paper, any method that can solve problems (8) and (9) becomes a decentralized optimization algorithm.

3 Decentralized Optimization Algorithm Based on Log-Linear Model

Compared to the probability optimization algorithm, the probability decision rule algorithm can not only avoid the random migration that appears in the heuristic algorithm but also can guide the algorithm search. Therefore, a decentralized optimization algorithm for parallel pumps based on log-linear model is proposed under the decentralized control network of parallel water pump shown in Fig. 1.

3.1 Log-Linear Model

The log-linear model (maximum entropy model) is one of the most widely used models in the neuro-linguistic program (NLP) field. Model training is usually based on maximum likelihood criterion, and its goal is to get the feature weights that maximize the conditional likelihood of training data. This model is represented as [12]

$$p(y|x) = \frac{e^{\sum_{i=1}^n w_i f_i(x,y)}}{\sum_y e^{\sum_{i=1}^n w_i f_i(x,y)}} \quad (11)$$

Combining Eqs. (9) and (11), we can define the updated equation for evaluating the probability distribution of each intelligent pump operation strategy set:

$$p_i(k) = \frac{1/e^{\mu\{W_i[x_i, x_j(k-1)] + \varphi_i[x_i, x_j(k-1)]\}}}{\sum_{s_i \in S_i} 1/e^{\mu\{W_i[s_i, s_j(k-1)] + \varphi_i[s_i, s_j(k-1)]\}}} \quad (12)$$

In the upper form, k is the number of iterations, and the greater the value of learning parameters, the more intelligent pumps tend to take the best path to search.

Therefore, it is suggested to take smaller values at the beginning of algorithm and increase the value of μ as the number of iterations increases.

In order to ensure the real-time performance of the parallel water pump system, the operation strategy space of each pump can be discretized in combination with the actual situation. Then the operation strategy of the i th pump is defined as $S_i = \{s_i^1, s_i^2, s_i^3, \dots, s_i^m\}$, $i \in \{1, 2, \dots, N\}$ and s_i^1 are the lower limit of operation strategy set for the i th pump. s_i^m is the upper limit of the operation strategy set for the i th pump.

3.2 Decentralized Optimization Algorithm

The convergence-based probabilistic decision rule algorithm can be explained as follows: the best optimization result is defined as s_{best} . It can be obtained as

$$s_{best} = \arg_{s \in S} \min ob_i \quad (13)$$

When the μ is big enough, it can be obtained as

$$\exp\{\mu[W_i(s_{best}) + \varphi_i(s_{best})]\} \ll \exp\{\mu[W_i(s_n) + \varphi_i(s_n)]\}, \forall s_n \in S \setminus s_{best} \quad (14)$$

It can be obtained as

$$1/\exp\{\mu[W_i(s_{best}) + \varphi_i(s_{best})]\} \gg 1/\exp\{\mu[W_i(s_n) + \varphi_i(s_n)]\}, \forall s_n \in S \setminus s_{best} \quad (15)$$

The combination of Eqs. (13) and (15) can be obtained as

$$\lim_{\mu \rightarrow \infty, k \rightarrow \infty} p_i^{s_{best}}(k) = \lim_{\mu \rightarrow \infty, k \rightarrow \infty} \frac{1/e^{\mu\{W_i[x_i, x_j(k-1)] + \varphi_i[x_i, x_j(k-1)]\}}}{\sum_{s_i \in S_i} 1/e^{\mu\{W_i[s_i, s_j(k-1)] + \varphi_i[s_i, s_j(k-1)]\}}} = 1 \quad (16)$$

The above equations show that the algorithm based on log-linear probabilistic decision rule converges to s_{best} according to probability, that is to say, in this case minimizing objective function is guaranteed by a large probability.

In the iterative process, if the r th operation strategy has a greater contribution to the minimum value of the target function than the other policies, the probability value of the strategy is greater than that of the other strategies. The update rule of the optimization algorithm: with the increase of the number of iterations, the probability of any strategy corresponding to any strategy of the N intelligent pump running strategy (its value is close to 1) is far greater than the other values, and then the iteration stops.

The decentralized optimization algorithm of parallel pumps based on log-linear model is described as follows:

- Step01:* Set the iteration number $k = 0$ and equal probability to select the initial operation strategy $S_i(0)$ from the operation strategy set S_i .
- Step02:* According to the probability distribution $p_i(k)$ corresponding to the operation strategy set S_i , the initial probability $1/m$ of $p_i(k)$ is selected.
- Step03:* In the k iteration, the local load difference is calculated by using the strategy $S_i(k)$ selected in *Step02* and transmitted to the neighbor intelligent pump at the same time.
- Step04:* After receiving the difference between the adjacent pump loads, the fitness values including all the running strategies are calculated, and the probability distribution is updated by (12).
- Step05:* If the number of iterations reaches the maximum number set, the algorithm terminates or returns to *Step02*

4 Case Study

4.1 Experimental Setup

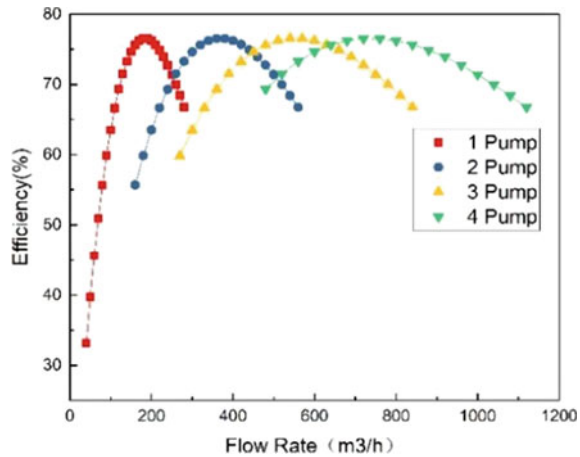
Take the chilled water circulation pump system of a public building as the research object. The chilled water circulating pump group is composed of four parallel pumps with the same size, the type of pump is ISW 150-315 (impeller radius: 335 mm), its rated flow rate is $200 \text{ m}^3/\text{h}$, rated head is 32 m, and rated efficiency is 0.78. The control method of the pump group is the pressure difference control. Performance parameters of each pump are identified by the actual operation data. As shown in Table 1, the parallel pump coefficient of the same size is similar. The design flow of chilled water is $783 \text{ m}^3/\text{h}$, and the design head is 30 m.

According to the working characteristics of the parallel pump, when the four pumps are in parallel, the total flow rate of the water pumps is summed up and the working pressure of each pump is equal. In theory [8], the system has the highest efficiency when all pumps of parallel pump station speed are equal. Therefore, the efficiency curve of the different number of parallel pump system can be drawn easily. Under the given terminal flow demand, it can be obtained from the diagram that the

Table 1 List of characteristic parameters of all pumps

Parameter	1. ISW 150-315	2. ISW 150-315	3. ISW 150-315	4. ISW 150-315
a	−0.000241	−0.000253	−0.000231	−0.000240
b	0.02332	0.02400	0.02373	0.02298
c	36.51	36.01	36.35	36.68
j	−2.03E−05	−2.01E−05	−2.019E−05	−2.028E−05
k	0.007778	0.007889	0.007008	0.007438
l	0.04197	0.04110	0.04207	0.04199

Fig. 2 Efficiency curves of running numbers of parallel pump with same type at head $H = 30$ m



system finally runs several pump systems, and then the pump speed can be calculated according to the characteristic curve of the pump. Therefore, this conclusion can be used to verify the effectiveness of the decentralized optimization algorithm for multiple parallel chilled pumps proposed in this paper.

The verification scheme adopts 25, 45, 65, 85, and 100% of the design flow separately as the testing conditions of the algorithm. According to Fig. 2, it can be determined that the number of pumps running in this system is 1, 2, 3, 4, and 4, respectively. This conclusion is used to verify the effectiveness of the proposed algorithm.

4.2 Results and Analysis

In order to ensure the real-time performance of the algorithm, the feasible strategy space of the pump can be discretized to form a finite set according to the actual situation. According to the pump characteristic curve, the lower speed limit of this type of pump can be calculated to be 0.915 under the condition of maintaining the working pressure head $H = 30$ m. Otherwise, the working pressure head cannot meet the requirements of the working pressure head. Therefore, the pump operation strategy is set to $S_1 = \{0, 0.915, 0.925, 0.935, 0.945, 0.955, 0.965, 0.975, 0.985, 1\}$. And because each pump is the same type, so the strategy collection of the four pumps is the same.

In this case study, when the intelligent pump will detect the pressure difference H_o that is not equal to H_{set} , the algorithm starts to execute. The Q_o can be obtained by the flow meter on the chilled water pipe, and Q_i can be solved by (1). The decentralized optimization algorithm can be downloaded into every intelligent pump and can be activated by itself. Each intelligent pump can communicate and cooperate with other

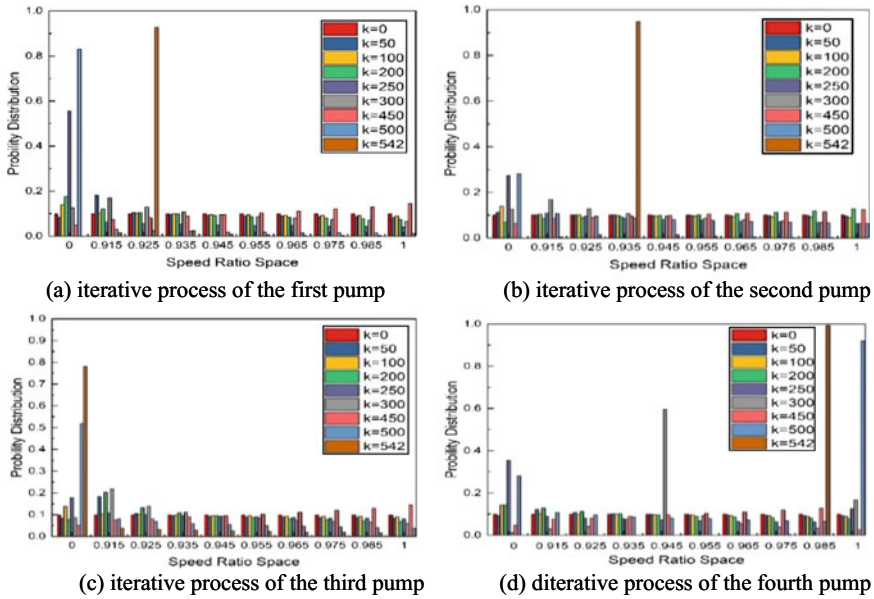


Fig. 3 The iteration process of the decentralized algorithm of parallel pumps when the test condition is $509 \text{ m}^3/\text{h}$

intelligent pump to optimize the corresponding assignment. Figure 3 is the probability distribution process of each pump running strategy set under different iteration times when the terminal demand flow is $509 \text{ m}^3/\text{h}$.

As shown in Fig. 3, the probability distribution difference on running strategy set of each pump gradually increases with algorithm iterating when the air-conditioning terminal demand flow is $509 \text{ m}^3/\text{h}$. From Fig. 3, the final optimization result is that opening 1#, 2#, 4# pump, respectively, and closing the 3# pump in the chilled water circulating pump system. This is consistent with the expected results in the verification scheme. It shows that the algorithm can optimize the number of parallel pumps under the test condition.

Similarly, the simulation results of other operating condition are recorded in Table 2. It can be seen that the algorithm basically meets the requirement of air-conditioning terminal flow, and the efficiency of each pump is high. At the same time, the optimization results are consistent with the expected results in the experimental scheme, which shows that the algorithm can realize the optimization of the number of parallel pumps.

After all test conditions, it can be found that the algorithm can optimize the number of parallel pumps, but the speed optimization results of the pumps in each platform in Fig. 3a–d are 0.925, 0.935, and 0.985, respectively, which are different from the conclusions given in literature [8]. The reason is that the granularity of the pump operation strategy set is coarse when simulating.

Table 2 The results under typical working conditions

PUMP	Test condition and result	25%	45%	65%	85%	100%
1#PUMP	n Hz	0	46.75	50	47.25	47.25
	Qm ³ /h	0	145.1	219.5	159.1	171.8
	η	0	0.760	0.771	0.776	0.784
2#PUMP	n Hz	0	0	46.75	49.25	47.75
	Qm ³ /h	0	0	145.1	204.9	171.8
	η	0	0	0.760	0.781	0.784
3#PUMP	n Hz	48.75	0	0	47.75	50
	Qm ³ /h	194.5	0	0	171.8	219.5
	η	0.786	0	0	0.784	0.771
4#PUMP	n Hz	0	49.25	46.75	46.25	50
	Qm ³ /h	0	204.9	145.1	129.1	219.5
	η	0	0.780	0.760	0.732	0.771

5 Conclusions

In this study, a fully decentralized control strategy has been applied to the parallel pump systems with same type in which the traditional master controllers have been substituted by independent smart node embedded in each pump. First, a new fully decentralized mathematical model for optimal distribution of parallel pump and decentralized flat architecture is presented. Second, decentralized optimization algorithm based on log-linear model is used to solve the parallel pump optimal load distribution. In order to achieve effectiveness of global goals, the algorithm can optimize number of opening pumps and speed ratio of each pump without a central controller when knowing the load requirement of terminal by many sensors. Finally, taking a chilled water circulating pump system as an example, the algorithm is verified by the 25, 45, 65, 85, and 100% of design flow of the pump system. The results show that the algorithm can satisfy the demand of the terminal flow, optimize the number of running stations and speed ratio of the same type parallel pump, and ensure the system running in the high-efficiency interval. However, the iteration time is longer and the number of iterations is larger. Therefore, it needs further improvement.

Acknowledgements This work is supported by National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques).

References

1. Koor, F., Vassiljev, S., Koppel, T.: Optimization of pump efficiencies with different pumps characteristics working in parallel mode. *Adv. Eng. Softw.* **101**, 69–76 (2016)
2. Olszewski, F.: Genetic optimization and experimental verification of complex parallel pumping station with centrifugal pumps. *Appl. Energy* **178**, 527–539 (2016)
3. Zhang, F., Kusiak, S., Zeng, T.: Modeling and optimization of a wastewater pumping system with data-mining methods. *Appl. Energy* **164**, 303–311 (2016)
4. Viholainen, F., Tamminen, S., Ahonen, T.: Energy-efficient control strategy for variable speed-driven parallel pumping systems. *Energ. Effi.* **6**(3), 495–509 (2013)
5. Shen Qi, F.: An analysis of the optimization disaggregation algorithm in the estimation related to energy consumption of appliances in buildings. Elsevier Science Inc. (2014)
6. Dai, Y.F., Jiang, Z.S., Shen, Qi T.: A decentralized algorithm for optimal distribution in HVAC systems. *Build. Environ.* **95**, 21–31 (2016)
7. Wang, Y.F., Jang, Z.S., Xing, J.T.: Decentralized algorithm for sensors fault detection in HVAC system. *J. Refrig.* **37**(2), 30–37 (2016)
8. Koor, M.F., Vassiljev, A.S., Koppel, T.: Optimal pump count prediction algorithm for identical pumps working in parallel mode. *Procedia Eng.* **70**, 951–958 (2014)
9. Zhang, X.F., Liu, J.S., You, Z.T.: Study on the influence of distribution lines to parallel inverter systems adopting the droop control method. *J. Power Electron.* **13**(4), 701–711 (2013)
10. Reardon, L.F.: Particle swarm optimization algorithms with decreasing inertia weight based on gaussian function. *Appl. Res. Comput.* **29**(10), 3710–3712 (2012)
11. Werfel, J.F., Petersen, K.S., Nagpal, T.: Designing collective behavior in a termite-inspired robot construction team. *Science* **343**(6172), 754–758 (2014)
12. Luan, X.F., Li, Z.S., Liu, T.T.: A novel attribute reduction algorithm based on rough set and improved artificial fish swarm algorithm. *Neurocomputing* **174**, 522–529 (2016)

Partial Fault Detection of Cooling Tower in Building HVAC System



Liangliang Sun, Haiqi Jia, Hang Jin, Ye Li, Junning Hu and Congxin Li

Abstract The high false alarm rate and the difficulty of modeling are the main problems in the field of cooling tower system fault detection which is an important energy consumption optimization method in heating, ventilation, and air-conditioning (HVAC) system. This paper proposes an effective solution that is used to reduce the false alarm rate and built a gray box model which simplified from the physical principle of a cooling tower. The Kalman filter is used to forecast the running state of the cooling tower system, and the dynamic control limit set by the statistical process control (SPC) is used to reduce the false alarm rate. Through the final experimental results in the Sino-German building, located in the northeastern part of China, it can be seen that the control limit can be effectively adjusted according to the fluctuation of the natural environment, and the false alarm rate can be well controlled.

Keywords HVAC system · Cooling tower · Fault detection · Kalman filter · SPC rule

L. Sun (✉) · H. Jia · H. Jin · Y. Li · J. Hu · C. Li
Faculty of Information and Control Engineering, Shenyang Jianzhu University, Shenyang 110168, China

e-mail: swinburnsun@163.com

H. Jia

e-mail: 1223450822@qq.com

H. Jin

e-mail: 1182223197@qq.com

Y. Li

e-mail: 1015941310@qq.com

J. Hu

e-mail: 1191018813@qq.com

C. Li

e-mail: 1095038606@qq.com

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_21

1 Introduction

Owing to the growing serious problem of energy shortage, people are increasingly focusing on the energy sector. The increasing demands to fulfill a modern lifestyle have led to the widespread use of HVAC systems around the world; the energy consumption of an HVAC system is higher than can be imagined. Presently, the building energy consumption constitutes about 33% of world energy consumption, whereas the HVAC system energy consumption constitutes about 40% of the building energy consumption [1].

The consumption due to the faulty mechanism of an HVAC system constitutes a large proportion of the consumption of the entire HVAC system, and this can be avoided by the fault detection method. Owing to the waste of energy caused by the fault in an HVAC system, the energy consumption is considerably higher than the normal demand of an HVAC system in a commercial building. If the system can always run in a fault-free condition, it will require less energy. However, a fault in an HVAC system can be caused by many reasons. A fault in an HVAC system can be broadly classified as an electrical system fault, pipe system fault, mechanical system fault, or valve system fault. The faults cause a reduction of comfort and an increase in energy consumption for the building. The faults in an HVAC system cause the various factors of an indoor environment to deviate from their target values. Timely detection of the faults in an HVAC system can significantly reduce unnecessary energy consumption, and the cost of energy consumption in daily life can be reduced. The main components of the HVAC system are the cooling tower, chiller, and air handler unit. Among them, the cooling tower is directly related to the HVAC system cooling efficiency. The cooling tower can directly influence our life; it is a very important subsystem. This study mainly focuses on the fault detection of the cooling tower.

The first difficulty in a cooling tower fault detection is how to reduce the impact caused by outdoor environmental factors. Most buildings are constructed in open natural environment, which is governed by time-varying conditions, and hence the parameters of construction vary all the time. The first problem in fault detection is how to minimize the interference caused by fluctuating environmental factors and ensure the accuracy of cooling tower fault detection.

Another difficulty is how to increase the feasibility of cooling tower fault detection. Under normal circumstances, commercial buildings cannot provide enough historical data for analysis. If a cooling tower is run for considerable time before installation in order to get enough historical data, many resources will be consumed, and the life of the cooling tower system will be reduced. The second problem that we need to solve in fault detection is the estimation of accurate data from limited available original data.

2 Development Status

The fault detection method has developed more systematically in recent years. The fault detection methods of the HVAC system are mainly developed for the optimal system control operation, timely detection, and low fault alarm rate. Additionally, the fault detection method can also reduce the energy consumption of the system and extend the service life of the system equipment. It can improve user comfort by reducing the probability of a fault occurring at the origin itself. Most fault detection methods follow a similar pattern: The control system estimates the running state of the HVAC system at the given time through the measured value of the HVAC system. The running state is then entered into the fault detection system, which is established by the designer. If the state does not meet the estimated rule, it will be set as a fault [2].

The complexity of an HVAC system gives rise to many impact factors. This causes an increase in the complexity of the HVAC system fault detection method; therefore, it may be required to detect the fault for a long time and it is difficult to ensure the detection accuracy. To solve these problems, many fault detection methods for the HVAC system are developed.

The first fault detection method is rule-based. This method is easy to maintain and can promote all kinds of appearing rules; however, its drawback is that the system does not recognize an undefined fault. The second fault detection method is model-based. The advantage of this method is that the system can be accurately detected at all levels; however, its disadvantage is that it is difficult to build the model of the target system. The third method is based on the fault tree. The advantage of this method is that the fault search should be complete; however, the disadvantage is that the fault tree is very complex and needs considerable time to improve the detection tree. The fourth fault detection method is the case-based method. The advantage of this method is its fast detection speed, and the disadvantage is that it needs a large number of cases to build the database. The fifth method is based on the neural networks. The advantage of this method is that it can adaptively form a decision area through the historical data, and the disadvantage is that the curve is difficult to converge [3–5].

In the proposed method of our research work, a gray model is established by the physical principles of the target system, which is then simplified by a physical model and the parameters that can reflect the physical relationship between the input and output are selected as fault indicators. Further, the state estimation of the target system is calculated by the Kalman filter method, and dynamic control limit is established according to the SPC rule. Finally, if the state estimation exceeds the dynamic control limits for a certain time period, the state at the then current time will be considered as a fault. In this method, the Kalman filter can obtain a more accurate state estimation value with less historical data, and the dynamic control limits constructed by the SPC control rule can reduce the influence of the fluctuation of natural factors to a large extent. The combination of these two methods can greatly increase the accuracy and simplicity of the HVAC fault detection and reduce the false alarm rate of the fault detection system.

3 Model Description

The research object in this study is the HVAC system of the Sino-German building of the Shenyang Jianzhu University, which is located in Shenyang, China.

In winter, the ground source heat pump system and the air source heat pump system work together during the day time, and excess heat is stored in the phase change tank.

During the night, the ground source heat pump system continues to heat the building, and the phase change tank releases heat for heating the building as an auxiliary device. In summer, only the ground source heat pump works. The details of interaction between the cooling towers, chillers, and air handling units are shown in Fig. 1.

In this study, a gray box model of the cooling tower was developed. The gray box model is a simplified form of the physical model, which reflects the physical relationship between the input and output; the parameters in the model have their own physical significance [6].

In summer, the ground air heat pump does not participate in indoor temperature regulation owing to the weather, and the work of the cooling tower is mainly borne by the buried coil. We obtained the following mathematical heat transfer model of the ground air heat pump which is suitable for this study [7]:

$$AK_c \Delta t_m = a_1 \frac{v_{cw}}{d} (T_{cwo} - T_{cwi}) C_{cw} + a_2 \frac{\lambda}{de} (T_{ao} - T_{ai}) Re^{0.45} \quad (1)$$

where a_1 and a_2 are the parameters of cooling tower, and the parameter a_1 reflects the effect of the ground heat pump on the heat transfer coefficient K_c , where a_2 is the effect of air heat pump on the heat transfer coefficient. v_{cw} is the flow rate of cooling water in ground heat pump, d is the internal diameter of cooling water pipe, T_{cwo} is

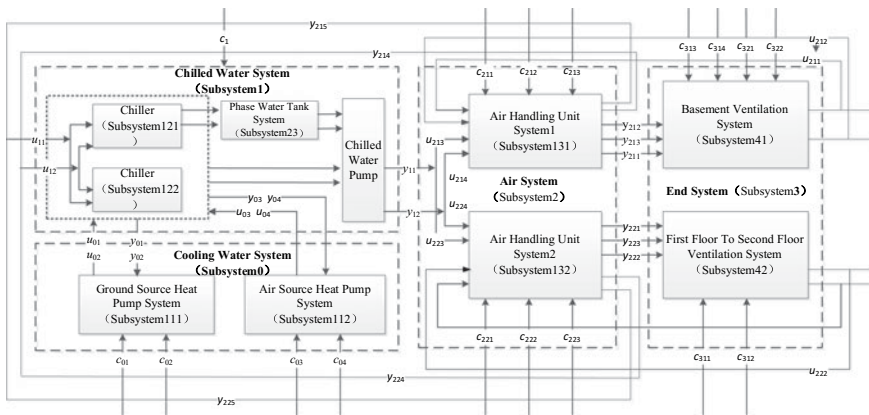


Fig. 1 HVAC system in the Sino-German building [6]

the outlet cooling water temperature, T_{cwi} is the inlet cooling water temperature, C_{cw} is the specific heat capacity of cooling water, λ is the average thermal conductivity of air in the air heat pump, de is the equivalent diameter of air channel in air heat pump, T_{ao} is the outlet air temperature in air heat pump, T_{ai} is the inlet air temperature in air heat pump, Re is Reynolds number, and Re is as follows:

$$Re = v_a de / \gamma \quad (2)$$

where γ is the average air kinematical viscosity and v_a is the flow rate of air.

And in Eq. (1), A is the heat transfer area of cooling tower, Δt_m is average transfer temperature difference, and K_c is heat transfer coefficient of cooling tower:

$$K_c = \xi / D \quad (3)$$

In Eq. (3), ξ is the average thermal conductivity of heat transfer pipe, and D is the average thickness of heat transfer pipe wall.

In Eq. (1), the left side denotes the system output. In the right side of the equation, a_1 and a_2 denote system parameters. These parameters have practical physical significance and can express the states of the system, and hence, we consider these parameters as the states. The Kalman filter method obtains the estimated value. Further, we can get the estimated states, and the fault detection is performed by comparison with the dynamic control limit. The details are stated in Sect. 4.

4 Methodology

In this study, the parameters are estimated by the Kalman filter with historical and the current measurement data. We then use the SPC to set the dynamic control limit and observe whether the next parameter estimates are beyond the scope of the set dynamic control limits. If it is exceeded, a fault alarm is issued.

The Kalman filter is a mature technology that is widely used in the world. This method has the characteristics of automatic update iteration without much historical data. The Kalman filter can automatically correct the error in the iterative process, making the estimated value approximately equal to the true value.

The state space expression of the standard Kalman filter is as follows [8]:

$$x(k) = Ax(k-1) + w(k) \quad (4)$$

$$x(k) = x(k-1) + (x(k|k-23) - x(k|k-24)) + w(k-1) \quad (5)$$

where $x(k)$ is the estimated state of the cooling water in ground heat pump and the air flow in air heat pump at time k . A is the parameter of the Kalman filter. The state variables in the model obey a positive distribution, therefore $A = 1$, and $w(k)$ is the

matrix of process noise at time k . The object of study is located in Shenyang, China; extreme weather changes affect the accuracy of fault detection system. In order to reduce the false alarm rate caused by fluctuation of temperature in the morning and evening, this study eliminates a 24-hour cycle of process noise fluctuations.

Additionally, the system output equation is

$$z(k) = H(k)x(k) + v(k) \quad (6)$$

where $z(k)$ is the heat get from the outdoor environment at time k , $H(k)$ is the matrix of the measured parameters at time k , such as cooling water temperature and air flow rate, $v(k)$ is the measured noise at time k , such as error caused by the lack in sensor accuracy.

We get the following equations from the cooling tower model:

$$z(k) = AK_c \Delta t_m \quad (7)$$

$$H(k) = \left(\frac{v_{cw}}{d} (T_{cwo,k} - T_{cwi,k}) C_{cw}, \frac{\lambda}{de} (T_{ao,k} - T_{ai,k}) Re^m \right) \quad (8)$$

The state variable estimate $x(k)$ of cooling tower is expressed by the vector of a_1 and a_2 , and it is represented as follows:

$$x(k) = (a_{1,k}, a_{2,k})^T \quad (9)$$

where a_1 and a_2 are the parameters in the model of the cooling tower, which are described in Sect. 3. The fault detection system can get the status of the cooling tower system in real time by using these parameters.

The state variable estimate $x(k+1|k)$ at time $k+1$ can be obtained from the Kalman filter method. The equation is as follows [8]:

$$x(k+1|k) = x(k) + K(k+1)(z(k+1) - H(k+1)x(k)) \quad (10)$$

where $x(k+1|k)$ is the estimated state at time $k+1$, which is estimated by the state at time k , $K(k+1)$ is the Kalman gain at time $k+1$, and can be calculated as follows:

$$K(k+1) = \frac{P(k|k)H(k+1)^T}{(H(k+1)P(k|k)H(k+1)^T + R(k+1))} \quad (11)$$

where $P(k|k)$ is the covariance matrix of $x(k)$ at time k :

$$P(k|k) = P(k) + Q(k) \quad (12)$$

where $P(k)$ is the covariance matrix of the state variable estimate $x(k)$ at time k , and the $Q(k)$ is the process noise covariance matrix.

After the state at time $k + 1$ is estimated, the $P(k + 1)$ can be calculated as follows:

$$P(k + 1) = (I - K(k + 1)H(k + 1))P(k|k) \quad (13)$$

After the state estimation values for a period of time are obtained, the dynamic control limits of the cooling tower can be obtained by calculating the sliding average [9, 10].

SPC rules are generally used in production. In this study, we use this rule to perform fault detection analysis of the cooling tower. According to mathematics, the distribution of parameters is subject to the normal distribution; therefore, we take control limits three times. We take $logb$ as an example and build the dynamic control limits as follows:

$$[\sigma_{logb,k} - 3\mu_{logb,k}, \sigma_{logb,k} + 3\mu_{logb,k}], \quad k \in N \quad (14)$$

where $\mu_{logb,k}$ is the control limit size of the $logb$ at time k , and $\sigma_{logb,k}$ is the sliding average value of the standard deviation of $\widehat{logb}_i$ over the last K hours. The $\mu_{logb,k}$ can be calculated as follows:

$$\mu_{logb,k} = \frac{\sum_{i=k-K+1}^k \widehat{logb}_i}{k} \quad (15)$$

The dynamic control limits can be set by dynamic values according to the SPC rules. Here K takes the value up to 24 hours.

If a parameter located in the range of control limits, the state is considered as stable.

When the parameter exceeds the set error limit for six consecutive hours, we determine the system fault and issue a fault alarm.

5 Simulation Studies

The Sino-German building in the Shenyang Jianzhu University is considered for testing the proposed method in this study. The Sino-German building was designed by the Shenyang Jianzhu University, China, and Darmstadt University of Technology, Germany. The building is located in Shenyang, China, which is located at high latitudes. The temperature in winter is very low; the average temperature is -15°C , and the difference value in temperature may reach up to 20°C in a day. In January, the outdoor temperature can reach -30°C . In such weather condition, the HVAC system has to bear a heavy workload every day, and therefore it can easily get fail. In this building, there are two cooling tower systems in the HVAC system: one is the ground source heat pump unit, which is located underground, and the other is the air source heat pump unit, which is installed on the southern wall of the Sino-German

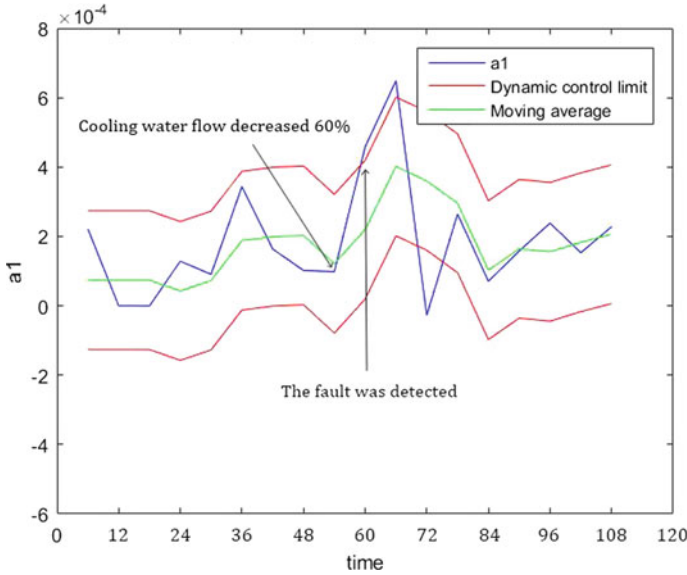


Fig. 2 Parameter diagram of a_1

building. The data of the system can be obtained from the sensors and equipment in the building. Through the simulation of the test object, we can obtain the graphs in Figs. 2 and 3.

The experiment performed sampling every 6 h. In the experiment, 24 sets of initial data were set randomly, and the data were estimated according to the measured data and equipment parameters. In Fig. 2, the blue line is the variation of a_1 , the green line is the moving average of a_1 , and the red lines are the dynamic control limits which are set by the SPC. In the experiment, we simulated to decrease 60% of the cooling water flow at the 54th hour. By analyzing Fig. 2, we can see that the blue line exceeded the red line at the 60th hour, and it did not return to the limit until 72 h. This means that the variation of a_1 exceeds the dynamic control limits for at least 12 h, and the fault is detected by a_1 after 6 h.

The states x_k consist of a_1 and a_2 ; therefore, we can obtain the variation of a_2 at the same time; this is shown in Fig. 3.

The meaning of each curve in Fig. 3 is similar to that in Fig. 2; the blue line is the variation of parameters, the green line is the moving average, and the red lines are the dynamic control limits. We did not simulate a fault in the air heat pump. By analyzing Fig. 3, we can see that the variation of a_2 did not exceed the control limits. Consequently, all the parameters in the states need to be considered in this method.

In Figs. 2 and 3, we can see that there are no parameters that exceed the dynamic control limits when the fault did not occur; therefore, we can conclude that the dynamic control limits can have a good control of the false alarm rate.

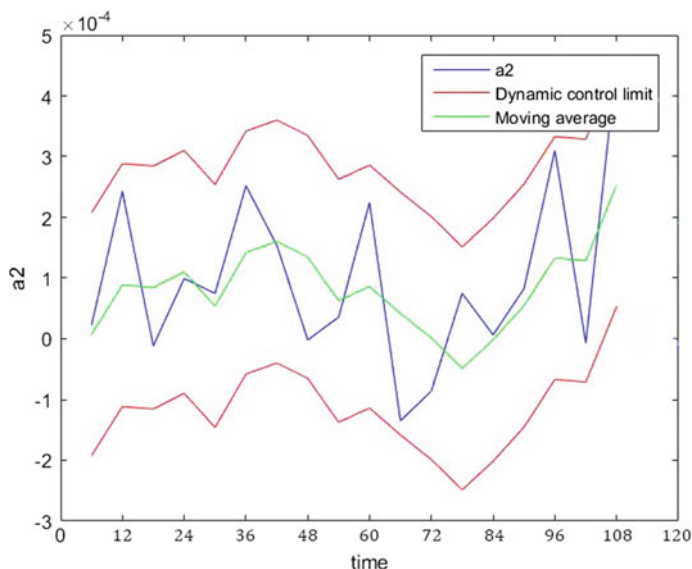


Fig. 3 Parameter diagram of a_2

6 Conclusions

In the context of energy saving, reducing the energy consumption of a cooling tower system caused by a fault counts as a feasible method. The method introduced in this paper uses the Kalman filter and SPC control rules to reduce the influence of environmental factors on the fault detection and simplifies the fault detection procedure. The results of the experiment in the Sino-German building demonstrate that this method can detect the fault. However, in this method, we need to design the fault detection method for every part of the HVAC system. Therefore, we plan to find an effective way to detect the fault for the entire HVAC system in the future work.

Acknowledgements The authors would like to thank the referees for their constructive comments and suggestions. The authors also gratefully acknowledge the contributions of Prof. Tianyou Chai of Northeastern University, Shenyang, China. The research is financially sponsored by the National Natural Science Foundation of China (61873174, 61503259), China Postdoctoral Science Foundation Funded Project (2017M611261), Science and Technology Projects of Ministry of Housing and Urban Rural Development (2018-K1-019), SJZU Postdoctoral Innovation Fund Project (SJZUBSH201705), Chinese Scholarship Council, and Hanyu Plan of Shenyang Jianzhu University.

References

1. Pérez-lombard, L., Ortiz, J., Pout, C.: A review on buildings energy consumption information. *Energy Build.* **40**(3), 394–398 (2008)
2. Li, Z.S., Zhang, G.Q., Liu, J.L.: Application and prospect on fault detection and diagnosis (FDD) technology for HVAC. *Fluid Mach.* **34**(6), 74–81 (2006) (in Chinese)
3. Lee, W.S., Grosh, D.L., Tillman, F.A., Lie, C.H.: Fault tree analysis, methods, and applications: a review. *IEEE Trans. Reliab.* **34**(3), 194–203 (2009)
4. Isermann, R.: Model-based fault-detection and diagnosis-status and applications. *Ann. Rev. Control* **29**(1), 71–85 (2005)
5. Namburu, S.M., Azam, M.S., Luo, J., Choi, K., Pattipati, K.R.: Data-driven modeling, fault diagnosis and optimal sensor selection for HVAC chillers. *IEEE Trans. Autom. Sci. Eng.* **4**(3), 469–473 (2007)
6. Sun, L.L., Wu, J.H., Jia, H.Q., Liu, X.B.: Research on fault detection method for heat pump air conditioning system under cold weather. *Chin. J. Chem. Eng.* **2017**(12), 1812–1819 (2017)
7. Yan, Q.S., Shi, W.X., Tian, C.Q.: Refrigeration technology for air conditioning, 4th edn. China Architecture & Building Press, Beijing (2010) (in Chinese)
8. Sun, B., Luh, P.B., Jia, Q.S., Zheng, O.N., Song, F.T.: Building energy doctors: an SPC and Kalman filter-based method for system-level fault detection in HVAC systems. *IEEE Trans. Autom. Sci. Eng.* **11**(1), 215–229 (2014)
9. Colledani, M., Tolio, T.: Performance evaluation of production systems monitored by statistical process control and off-line inspections. *Int. J. Prod. Econ.* **120**(2), 348–367 (2009)
10. Pfohl, H.C., Cullmann, O., Stölzle, W.: Inventory management with statistical process control: simulation and evaluation. *J. Bus. Logist.* **20**(1), 101–120 (1999)

Application of Information Network and Control Network Integration Technology in Central Air Conditioning Data Management System



Yan Bai, Zhengmin Liu and Qingchang Ren

Abstract With the rapid development of intelligent buildings and the continuous improvement of management requirements, demand for real-time information management and rationality of the storage strategy design in a central air conditioning system are increased simultaneously. Aiming at the characteristics of the central air conditioning data management system with large numbers of detecting points, large capacity of storage, variety in subsystems, and management complexity in intelligent buildings, the data acquisition module was constructed and the data management system was designed and implemented in the research based on OPC remote access server technology and the zero configuration and high efficiency of SQLite database. The experimental and practical results show that the designed system can be used for data acquisition and information management in central air conditioning monitoring systems and improve the level and efficiency of management to a certain degree.

Keywords Central air conditioning system · Data acquisition and management · OPC · SQLite

1 Introduction

With the rapid development of intelligent buildings and the continuous improvement of control and management requirements, the data fusion and integration of heterogeneous information network and control network has become an important direction in enterprise automation system and intelligent buildings, and the inevitable trend of academic development of control theory as well [1].

Y. Bai (✉) · Z. Liu

School of Science, Xi'an University of Architecture and Technology, Xi'an 710055, China

e-mail: yb_xauat@126.com

Q. Ren

School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_22

The information of all subsystems of the distributed network is integrated into intelligent building data acquisition and management system, by which the subsystems, key equipment, and monitoring points are centralized monitored. The related information is recorded by accessing to the database, and the linkage function among subsystems can be realized to coordinate and optimize the operation of each subsystem. So as to provide safe, comfortable and convenient living conditions and an efficient working environment for people by means of economical and intelligent operation and maintenance [2].

At present, the gateway technology, dynamic data exchange technology, unified protocol standard technology, network database technology, and OPC (Object Linking and Embedding for Process Control) technology are commonly used for control and information network integration. F. Lin et al. have analyzed and compared industrial control network and information network convergence technology. Li, S. J. et al. proposed an integrated framework of the control network and information network based on Web database technology and applied it to the network integration of the central air conditioning system. Qiao, F. Q., Zhao, X. et al. introduced a real-time database Proposed and designed a real-time data acquisition system based on OPC technology

The above research has made achievements in the fields of control network and information network integration, the applications of web database and real-time database on the basis of OPC technology. However, the research of customized data acquisition and management system and optimization of storage strategy for the central air conditioning system still need further study.

In this paper, aiming at the characteristics of the central air conditioning data management system with large numbers of detecting points, large capacity of storage, variety in subsystems and management complexity in intelligent buildings, the customized data acquisition and management system and the storage optimization strategy have been designed and deployed on the VAV (Variable Air Volume) central air conditioning system experimental platform in Intelligent Building Institute in Xi'an University of Architecture and Technology (XAUAT).

2 Overall Architecture

This experimental research platform (divided into water systems, wind system, and terminal system) with the total area of 135 m², consists of 2 chillers, 2 Air-Handling Units (AHU), 6 VAV boxes, 3 fixed-frequency pumps, 1 variable-speed pump, 1 hot water cylinder, several air diffusers, pipes, and various kinds of sensors, controllers, actuators, Industrial Personal Computers (IPS), etc. Figure 1 shows the Architecture of VAV Central Air Conditioning System in XAUAT.

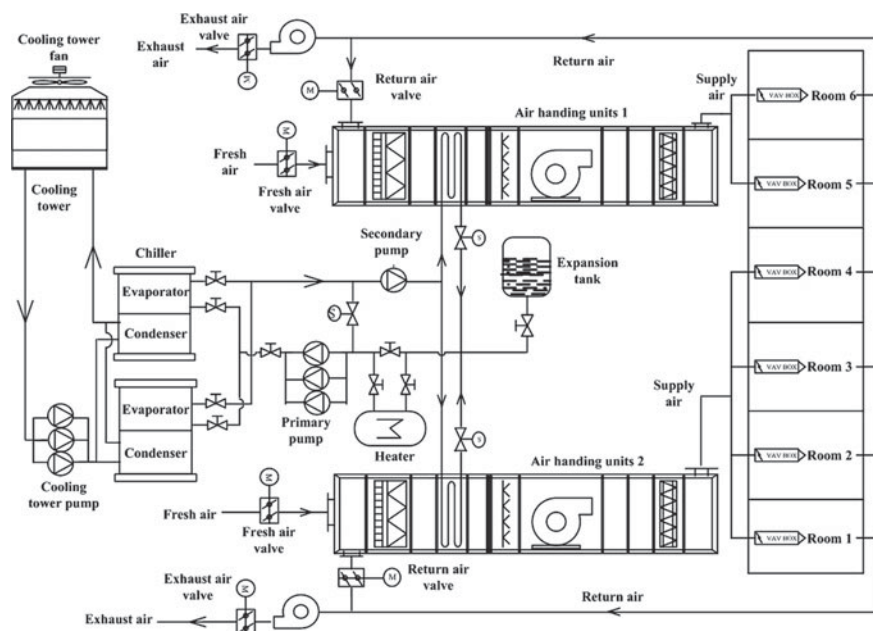


Fig. 1 Architecture of VAV central air conditioning system in XAUAT

3 System Overall Design

The system is responsible for the acquisition and online processing of control information and state information in the air conditioning system. And it must keep up with the data rate from the continuously operating devices.

The users of this data acquisition and management system mainly include ordinary users in terminal rooms, system administrators, system operators, etc. Users of different roles have different operations. Ordinary users are concerned about the trend of energy consumption and state parameters of their terminal systems. The system operators focus on the relevant control parameters and state parameters of each subsystem in the control process. System administrators mainly focus on historical data and abnormal data of the system, in addition to user authority assignment and management.

Therefore, the system mainly includes the functions of data acquisition, data storage, data query, data display, user authority management, etc., based on the creation of engineering and the level division of control point. In addition, in view of the characteristics of real-time monitoring and large data volume, the storage strategies should be optimized to improve the management efficiency of the system [3].

3.1 Performance-Based Design

In this research, the data acquisition module was constructed and the data management system was designed and implemented in the research based on OPC remote access server technology and the zero configuration and high efficiency of SQLite database [4]. In addition, aiming at the characteristics of the central air conditioning data management system with large numbers of detecting points and large capacity of storage, the optimized storage strategy was designed for data redundancy reducing and system management efficiency improvement.

An efficient data acquisition and management system was constructed for different users' authorities. And different types of system parameters and energy consumption data can be customized for collection, analysis, and processing to meet the needs of different users. The main functions include data acquisition, data storage, data display, data query, data export, and authority management. The communication between the upper computer software and Omron PLC is established through NI OPC Server and DSC module. The overall structure of the software shows in Fig. 2.

3.2 Architecture Design

The designed system is an OPC client program, which provides data query, display, and export functions to users for different requirements on the basis of OPC DA 3.0 specification. It was divided into the basic environmental layer, data resource layer, interaction support layer, application layer, and service layer. Figure 3 shows the System Architecture of software.

The basic environment layer was responsible for the communication between the OPC client and the OPC server which include the communication with the local OPC server and with the OPC server on the remote host [5–7]. The data resource layer includes the SQLite database and system configuration files. Among them, the SQLite database was responsible for recording the collected data, and the system configuration files were responsible for recording the configuration information of the project [8, 9]. The interaction support layer was responsible for the interaction with the user. The data application layer was responsible for responding to the data

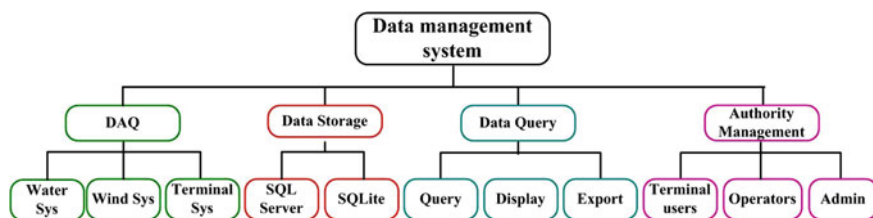


Fig. 2 The structure chart of the software

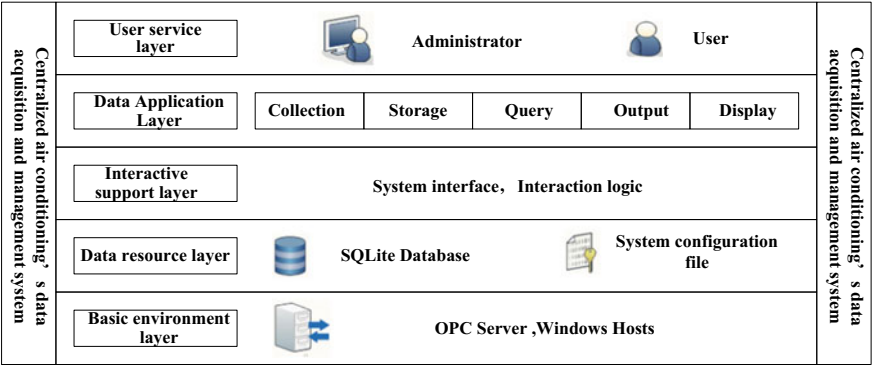


Fig. 3 System architecture of software

processing requests of the users and feedback to the users. The user service layer was responsible for responding to different user's request for creating project, configuring, and login.

3.3 Modules Design

According to the overall architectural design of the system, the following modules will be designed to support its implementation:

- (1) XML: Responsible for the analysis and generation of system configuration files;
- (2) DBMS: Responsible for the database system and data management or query operations;
- (3) GUI: Responsible for the system's interactive interface and interaction logic;
- (4) OPCDA: Responsible for communication and data exchange between system and OPC server;
- (5) XLS: Responsible for exporting the database information to the XLS/XLSX files;
- (6) CHART: Responsible for the display of data trend diagram and data storage;
- (7) BASIC: Responsible for the maintenance of functions or foundation classes in the system.

The relationship of module invoke was shown in Fig. 4.

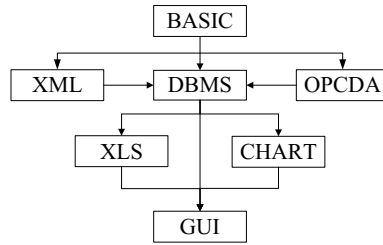


Fig. 4 Relationship of module invoke

4 Software Implementation

4.1 Development Environment and Tools

- (1) Programming Language: C++;
- (2) IDE (Integrated Development Environment) Tools: VS2015;
- (3) Development Platform: Windows 10;
- (4) Target platform: Windows operating system above Windows XP;
- (5) GUI Framework: wxWidgets;
- (6) Dependent libraries: WT Client, TinyXml2, Chart Director, Libxl;
- (7) Database: SQLite database;
- (8) Packaging Tool: Inno Setup.

4.2 Storage Optimization Strategy

Aiming at the characteristics of real-time monitoring and large capacity of storage in the central air conditioning system, the following optimized storage strategy was designed to improve the management efficiency of the system.

- (1) When data was stored, for each data point in the same system, a change limit can be specified. When the data value of a node in the system exceeds the limit of the previously stored value, the record was triggered.
- (2) The authority of interval setting for data recording could be assigned to the user. The system records data at a given time–frequency, when the user specifies a system data refresh rate.
- (3) Database detection and data transfer. When the database reaches the specified capacity, the user could transfer the current database to another location or other storage. And then, a database with the same structure as the transferred database would be regenerated and the future data would be stored in the new one.

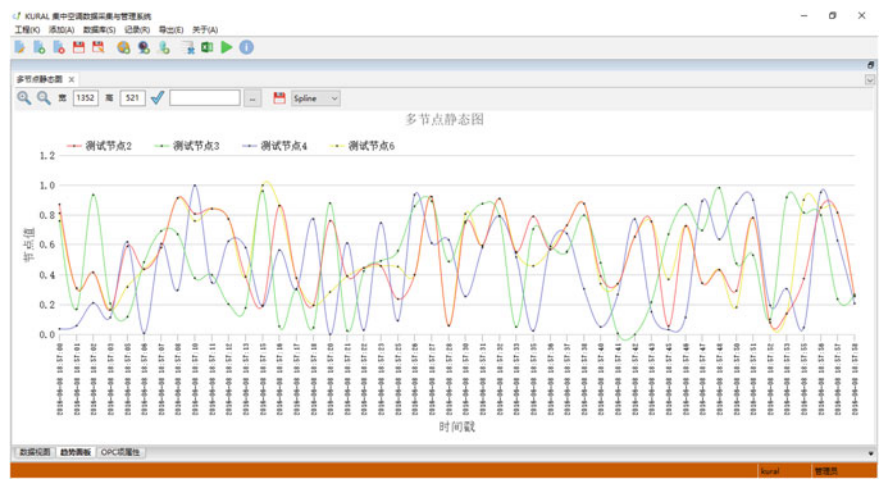


Fig. 5 Panel display page

Combining the above three strategies, under the premise of ensuring that important changes to the data are not overlooked, data memory usage can be reduced by 20–60% compared to previous data storage methods.

4.3 The Implementation and User Interface

The data view page was responsible for displaying the data information of the system. It also provided query functions and data view export functions.

As shown in Fig. 5, the display panel is responsible for the display of trend graphs, including static diagrams and dynamic diagrams, and provides the export function of trend diagrams as well.

The system provided three kinds of drawing methods including Line, Spline, and Step Line. The function of zoom in and zoom out was supported by this system. And the maximum supported pixel was 32,000 × 16,000, the smallest pixel was determined by the size of the panel.

5 Conclusions

In this research, the data acquisition and management system for a central air conditioning system were designed and implemented based on OPC DA3.0 specification. Functions of data acquisition, data storage, data query, data export, etc., were designed and implemented to satisfy the needs of different users. Aiming at the char-

acteristics of the central air conditioning, data management system with the real time and large capacity of storage, the optimized storage strategy was designed.

The experimental and practical results show that the designed system can be used for data acquisition and information management in central air conditioning monitoring systems and improve the level and efficiency of management to a certain degree. At the same time, it provides case support for the design and development of a customized OPC client program.

Acknowledgements This work was supported by the Shaanxi Provincial Natural Science Foundation for Young Scientist (NO. 2017JQ5075), the Ministry of Housing and Urban-Rural Development of China (NO. 2016-K1-013), the 13th Five-Year Plan Project of Education Science in Shanxi Province (SGH18H111) the Basic Research Foundation (No. JC1516), and the Youth Scientist Foundation of Xi'an University of Architecture & Technology (No. RC1515).

References

1. Perumal, T., Ramli, A.R., Leong, C.Y., Samsudin, K., Mansor, S.: Middleware for heterogeneous subsystems interoperability in intelligent buildings. *Autom. Constr.* **19**, 160–168 (2010)
2. Yang, T., Clements-Croome, D., Marson, M.: Building energy management systems. In: *Encyclopedia of Sustainable Technologies*, pp. 291–309 (2017)
3. Hoffmann, M., Büscher, C., Meisen, T., Jeschke, S.: Continuous integration of field level production data into top-level information systems using the OPC interface standard. *Procedia CIRP* **41**, 496–501 (2016)
4. Mehta, B.R., Reddy, Y.J.: *Industrial Process Automation Systems: Design and Implementation*, pp. 459–477 (2015)
5. Liu, T.D., Cai, G.G., Peng, X.F.: OPC server software design in DSC. In: *Proceeding of 2009 4th International Conference on Computer Science & Education*, vol. 4, pp. 456–458 (2009)
6. Cavalieri, S., Stefano, D.D., Salafia, M.G., Scroppo, M.S.: Integrating OPC UA with web technologies to enhance interoperability. In: *Computer Standards and Interfaces*, pp. 1206–1211 (2017)
7. Yin, Y., Zhou, B.: The analysis and research of OPC XML-DA server. *Energy Procedia* **16**(2), 1535–1540 (2010)
8. Nemetz, S., Schmitt, S., Freiling, F.: A standardized corpus for SQLite database forensics. *Digit. Investig.* **24**, 121–130 (2018)
9. Wang, W., Zhang, X., Li, Y., Li, Y.: Open CNC machine tool's state data acquisition and application based on OPC specification. *Procedia CIRP* **56**, 384–388 (2016)

MFAC and Parameter Optimization for a Class of Models in HVAC



Zengxi Feng, Junqi Yu, Zhongtian Rao and Anjun Zhao

Abstract That the aging equipment, nonlinearity, and other external factors combine to make one-order plus time-delay process model be uncertain in Heating, Ventilation, and Air Conditioning (HVAC) makes the initial control parameters ineffective. Aiming at one-order plus time-delay process model with uncertainty in HVAC, the Model-Free Adaptive Control (MFAC) is introduced in studying the control characters. The changes on the overshoot and settling time are simulated and compared with PID, which shows that MFAC has good stability and anti-interference and is insensitive to the change of time-delay. Therefore, it is proved that MFAC is suitable for solving the problem of control failure caused by changing time-delay. However, there are no good methods for the parameters setting of MFAC, which makes it difficult to find the optimal controller parameters. Aiming at the condition, the simplex method is used to optimize the controller parameters and the first-order inertia plus time-delay model is regarded as the controlled object in this paper. The simulation results show that the MFAC parameters optimized by the simplex method have good control results.

Keywords Model-free adaptive control · Uncertainty · Anti-interference · Simplex · Optimization

Z. Feng · J. Yu (✉) · Z. Rao · A. Zhao

School of Building Services Science & Engineering, Xi'an University of Architecture and Technology, Xi'an, Shaanxi, China

e-mail: junqiyu@126.com

Z. Feng · J. Yu

State Key Laboratory of Green Building in Western China, Xi'an University of Architecture and Technology, Xi'an, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_23

1 Introduction

One-order plus time-delay process model is the common mathematical model in Heating, Ventilation & Air Conditioning (HVAC) control systems. For example, the pressure difference control model, the temperature difference control model, the air supply temperature control model and fan static pressure control model in a central air conditioner are one-order plus time-delay process models. If the mathematical model of one-order plus time-delay is accurate, there are some good control methods dealing with the time-delay. For example, when the time-delay is small and accurate, PID and derived PID algorithms can get good control results. When the time-delay is longer and accurate, Dalin algorithm and Smith predictive control algorithm are very suitable. However, in practice, there are many uncertainties of mathematical models in many systems and the model is difficult to establish, which results in the partial failure of control algorithm. For example, Smith predictive control relies heavily on the precise mathematical model. When the estimated model and the actual object do not match, the control quality will deteriorate significantly or even diverge, and it is also very sensitive to external disturbances [1].

Aiming at the uncertainty, some results had been reported and some general conclusions had been got [2]. However, because the mathematical models of the controlled systems are diverse and different models have different parameters and structures, the specific effects on the performance of the Model-Free Adaptive Control (MFAC) are not still very clear. For example, there are no results on MFAC dealing with the uncertainty of the one-order plus time-delay process model caused by the aging equipment, nonlinearity, and other external factors in Heating Ventilation Air Conditioning (HVAC). In order to explore the adaptability of the MFAC dealing with the uncertainty of the one-order plus time-delay process model, the control performance is studied by the simulation when the model parameters change, such as the time-delay, process gain, and system time constant. This study result is helpful to refine the application scope of MFAC and choose control methods for one-order plus time-delay delay systems with uncertainty.

Based on the above study, how to optimize the MFAC controller parameters is studied. Although MFAC is an effective, widely applicable and advanced control strategy that does not rely on the mathematical model of objects, its controller parameters are not easy to optimize, which sometimes affects the control performance of MFAC. Therefore, how to set and optimize the parameters of the MFAC controller is an important issue. Hou refers to the parameters tuning method of PID controller and proposes four methods of tuning the controller parameters of MFAC: parameters tuning method based on the Z-N empirical equation, Cohen-Coon parameters tuning method, parameters tuning method based on Z-N Critical Proportional band and parameters tuning method based on system performance index [3]. These methods have played a positive role in the promotion of MFAC. However, the choice of the initial value of MFAC controller has a great influence on its control performance.

If the initial value changes, the control effect will also change. It leads to the difficulty during tuning the controller parameters of MFAC. Therefore, in this research work, it attempts to use the simplex optimization algorithm to automatically optimize parameters for different initial values of the MFAC controller parameters.

2 The Model-Free Adaptive Control

MFAC uses a novel concept pseudo partial derivative (PPD) (or pseudo-gradient or pseudo-Jacobi matrix) and an equivalent dynamic linearization data model at current operation point. The optimal controller is designed in the light of the estimation of the PPD online, which only makes use of the output and input data of the controlled plant [4–7].

The MFAC has the following advantages. (1) It only employs the input and output data of the controlled system. This shows that we are able to design the controller independently and we are able to build a universal controller. (2) Any training process and any external testing signals are not needed in the MFAC mechanism. This shows that the MFAC is low cost and less expensive. (3) It is simple and can be easily employed and carried out, and it has strong robustness and the minimum computational burden [4–7].

Following discrete-time SISO nonlinear systems is considered:

$$y(k+1) = f(y(k), y(k-1), \dots, y(k-n_y), u(k), u(k-1), \dots, u(k-n_u)) \quad (1)$$

where n_y , n_u are the orders of output $y(k)$ and input $u(k)$, respectively, $f(\cdot \cdot \cdot)$ is a nonlinear function.

Rewrite Eq. (1) as:

$$y(k+1) - y(k) = \phi(k)[u(k) - u(k-1)] \quad (2)$$

where $\phi(k)$ is the PPD of controlled plant [4].

Aiming at the controlled system as Eq. (2), based on the MFAC, the control law can be given as follows [4]

$$u(k) = u(k-1) + \frac{\rho}{\lambda + |\hat{\phi}(k)|^2} \hat{\phi}(k) [y^* - y(k)] \quad (3)$$

where ρ is the control parameter, y^* is the desired output value of the system, λ is a proper constant, $\hat{\phi}(k)$ is the following PPD estimation of $\phi(k)$. $\hat{\phi}(k)$ can be calculated by Eq. (4).

$$\hat{\phi}(k) = \hat{\phi}(k-1) + \frac{\rho}{\mu + \Delta u(k-1)^2} \Delta u(k-1) \left[\Delta y(k) - \hat{\phi}(k-1) \Delta u(k-1) \right] \quad (4)$$

where $\hat{\phi}(k) = \hat{\phi}(1)$, if $|\hat{\phi}(k)| \leq \varepsilon$, or $|\Delta u(k-1)| \leq \varepsilon$. $\Delta u(k-1) = u(k-1) - u(k-2)$, η is the parameter which is added to increase flexibility of the algorithm, μ is a proper constant, $\Delta y(k) = y(k) - y(k-1)$, ε is a sufficiently small positive number.

3 Simulation Studies on Controlling Object with Uncertainty

3.1 Simulation Object and Parameter Settings

The static pressure model of central air conditioning is shown in Eq. (5).

$$G(s) = \frac{y(s)}{u(s)} = \frac{K}{Ts + 1} e^{-\tau s} \quad (5)$$

where u is the input frequency of fan, y is the duct static pressure, K is the process gain, T is the system time constant, τ is the time-delay.

Due to the uncertainty resulting from the aging equipment, nonlinearity and other external factors, the parameters of Eq. (5) maybe change. The static pressure model built by identification during a certain period is shown in Eq. (6) [8]. Based on the results of reference [8], the final scope of changing parameters, K , T and τ are limited to be [3, 6], [4, 8], and [2, 10], respectively.

$$G(s) = \frac{4.20}{6.21s + 1} e^{-4.87s} \quad (6)$$

In order to compare the control effects of the traditional PID and MFAC in the case of different changes of different parameters, the change of K , T , and τ is divided into three categories. The specific changes in the model parameters are shown in the first 3 columns of Table 1. The initial values of the controller parameters of PID and MFAC are assigned as the following. $k_P = 0.1842$, $k_I = 0.023$, $k_D = 0.2218$. $\mu = 1.55$, $\lambda = 1.65$, $\eta = 0.009$, $\rho = 0.14$.

Simulation studies were carried out with the PID and MFAC control strategies, respectively, in the case of parameter changes. The simulation results are shown in Table 1. Due to the limitation of paper length, four kinds of simulation results are shown here. The simulation results with the original parameters of the model are shown in Fig. 1. The simulations with a single parameter change of process gain K , time constant T and time-delay τ , respectively, are shown in Figs. 2, 3 and 4.

Table 1 The model parameters and simulation results

Model parameters	Parameters change	PID control		MFAC	
		Settling time/s	Overshoot/%	Settling time/s	Overshoot/%
$K = 4.20, T = 6.21, \tau = 4.87$	Primitive parameters	37	0	163	0
$K = 6, T = 6.21, \tau = 4.87$	One parameter changes	48	8.25	151	2.51
$K = 3, T = 6.21, \tau = 4.87$		62	0	258	0%
$K = 4.20, T = 8, \tau = 4.87$		54	3.30	205	0.99
$K = 4.20, T = 4, \tau = 4.87$		54	0	178	0
$K = 4.20, T = 6.21, \tau = 10$		105	39.96	206	2.04
$K = 4.20, T = 6.21, \tau = 2$		55	0	183	0

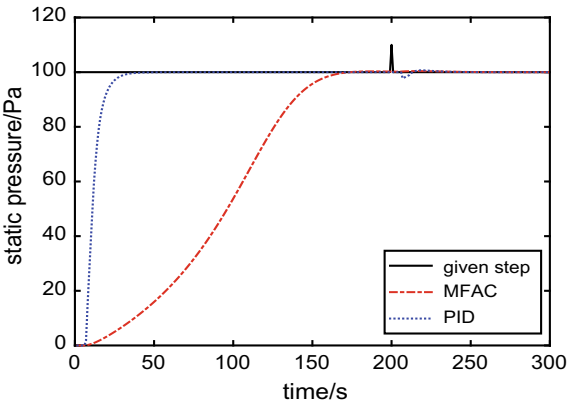


Fig. 1 The simulation based on primitive parameters

Fig. 2 The simulation based on changed parameter K .

1—PID: $K = 6$, $T = 6.21$, $\tau = 4.87$; 2—PID: $K = 3$, $T = 6.21$, $\tau = 4.87$; 3—MFAC: $K = 6$, $T = 6.21$, $\tau = 4.87$; 4—MFAC: $K = 3$, $T = 6.21$, $\tau = 4.87$

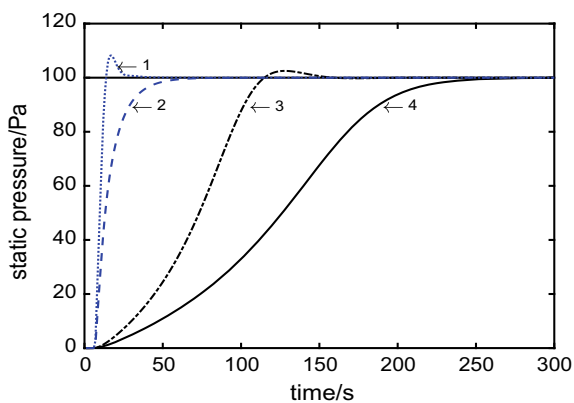


Fig. 3 The simulation based on changed parameter T .

1—PID: $K = 4.2$, $T = 8$, $\tau = 4.87$; 2—PID: $K = 4.2$, $T = 4$, $\tau = 4.87$; 3—MFAC: $K = 4.2$, $T = 8$, $\tau = 4.87$; 4—MFAC: $K = 4.2$, $T = 4$, $\tau = 4.87$

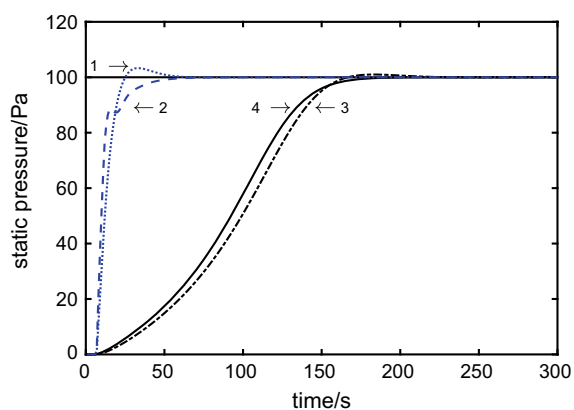
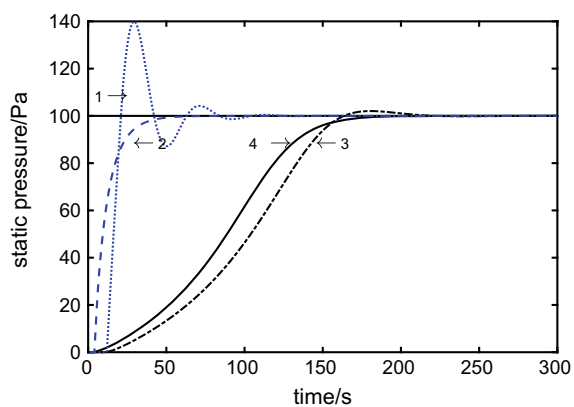


Fig. 4 The simulation based on changed parameter τ .

1—PID: $K = 4.2$, $T = 6.21$, $\tau = 10$; 2—PID: $K = 4.2$, $T = 6.21$, $\tau = 2$; 3—MFAC: $K = 4.2$, $T = 6.21$, $\tau = 10$; 4—MFAC: $K = 4$, $T = 6.21$, $\tau = 2$



3.2 Analysis of Simulation Results

For the original model, the results of control simulation in this study show that the overshoots of both methods are 0%. The settling time is 37 s by PID control, and it is 163 s by MFAC control. For anti-interference capability, when a perturbation amplitude of 10 is added in input at $t = 200$ s, the output fluctuates by PID control, its size is -2.26 and the settling time of this disturbance is 17 s and the output by MFAC control has a smaller fluctuation with a size of 0.39, and the settling time for this disturbance is 15 s. This indicates that when MFAC is adopted, the system has good anti-interference performance, good stability, and short recovery time.

Based on the above analysis, the following conclusions can be obvious. (1) The anti-interference performance of MFAC is better than that of PID. That is to say, if the system is disturbed, the system controlled by MFAC method has good immunity, stability and short recovery time. (2) From the effect of settling time and overshoot, For MFAC, the change of process gain K has a great influence on the settling time. The change of inertia time constant T and time-delay τ has little effect on the settling time, and the change of three parameters has little effect on the overshoot. For PID control, the change of process gain K and time-delay has a greater impact on overshoot and settling time, and the inertia time constant T has a great effect on overshoot and has little effect on the settling time.

It is obvious that the one-order plus time-delay process model with uncertainties in HVAC systems, especially with high anti-interference requirements, large delays and uncertain delays, the control performance of MFAC is superior to that of PID.

4 MFAC Parameter Optimization

So far, there are no very simple and useful methods for tuning and optimizing the controller parameters, λ , ρ , η , and μ , of MFAC. In addition, the initial values of controller parameters dramatically affect the control results. Therefore, it is hard to obtain the optimal controller parameters. In order to obtain the optimal controller parameters, the optimization theory is introduced to this paper. For multivariate optimization, there are some methods, conjugate gradient method, steepest descent method, random optimization method, and simplex method, to be chosen. In general, the random optimization method is suitable for optimizing many parameters. Conjugate gradient method and steepest descent method will take long time to optimizing parameters because of calculating the gradient of objective function in each step. The simplex method uses the comparison of the objective function values of different parameter points to determine the optimal direction, and does not require gradient information. It is more suitable for occasions where there are not many variables to be optimized [9]. Therefore, the basic simplex method is chosen to auto-optimize the controller parameters of MFAC for any given initial values.

4.1 The Basic Idea of the Simplex Method

Simplex is an optimization technique that does not require complex mathematical and statistical tools. It does not need to calculate the gradient of the objective function. There are four basic steps: expansion, contraction, compression, and discarding, during the optimizing process. The basic idea of simplex is described as follows. Take $(n + 1)$ points in the n -dimensional space to form the initial simplex, compare the values of the objective function at the $(n + 1)$ point, discard the worst point that is the point with the largest function value, and replace it with the new point to form a new simplex. By iterating, the value of the function at its vertex gradually decreases, and the vertex gradually approaches the minimum point of the objective function. If it is required to solve the maximum or minimum point of function, the values of function at several points can be obtained and compared, and the changing trend of the function can be determined based on their size relationship, which can decide the reference direction of the search. The minimum or maximum value will be found according to the reference direction of the search [10].

4.2 Simplex Parameter Optimization Design Based on Simulink

Selection of Objective Function. The control performance must be measured and determined according to the performance index, and the performance index usually adopts the objective function. In practice, the function of the deviation between the reference input and the actual system response is used as the objective function. It also called the error objective function. The objective function mainly includes the integration of the error square, the integral of the time and the square of the error, the integral of the absolute value of the error, the integration of the time and the absolute error. The integral function of time and squared error is used as the objective function in this paper. It is shown as Eq. (7).

$$J_{ITSE} = \int_0^{\infty} t e^2(t) dt \quad (7)$$

Programming

Main Program Design. The main program of the MFAC parameter optimization design based on the simplex method is shown as follows:

```
global rou;
global lambda;
global eta;
global miu;
```

```

global i;
i=1;
m0=[0.005 1 0.001 1]
[m,n,o,p]=fminsearch('optm',m0)

```

In the program, the four controller parameters, λ , ρ , η , and μ denoted as lambda, rou, eta, and miu, respectively, of MFAC are defined as global variables. i is a loop variable. The vector m0 is the initial value of λ , ρ , η and μ . The fminsearch function is used in the main program. It is a function for solving unconstrained nonlinear programming, and it uses a simplex algorithm. It works as the following:

$$[x, fval, exitflag, output] = \text{fminsearch}(\text{fun}, x0, \text{options})$$

In the function, fun is the objective function, x0 is the starting point for optimizing search. Options are option parameters for setting optimization. x is the return value of the optimization parameter. fval is the return function value at the optimal point x. exitflag is the termination flag of the return algorithm. Output is a return data structure of optimization algorithm information.

The Optimization Module Based on Simulink. The optimization program based on Simulink is shown in Fig. 5.

The objective function is an integral function of the time and error squared product constructed. The fcn2 function module is error squared and the clock module is the time. Subsystem is the package module of MFAC [11, 12].

4.3 Simulation on Optimizing the Controller Parameters of MFAC

Considering that one characteristic of model-free adaptive control is insensitivity to time-delay, the mathematical model of the simulation object still uses the previous fan static pressure model. It is shown as Eq. (6). This model is a one-order plus

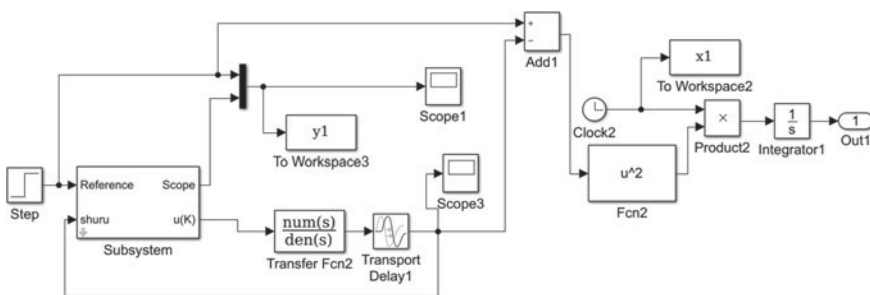


Fig. 5 The optimizing parameter based on Simulink

Table 2 The parameters of MFAC and control result

	Initial value		Optimal value	
ρ	0.15	0.5	0.2236	0.5935
λ	2.2	4	1.3956	3.7836
η	0.11	0.1	0.1183	0.1069
μ	2	4	1.3106	3.6486
Overshoot/%	0	0	5.02	4.79
Settling time/s	500	120	70	40

time-delay process model. Because of $f = \tau/T = 4.87/6.21 = 0.784 > 0.5$, the model has a large delay.

In order to verify the parameter optimization effect of MFAC based on the simplex method, 2 groups of initial parameters were given at random and the optimizations are realized. The initial values and optimal values of the parameters and the control effects are shown in Table 2. The settling time is calculated based on the relative error that is 5%.

The simulation results with 2 groups of initial parameters and 2 groups of optimized parameters are shown in Figs. 6 and 7.

The simulation of parameters optimization shows that the optimization results with the simplex method are affected by the initial values of the parameters and that the optimization results maybe change when the excitation signals change. In general, based on the simulation of one-order plus time-delay process model with the simplex method, the following conclusion can be concluded. (1) As long as the reasonable initial controller parameters of MFAC are given, the optimized parameters can be found based on the simplex method. (2) The constraint range of parameter values, $\rho, \eta \in (0, 2)$, is given in Ref. [2], which ensures that any system that satisfies the

Fig. 6 The simulation result with the first set of parameters

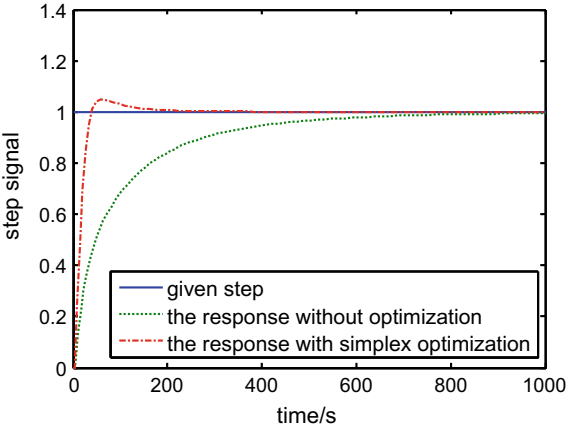
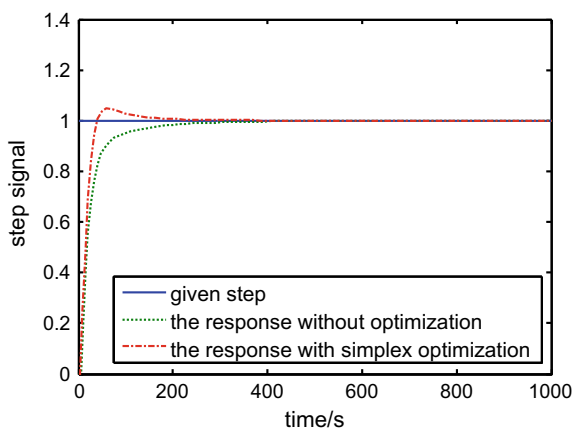


Fig. 7 The simulation result with the second set of parameters



constraints can achieve bounded parameter estimation error and the convergence of output error. The value of ρ and η can change according to the controlled system.

5 Conclusions

The performance of MFAC is studied based on the one-order inertial plus time-delay model with uncertainties in an HVAC system. The results show that MFAC has good stability and anti-interference performance and that MFAC is insensitive to delay variation. At the same time, aiming at the blindness of setting controller parameters of MFAC, the simplex method is used to optimize MFAC controller parameters. The simulation results show that the initial values of MFAC controller parameters are different and the optimized values are different, but the control results with the same optimized parameters are good.

Acknowledgements This study was supported by the National Key Research and Development Program Project (2017YFC0704100), the Special Research Project of Shaanxi Science and Technology Department (2017JM6106), and the Xi'an University of Architecture and Technology Foundation Fund Project (JC1706).

References

1. Sanz, R., García, P., Albertos, P.: A generalized smith predictor for unstable time-delay SISO systems. *ISA Trans.* **72**(1), 197–204 (2018)
2. Hou, Z.S., Jin, S.T.: A novel data-driven control approach for a class of discrete-time nonlinear systems. *IEEE Trans. Control. Syst. Technol.* **19**(6), 1549–1558 (2011)
3. Hou, Z.S., Zhu, Y.M.: Model based control and MFAC, which is better in simulation? *IFAC Proc.* Vol. **46**(13), 82–87 (2013)

4. Hou, Z.S., Xu, J.X.: On data-driven control theory: the state of the art and perspective. *Acta Autom. Sin.* **35**(6), 650–667 (2009)
5. Han, Z.G.: An integrated approach to modeling and adaptive control. *Acta Autom. Sin.* **30**(3), 380–389 (2006)
6. Hou, Z.S., Gao, H.J.: Data-driven control and learning systems. *IEEE Trans. Industr. Electron.* **64**(5), 4070–4075 (2017)
7. Hou, Z.S., Chi, R.H.: An overview of dynamic-linearization-based data-driven control and applications. *IEEE Trans. Ind. Electron.* **64**(5), 4076–4090 (2017)
8. Yang, S.Z., Ren, Q.C.: Robust control based on fan model uncertainty. *Meas. & Control. Technol.* **32**(10), 69–74 (2013). (in Chinese)
9. Kudinov, Y.I., Kolesnikov, V.A., Pashchenko, F.F.: Optimization of fuzzy PID controller's parameters. *Procedia Comput. Sci.* **103**, 618–622 (2017)
10. Aekarin, S., Wudhichai, A.: Particle swam optimization based optimal PID parameters for air heater temperature control system. *Procedia Comput. Sci.* **86**, 108–111 (2016)
11. Li, C.Q., Liu, G.S.: Development & simulation of MFA control module based on Matlab-Simulink. *Control. Eng. China* **15**(9), 61–63 (2008). (in Chinese)
12. Liu, X.Q., Wang, Y., Mu, S.Y.: Design of optimal parameters for PID controller Based on Simplex Search. *Comput. Simul.* **21**(11), 191–194 (2005). (in Chinese)

A Controller Algorithm (ILC) for the Variable Differential Pressure Control of Freezing Water in a Central Air Conditioning System



Qingchang Ren and Hongmei Jiang

Abstract In actual operation, due to the change of many factors, the central air conditioning system runs at non-designed working conditions in most of the time. Usually, it works under partial load, cannot meet the maximum load, causes great waste of energy. This paper proposes an Iterative Learning Controller algorithm (ILC) for the air conditioning water system, deal with the variable frequency control for the secondary pump. Optimization settings of water pressure differential value are given according to customer demand and based on the water valve features, thus make water valve in the chilled system having the largest opening as far as possible to provide the required minimum water differential pressure. By this way not only good control effects can be obtained, but also the energy consumption of pump delivery can be reduced.

Keywords Air conditioning system · Iterative learning controller algorithm (ILC) · Variable differential pressure · Energy saving

1 Introduction

In recent years, with the rapid development of the economy in China, the application of the air conditioning system in the buildings is increasing. Air conditioning energy consumption becomes more and more. This made the demand for energy increased. The energy supply appears serious. Reducing the energy consumption of an air

Q. Ren (✉)

School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China
e-mail: renqch@163.com

H. Jiang

College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou, China

Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_24

261

conditioning system is necessary, therefore, the research on energy consumption of air conditioning water system is urgent.

In the actual operation, due to the change of many factors, most of the air conditioning systems are not in the designed running condition. They are generally working under the partial load, cannot meet the requirement of the maximum load, and cause a big waste in energy. The reasons are the indoor personnel flow change, the change of outdoor temperature and the indoor heat source and so on. Many factors will make the load of the air conditioning room deviate from the designed condition of the load. Therefore, the design of the air conditioning system should be adjusted along with the load condition, real-time tracking of the load change. The tracking way does not just adjust the air output, that is, the so-called variable air volume (VAV) air conditioning system. The chilled water also can be realized as the variable flow by the control strategy. This paper proposes an iterative learning controller (ILC) on secondary pump frequency control, according to the customer's demand for the differential pressure value of optimization settings. This makes the position of chilled water valve close to fully open state, and provide the required minimum water differential pressure for the system. This method can realize the energy saving of the water pump [1, 2].

2 Variable Water Volume Air Conditioning Systems

Variable water volume (VWV) system refers to the system in which supply and return water temperature difference remains unchanged. If the air conditioning loads change, the water flow can be changed to adapt to the change of the terminal load of the air conditioning system. When the terminal load is reduced, water flow in the water system reduced. The cold quantity is reduced to feed the actual load, therefore, can meet the requirements for the reduction of load. Because the decrease in water flow can reduce the transportation energy consumption of the water system, thus it has significant effects on energy saving.

The main steps of the variable flow technique are based on the actual load changes to adjust the speed of the frozen water pump, in order to realize the chilled water flow control. Usually, the differential pressure is chosen for reflecting the system load changes that means to adjust the speed of the water pump according to the change of the differential pressure. If the selection of the differential pressure value is too large, it will not be able to fully excavate the pump energy-saving potential under the partial load. If the selection of differential pressure value is too small, it may not be able to meet each user's requirements under some conditions. Therefore, how to make balance on the load demand and the pump energy-saving effects is the key problem of the application of variable flow technology. Obviously, the optimal settings of the differential pressure value are very important.

3 Variable Differential Pressure Control

The principle of variable differential pressure control is as follows: when the temperature in the room decreases, the temperature controller turns down the chilled water valve. So it reduces the amount of cold so as to fit the change of the room temperature. The changes of valve opening can be detected by the detector in the valve control system, and then compare measured temperature with the setting value. By means of the frequency conversion controller, the pump speed is adjusted to reduce the indoor cold quantity to make the room temperature rise. When the room temperature controller detects the rise of the room temperature, it will turn the valve up. Thus, it not only meets the air conditioning user's need of load variation, remains the biggest valve opening, lowers the valve resistance loss, and saves the energy consumption of pump delivery.

The schematic diagram of the variable differential pressure control system is shown in Fig. 1. The idea of the variable differential pressure control system is implemented by a certain control algorithm, changing the differential pressure setting value. Under the premise of regulating and stable performance, the guarantee system tries to make the selection of regulating valve in fully open position, so as to minimize the valve throttling losses, and the biggest energy-saving effects can be obtained. Because the change of valve opening reflects the size of the load change to a certain extent, it can be used as adjusting parameters [3, 4].

Considering the differential pressure setting value and valve opening, the setting value can be considered as a variable. In an actual control system, if the adjustment of the valve is in high frequency, the adjustment of the differential pressure value should not set too frequently. That is to guarantee the stability of the water pump

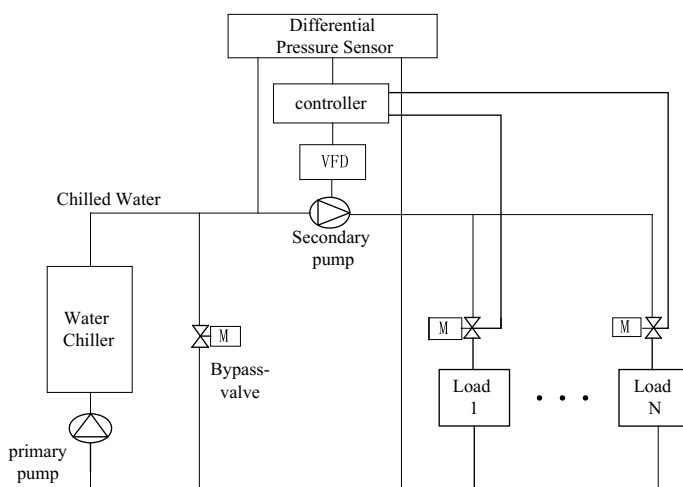


Fig. 1 Schematic diagram of water differential pressure-chilled air-conditioning system

control. As shown in Fig. 1, in the control mode, the specific practices are as follows: (1) Any time all the valve opening is less than 70%, and this state maintains for 30 min, the differential pressure setting value will be decreased by 10%. (2) At least one of the terminal valve position between 70 and 90%, the differential pressure setting value remains the same. (3) At any time the valve opening degree is more than 90%, and this state maintains for 30 min, the differential pressure setting value will be increased by 10%.

Because the selection of the valve in the different branches is different, so the definition of valve fully opening is different. Some of the valve reaching more than 75% is considered as the fully opening valve. So the variable pressure difference control needs not only to detect the valve opening of each branch, but judges the logic value of the valve opening also, takes control algorithm to control the change of pressure difference value according to the requirements of reliability.

4 Variable Differential Pressure Control Based on Iterative Learning Algorithm

In order to compensate process model error caused by the time change, prevent strong shock and avoid system instability, it is required the setting values gradually add to the differential pressure control system. The differential pressure control process is according to the given target trajectory in advance from a steady state to another steady-state operation.

For the differential pressure optimal control process, the controller's setting point produced from the differential pressure setting algorithm is the step sequence of different amplitude. Therefore, the differential pressure amplitude of ideal trajectory will change with the controller set value. At the same time, the ideal trajectory of differential pressure will have good dynamic quality, such as little or no overshoot, quick response in speed, short transition in time and other good characteristics [5, 6]. Therefore, an iterative learning control algorithm (ILC) for variable differential pressure control can be used, and its schematic diagram is shown in Fig. 2, in which $p_1, p_2, \dots, p_k, p_{k+1}$ refers to the differential pressure value sequence, $y_{dk}(t)$ is the ideal trajectory, $y_k(t)$ is the real output value, and $x_k(t)$ is the input of the last iteration.

4.1 Selection of Ideal Trajectory of Differential Pressure

Pressure difference optimization algorithm is to produce differential pressure value sequence $p_1, p_2, \dots, p_k, p_{k+1}$, in order to prevent the strong shock and even instability in the system, it is requested every time finding the value to gradually match the pressure difference control system.

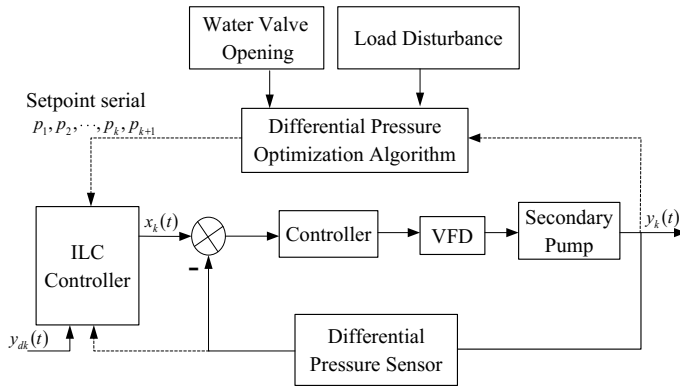


Fig. 2 Schematic diagram of the control system with iteration algorithm

For the differential pressure optimal control process, the ideal trajectory of differential pressure should have good dynamic quality. Thus, the selection of the ideal trajectory of the differential pressure method can be as follows: if the differential pressure corresponds to the set point p_1 , select $y_{d1}(t)$; then $y_{dk}(t)$ corresponds to the set point p_k , $y_{dk}(t) = p_k(1 - \frac{1}{(1+t)^n})$, ($n \geq 3$). For $y_{dk}(t) = p_k(1 - \frac{1}{(1+t)^n})$, the value of n is greater, the response speed is faster.

4.2 The Iterative Learning Control Unit

The differential pressure setting algorithm produces the sequence of controller setting values $p_1, p_2, \dots, p_k, p_{k+1}$. The iterative learning control unit produces current input $x_{k+1}(t)$ of the control system based on the controller setting value p_{k+1} , the real output value of differential pressure controller, the differential pressure ideal track $y_{dk}(t)$ and the input of the last iteration $x_k(t)$. To improve the efficiency of iterative learning control, weighted closed-loop PID type of iterative learning control algorithm can be used:

$$x_{k+1}(t) = \alpha_{k+1}x_k(t) + \Gamma_p e_{k+1}(t) + \Gamma_i \int_0^t e_{k+1}(\mu) d\mu + \Gamma_d \dot{e}_{k+1}(t) \quad (1)$$

$$\alpha_{k+1} = \frac{p_{k+1}}{p_k} \quad (2)$$

$$e_k(t) = y_{dk}(t) - y_k(t) \quad (3)$$

where

k the number of iterations;

α_{k+1} the weighted coefficient;

- $e_k(t)$ the error of differential pressure ideal track and actual differential pressures output;
- Γ_p learning gains for the proportion;
- Γ_i learning gains for the integration; and
- Γ_d learning gains for the differential

When the control input $x_{k+1}(t)$ is imposed on the differential pressure control system which purpose is to make the steady-state value of $x_{k+1}(t)$ is equal to the controller setting value p_{k+1} when the transient process is ended.

4.3 Simulation Experiments

In order to verify the correctness of the algorithm, the algorithm simulation needs to be done before the experimental research. The simulation background is a VAV air conditioning system installed in the laboratory of Xi'an University of Architecture & Technology. Identify the model describing the relationship of secondary pump frequency $u(t)$ and the differential pressure $p(t)$ with M sequence. The model is

$$p(t) = 0.78p(t-1) - 0.2273p(t-2) + 0.5115u(t-1) - 0.2093u(t-2)$$

Fitting the differential pressure $p(t)$ and flow rate $Q(t)$ relationship, it can be seen as

$$Q(t) = -0.0001p(t)^2 + 0.0335p(t) + 1.9369$$

The experiment is on a summer day from 8:00 to 20:00. The load changed because the outdoor temperature and indoor load of the air conditioning area changed. The initial chilled water flow and the water pump frequency can be calculated based on the outdoor weather and indoor conditions.

According to the changes of every terminal load, detecting the changes of every frozen water valve opening, using the differential pressure setting algorithm, we can get the differential pressure setting serial as {17, 17, 19, 19, 21, 23, 26, 29, 32, 36, 40, 45, 50, 55, 55, 50, 45, 40, 40, 45, 40, 36, 32, 29, 26}. The effect of differential pressure adopted iterative learning control is shown in Fig. 3.

If the iterative learning control is not adopted, the control effect comparing with the iterative learning control case is shown in Fig. 4.

Figure 4 shows that under the variable pressure difference control system with a differential value of 45 kPa, using the iterative learning control, the maximum deviation is 7.181 kPa, in case of non-using iterative learning control the maximum deviation is 9.168 kPa. The adjusted time is 428 s and 563 s corresponding to using and non-using iterative learning control algorithm respectively. Using iterative learning

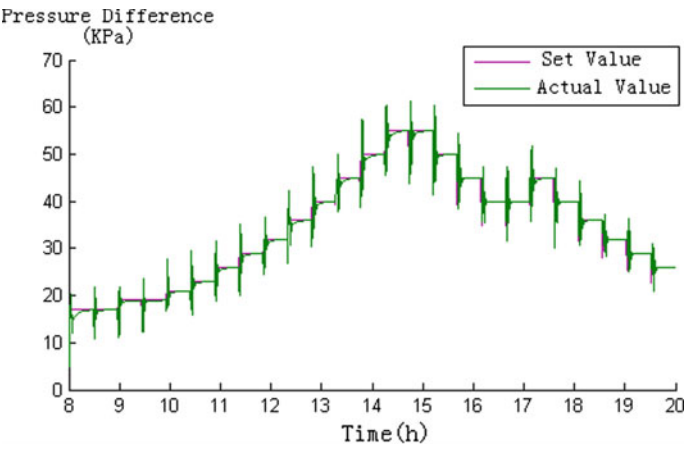


Fig. 3 Iterative learning control effect

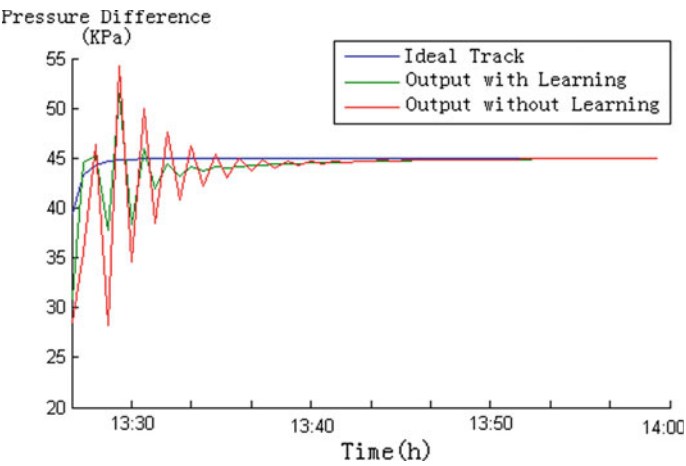


Fig. 4 Iterative learning control comparison chart

control, the steady-state error is 0.0142 kPa and 0.0173 kPa corresponding to using and non-using iterative learning control algorithm, respectively. And observing other indicators, the rise time is accelerated, transition time become shortened. In a word, the dynamic performance of the system is improved significantly.

5 Analysis of Energy Consumption

The differential pressure is associated with the frequency of secondary pump and energy consumption. For secondary pump experiment, the change of secondary pump frequency causes water supply and return pressure difference, and the frequency is obtained by the power measurement of secondary pump power. By the pressure difference and power experimental data, the relationship between differential pressure $p(t)$ and power $W(t)$ can be achieved as

$$W(t) = 0.0015P(t)^3 - 0.1293P(t)^2 + 15.0833P(t) + 322.4632$$

From experimental testing, the secondary pump power curve is shown in Fig. 5.

According to the actual change of pressure difference, the secondary pump power can be measured, which values are changed as {279.441, 274.441, 286.329, 279.441, 274.441, 309.615, 326.793, 343.861, 360.939, 383.937, 407.458, 438.033, 470.439, 511.237, 505.237, 479.439, 438.033, 407.458, 412.458, 438.033, 407.458, 383.804, 360.919, 348.650, 326.732}, its curve is shown in Fig. 5. It can be seen that using the variable and the constant differential pressure control, the energy consumption of secondary pumps is 9486.112 (W.h) and 13200 (W.h), respectively, under conditions of that, the differential pressure should meet the requirements of the maximum load and the measured load power 1100 (W.h). Thus, the energy-saving rate is 28.135%.

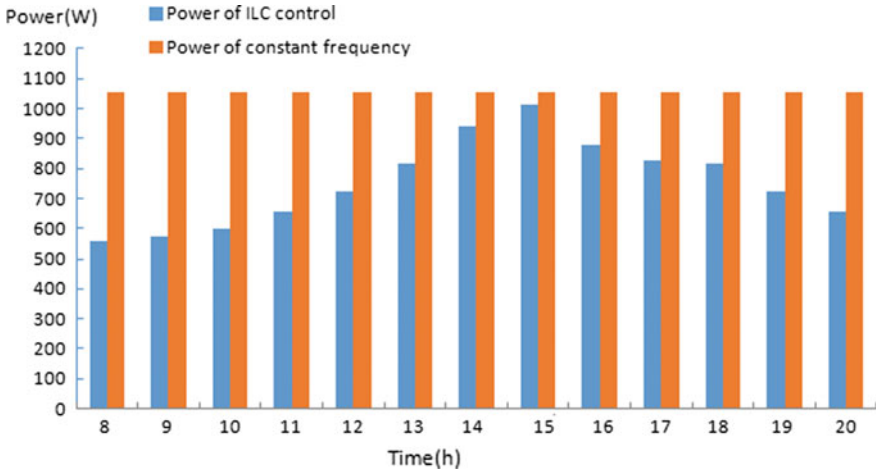


Fig. 5 Secondary pump power change curve

6 Conclusions

The variable water volume in the air conditioning system can be stably adjusted according to the change of the air conditioning system load with the variable differential pressure control algorithm based on the Iterative Learning controller algorithm (ILC). It automatically sets the differential pressure value, meets the demand of the air conditioning load from the users. It always makes the valve opening at the largest, it can lower the heat resistance loss, save the energy consumption of pump delivery. Compared with the conventional constant differential pressure control, this method can effectively reduce the energy consumption of the secondary pump.

Acknowledgements The work was supported by the Science and Technology Project of the Ministry of Housing and Urban-Rural Development, China (No. 2016 - K1 - 013) and the Key Laboratory of Gansu Advanced Control for Industrial Processes, China (Grant No. XJK201813).

References

1. Wang, Y., Liu, H.L., et al.: Adaptive algorithm design of smith predictor in HVAC control system. *Inf. Control* **40**(3), 408–412 (2011). (in Chinese)
2. Kuang, X.L.: The study on equality-renew method for two-level optimal control of central air-conditioning system. *Acta Sci. Nat. Univ. Sunyatsen.* **48**(3), 133–135 (2009)
3. Bian, W.F., Gong, Y.F.: Energy saving control and experimental research on VVW systems. *Energy Conserv. Technol.* **27**(3), 145–148 (2009)
4. Wang, J., Yuan, L.: Analysis of energy consume and control strategies of VVW system. *Refriger. Air Cond.* **24**(5), 47–50 (2010). (in Chinese)
5. Wang, Y.Q., Gao, F.R.: Survey on iterative learning control, repetitive control, and run-to-run control. *J. Process Control* **19**, 1589–1600 (2009). (in Chinese)
6. Bu, X.H., Hou, Z.S.: Iterative learning control for trajectory tracking of farm vehicles. *Acta Autom. Sin.* **40**(2), 368–372 (2014)

The Online Evaluation System of Chiller Plant in HVAC System



Jiaming Wang, Tianyi Zhao and Wei Li

Abstract In this paper, the online evaluation system for a chiller plant is proposed, which is significant for improving the energy efficiency and operation security of the chiller plant. For a general evaluation result, the system can estimate the chiller plant from three aspects: operating security, system performance, and cooling-supply quality, respectively. For each aspect, the paper presents a dynamic scoring method, which can convert the abstract evaluation indexes into the visualized contents by different evaluation demands. At the end of this paper, the effectiveness of the system is verified to perform excellently in practicability and operability, and it can also be used in most of the chiller plants.

Keywords Chiller plant · Evaluation method · Online evaluation system · Fault diagnosis · Scoring system

1 Introduction

The chiller plant is a complex and coupled system, which mainly consists of chillers, cooling water pumps (CWP), chilled water pumps (CHWP), cooling towers (CT), and water distributor and collector (WDC). For such a dynamical system, it's not scientific to evaluate it only by a single evaluation index. Likewise, it's unreasonable to evaluate the overall states in the security, efficiency, and performance of a chiller plant just by evaluating a single component. Thus, it's necessary to study the chiller plant in a deepgoing way, and take the impacts of each component's operating state into account for evaluating the overall system, and make the evaluation results as reasonable as possible.

In the study of evaluation of the chiller plant, there are many researches and reports have been done in the related field [1, 2]. However, most of the subjects mentioned in the studies are just for a single component, like the chiller or the pump, and the

J. Wang · T. Zhao (✉) · W. Li

Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian, China
e-mail: zhaotianyi@dlut.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_25

271

overall study of the chiller plant is few. The same situation in the fault diagnosis study of the chiller plant, many scholars have done enough researches in fault diagnosis of chillers, pumps, and sensors, but they all failed to consider the situation from the perspective of the overall chiller plant [3–5].

In this paper, we present an overall evaluation system which can evaluate the chiller plant from the aspects of operating security, system performance, and cooling-supply quality. Besides, for a better application of this overall evaluation system, an online evaluation system is proposed, which can gather the operation data of a chiller plant and give the evaluation results in time, also can find out the operating problems of the chiller plant timely. Due to the different evaluation demands needed by the different chiller plants or users, we also propose a dynamic scoring system, in which the weight of index, the boundary value of the evaluation interval, and the score value of each index can be defined or modified by the users, and this makes the system to be applied in most of the chiller plants.

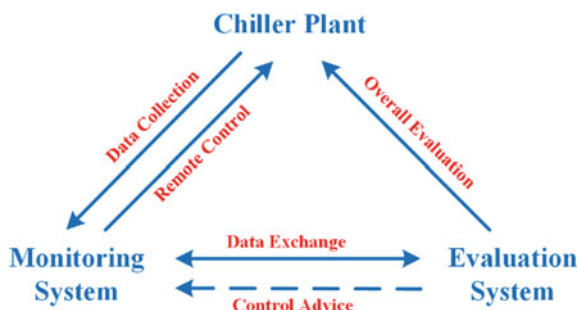
2 Architecture of the Evaluation System

The online evaluation system is just like a subsystem of the monitoring system, these two systems are mutually independent in function, and also mutually connected in the underlying structure, the relationship between these two systems is shown in Fig. 1. Here, we will introduce its architecture in two aspects. One is the architecture of software, which shows the interchangeable connections among logical structuring mechanisms. And, the other one is the architecture of function, which shows how the organizational structures of the online evaluation system are arranged.

2.1 Software Architecture

In this paper, we use the Siemens WinCC to the user interface programming, and the PLC is used to collect data and fulfill the logical operation for the chiller plant.

Fig. 1 The relationship among the systems



The software configuration of the online evaluation system for the chiller plant is shown in Fig. 2. It can be found that a three-layer architecture is applied in this online system. This typical multilayer architecture is widely used in most software development, which is composed of a presentation layer, a logic layer, and a data layer [6].

The data layer includes the mechanisms of data access and data persistence. At first, the monitoring system collects the data gathered by the sensors of the field and stores the data into the database. Then, the data is transferred to the logic layer for processing, and eventually delivered to the users.

In the logic layer, the PLC and WinCC can both access to the database and processing the data. In this paper, the PLC and WinCC are in charge of different parts of the logic process. The PLC is mainly in charge of the judgment in decision and some complex processes, while the WinCC is mainly in charge of some sample programming and calculation. After the data processing, the processed data will be transferred to the WinCC server for users.

In the presentation layer, the C/S (Client/Server) model is applied in the system framework, which supports multiclients using the online evaluation system at the

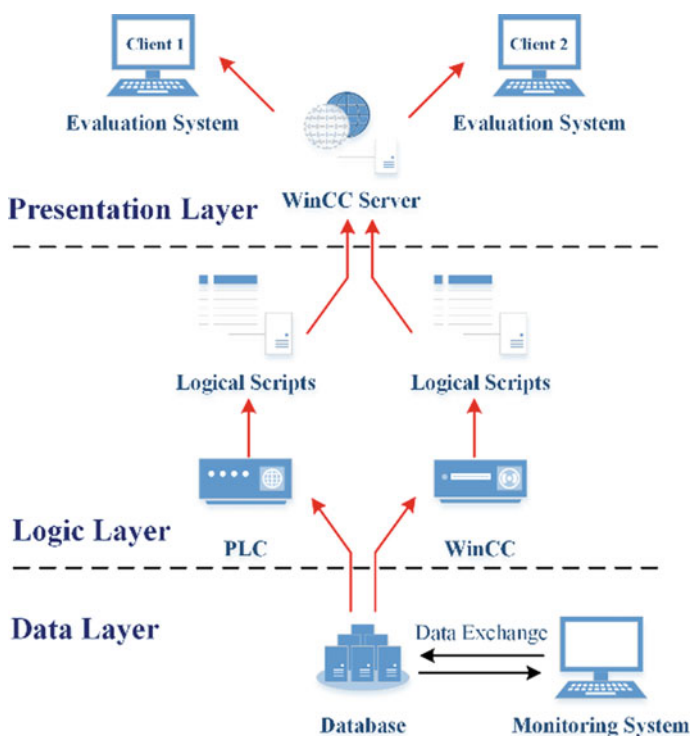


Fig. 2 The software configuration of the online evaluation system

same time. The interface of the online evaluation system is programmed by the WinCC software.

2.2 Functional Architecture

The functional architecture of the online evaluation system is shown in Fig. 3. In this paper, we design four main menus for the online evaluation system, including the overview of the chiller plant, operating security, system performance, and cooling-supply quality. In the overview menu, it's able to show the global operation states of the chiller plant briefly. For more detailed information, the users can check in the other three main menus. For each evaluation aspect, we also define several submenus on the basis of different main components in the chiller plant. And with these submenus, it's accessible for the users to check on the operating states of a single component.

The interface of the online evaluation system can be divided into five functional zones, besides the main menu and the submenu, it also includes the user login zone, the systemic function zone, and the detail presentation zone.

3 Evaluation Contents and Methods

3.1 Notations

S: the total score of chiller plant;

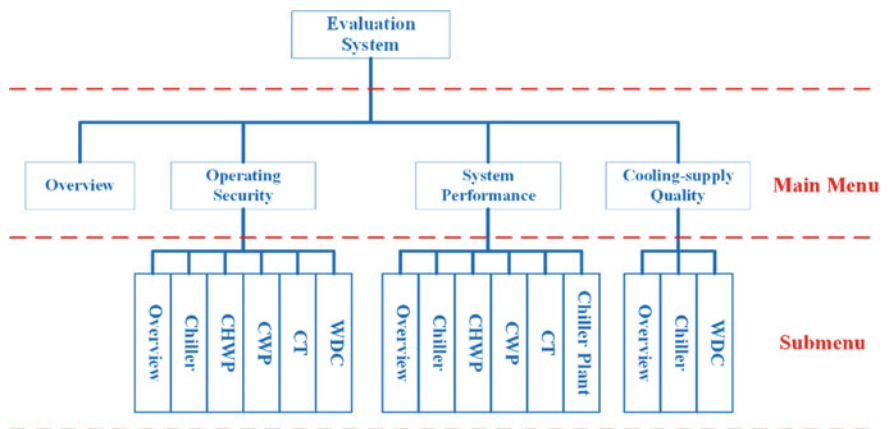


Fig. 3 The functional architecture of the online evaluation system

$S_s, S_p,$ and S_q : the scores of operating security, the performance, and the cooling-supply quality, respectively;

$W_s, W_p,$ and W_q : the weights of operating security, the performance, and the cooling-supply quality, respectively;

$S_1-S_5, P_1-P_5,$ and Q_1-Q_5 : the scores of operating security, the performance, and the cooling-supply quality for each component;

$s_1-s_5, p_1-p_5,$ and q_1-q_5 : the weights of operating security, the performance and the cooling-supply quality for each component;

$S_{11}-S_{17}, P_{21}-P_{29},$ and $Q_{21}-Q_{25}$: the scores of basic index subordinate to operating security, the performance, and the cooling-supply quality; and

$s_{11}-s_{17}, p_{21}-p_{29},$ and $q_{21}-q_{25}$: the weights of basic index subordinate to operating security, the performance, and the cooling-supply quality.

3.2 Evaluation Contents

For building a clear evaluation system, we’ve defined three levels of the index, and the sample tree view of the evaluation system is shown as Fig. 4. In this paper, we define 3 primary indexes (I index), 13 secondary indexes (II index), and 73 basic indexes (III index), and all of these indexes can be obtained online easily. The evaluation contents can be described as below.

Operating Security. The chiller plant consists of not only the main components mentioned above, but also many auxiliary devices (such as controllers, sensors, etc.)

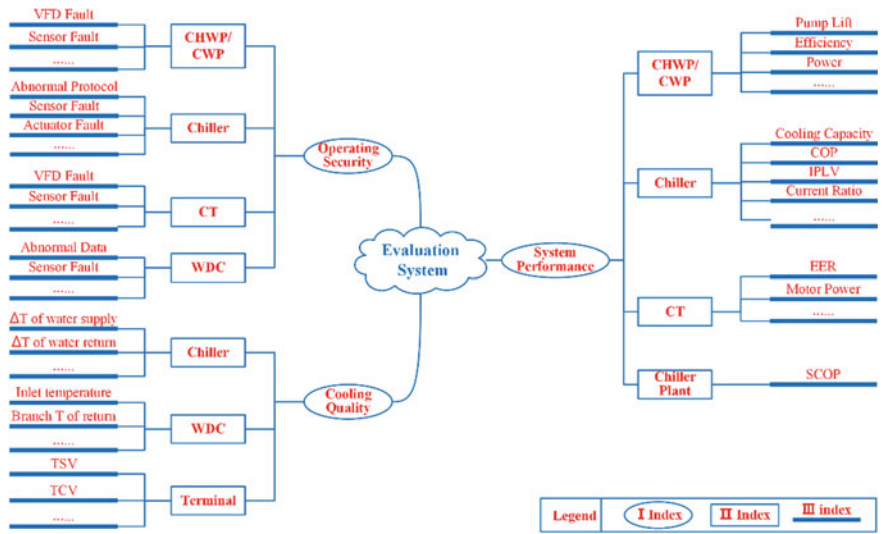


Fig. 4 The sample tree view of the evaluation system

which are indispensable to the chiller plant's monitoring and operating. The aim of the operating security is to find out the chiller plant's efficiency and security problems which may be caused by the components' abnormal state. For an overall evaluation in operating security, the secondary indexes are located in the main components of the chiller plant, and are, respectively, named as operating security of the chiller, operating security of CHWP, operating security of CWP, operating security of CT, and operating security of WDC. For each secondary index, there are several basic indexes subordinate to it, such as the sensor fault, the abnormal data, etc.

System Performance. In this part, the evaluation is mainly focused on the performance of the chiller plant system. Just like the aim of the operating security, the aim of the system performance is to evaluate the chiller plant from the perspective of energy efficiency. Which also includes five secondary indexes indicated the performance of chiller, CHWP, CWP, CT, and chiller plant. For each secondary index, there are also several basic indexes subordinate to it, such as the COP of the chiller, EER of the cooling tower, etc.

Cooling-supply Quality. For an overall evaluation, we also take into account of the user experience with the HVAC terminals, so the concept of the cooling-supply quality is given in this paper. The aim of the cooling-supply quality is to reflect the comfort level of the indoor thermal and humid environment. The evaluation will be conducted on the chiller, WDC, and HVAC terminal. For each secondary index, there are also several basic indexes subordinate to it, such as the temperature difference of chilled water supply/return, thermal sensation vote (TSV), etc.

3.3 Evaluation Methods

In order to present an intuitionistic evaluation results for the users, we apply a dynamic scoring system to the online evaluation system. By this way, the complex evaluation contents will convert into some sample scores. What the user directly gets is just a series of numbers, which synthesizes the overall operating states of the chiller plant. In this paper, we set the scores of the three primary indexes as S_s , S_p , and S_q , and the relevant weight as W_s , W_p , and W_q , then the final score of the chiller plant can be expressed as

$$S = W_s S_s + W_p S_p + W_q S_q \quad (1)$$

As for the values of S_s , S_p , and S_q , it can be obtained by the scores of secondary indexes. Here, we cite an instance of S_s . We assume that the scores of the operating security were s_1 – s_5 and the relevant weights for each index were s_1 – s_5 , then the score of the operating security could be expressed as

$$S_s = \sum_{i=1}^5 (s_i \times S_i) \quad (2)$$

Then, the same process for the value of S_1 – S_5 , it can be obtained by the scores of basic indexes. Here, we cite S_1 as an example. Assume that this secondary index got seven basic indexes subordinate to it, the scores of the basic indexes were S_{11} – S_{17} , and the weights were s_{11} – s_{17} , then the score of S_1 could be expressed as

$$S_1=\sum_{i=1}^7(s_{1i}\times S_{1i}) \tag{3}$$

4 Evaluation Case

Here, we set an example to illustrate the scoring procedure. We assume that the current situation of the chiller plant can be described as follows.

In the operating security part, as shown in Table 1, there are two basic indexes of chiller showing abnormally, the rest of the indexes all get 100 points.

Then, $S_2 = S_3 = S_4 = S_5 = 100$, $S_{11} = S_{12} = S_{14} = S_{15} = S_{17} = 100$, $S_{13} = 100 - 100 = 0$, $S_{16} = 100 - 30 - 30 = 40$, and the scores can be expressed as

$$\begin{aligned} S_1 &= 0.1 \times 100 + 0.1 \times 100 + 0.2 \times 0 + 0.1 \times 100 \\ &\quad + 0.2 \times 100 + 0.1 \times 40 + 0.2 \times 100 = 74 \end{aligned} \tag{4}$$

$$S_s = 0.2 \times 74 + 0.2 \times 100 + 0.2 \times 100 + 0.2 \times 100 + 0.2 \times 100 = 94.8 \tag{5}$$

In the system performance part, as shown in Table 2, there are three basic indexes of the chiller showing abnormally, the rest of the indexes all get 100 points.

Then, $P_1 = P_3 = P_4 = P_5 = 100$, $P_{21} = P_{24} = P_{26} = P_{27} = P_{28} = P_{29} = 100$, $P_{22} = 80$, $P_{23} = 80$, $P_{25} = 50$, and the scores can be expressed as

Table 1 The example scoring contents in operating security of chiller

Index	Weight	Initial value	Evaluation contents	Scores
Abnormal data S_{13}	0.2	100	Higher condensing pressure	−100
Sensor fault S_{16}	0.1	100	1# Sensor fault	−30
			2# Sensor fault	−30

Table 2 The example scoring contents in system performance of chiller

Index	Weight	Initial value	Evaluation contents	Scores
COP P_{22}	0.2	0	$[a_{22}, b_{22})$	80
IPLV P_{23}	0.05	0	$[a_{23}, b_{23})$	80
Current ratio P_{25}	0.2	0	$\geq b_{25}$	50

Table 3 The example scoring contents in cooling-supply quality of WDC

Index	Weight	Initial value	Evaluation contents	Scores
Temperature of water return Q_{23}	0.2	0	$[a_{23}, b_{23})$	80

$$P_2 = 0.1 \times 100 + 0.2 \times 80 + 0.05 \times 80 + 0.05 \times 100 + 0.2 \times 50 + 0.1 \times 100 \times 4 = 85 \quad (6)$$

$$S_p = 0.2 \times 100 + 0.2 \times 85 + 0.2 \times 100 + 0.2 \times 100 + 0.2 \times 100 = 97 \quad (7)$$

In the cooling-supply quality part, as shown in Table 3, there is one basic index of WDC showing abnormally, the rest of the indexes all get 100 points.

Then, $Q_1 = 100$, $Q_{21} = Q_{22} = Q_{24} = Q_{25} = 100$, as for Q_{23} , assume one of the three return pipes got abnormally, then the value of Q_{23} will be obtained by the arithmetic average of these three, so $Q_{23} = (80 + 100 + 100)/3 = 93.3$, and the scores can be expressed as

$$Q_2 = 0.2 \times 100 + 0.2 \times 100 + 0.2 \times 93.3 + 0.2 \times 100 + 0.2 \times 100 = 98.7 \quad (8)$$

$$S_q = 0.5 \times 100 + 0.5 \times 98.7 = 99.3 \quad (9)$$

Then we got the value of S_s , S_p , and S_q , and the final scores can be obtained by Eq. (1)

$$S = 0.4 \times 94.8 + 0.4 \times 97 + 0.2 \times 99.3 = 96.6 \quad (10)$$

5 Conclusions

In this paper, an online evaluation system of the chiller plant is proposed. By introducing the architecture, evaluation contents, and evaluation methods of the online evaluation system, this paper analyzes the feasibility of this online system. And, the application of the dynamic scoring system makes the online evaluation system much more visualized and practical, which also makes the system can be modified by different user demands and fit most of the chiller plants.

Acknowledgements This work is supported by the National Key Research and Development Project of China (Grant No. 2017YFC0704100, entitled “New Generation Intelligent Building Platform Techniques”) and “the Fundamental Research Funds for the Central Universities” (Grant No. DUT17ZD232).

References

1. Zmeureanu, R., Peragine, C.: Evaluation of interactions between lighting and HVAC systems in a large commercial building. *Energy Convers. Manag.* **40**(11), 1229–1236 (1999)
2. Deng, S.M., Burnett, J.: A study of energy performance of hotel buildings in Hong Kong. *Energy Build.* **31**(1), 7–12 (2000)
3. Salsbury, T.I., Diamond, R.C.: Fault detection in HVAC systems using model-based feedforward control. *Energy Build.* **33**(4), 403–415 (2001)
4. Reddy, T.: Evaluation and assessment of fault detection and diagnostic methods for centrifugal chillers-phase II. ASHRAE Research Project RP-1275 (2006)
5. Zhao, Y., Xiao, F., Wen, J.: A robust pattern recognition-based fault detection and diagnosis (FDD) method for chillers. *HVAC&R Res.* **7**(20), 798–809 (2014)
6. Chen, Y., Mu, X., Zhang, J., Lu, Z., Ma, J.: The software subsystem of building power consumption monitoring and management system, pp. 1–4. IEEE (2010)

The Power Consumption Model of Chiller with Elman Neural Networks for On-line Prediction and Control



Zhiyang Jia and Tianyi Zhao

Abstract In this paper, a new steady-state power consumption model using the Elman Neural Network (ENN) is proposed. The model is dependent on the external parameters of chiller, which are easily monitored and which are related to the global optimization of an air-conditioning water system. The simulation results show that the model can complete the training process within 3 s. In addition, it can be seen that the results of the model are in good agreement with the experimental values with the majority of the *RE* values within $\pm 3\%$. Therefore, this model is suitable for on-line prediction of the power consumption of chiller in on-field engineering.

Keywords Elman neural network · On-Line prediction · Chiller · Global optimization · Power consumption

1 Introduction

As we know, the power consumption of a chiller accounts for a large portion of the total energy consumption of an air-conditioning water system. In on-field engineering, the real-time power consumption of each chiller with different operating conditions is important to evaluate the power consumption of an air-conditioning water system.

Many researchers paid their attention to modeling and developed numerous models, which can be divided into two categories: gray-box model and black box model. For the gray box model, the performance model of a chiller developed by Ng et al. [1] and Lee [2] took the cooling capacity, the inlet chilled water temperature and the inlet cooling water temperature as input parameters. Besides, some researchers developed the power consumption model of a chiller using the cooling capacity, the inlet chilled water temperature, and the inlet cooling water temperature as the independent parameters, which have been applied in the optimal operation of an air-conditioning

Z. Jia · T. Zhao (✉)

Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian, China
e-mail: zhaotianyi@dlut.edu.cn

water system [3–5]. It is no doubt that the gray box model has satisfactory accuracy. However, developing the gray box model needs numerous parameters of a chiller, which is anyway difficult to obtain directly in many on-field operating chiller plants.

Compared with the gray box model, developing the black box model only depends on the operation data of chiller. So, it is convenient for the on-field engineers to use. As a typical black box model, the neural network model has been studied by many researchers. Swider et al. [6] developed the performance model of chiller based on radial basis neural network. However, few researchers used Elman Neural Network (ENN) to predict the chiller performance.

Based on the above discussion, this paper attempts to predict the power consumption of a chiller using ENN. At first, the input and output parameters of the ENN model are determined based on the online prediction of the power consumption of a chiller in many on-field engineering. Then, based on the selected input and output parameters, we determine the architecture of the ENN model. Finally, the ENN model is trained and tested by the sample data obtained from the field measures.

2 Description of the Elman Neural Network

ENN is one kind of typical dynamic recurrent neural network. It is different from Feedforward Neural Network (FNN). It is known that ENN has the function of recursion and dynamic memory. In view of the power consumption of chillers is influenced by many factors, and they are multivariables, strongly coupled and seriously nonlinear. So when we use ENN to predict the power consumption of a chiller, this model requires more simple structural parameters and shorter training time, relatives to FNN.

2.1 Notations

L_{chw}	the chilled water flow rate, m^3/s ;
Δp_{cw}	the cooling water pressure difference of chiller, kPa;
N_e	the experimental power consumption of chiller, kW;
N_m	the modeling power consumption of chiller, kW;
S	the resistance characteristic coefficient, kg/m^7 ;
DPS_{CHW}	the pressure difference sensor of chiller for the chilled water;
DPS_{CW}	the pressure difference sensor of chiller for the cooling water;
TS_{CHWi}	the temperature sensor for the inlet chilled water;
TS_{CHWo}	the temperature sensor for the outlet chilled water;
TS_{CWi}	the temperature sensor for the inlet cooling water;
TS_{CWo}	the temperature sensor for the outlet cooling water;
ET_{CH}	the electrical power transducer

3 Model Input and Output Parameters

As we know, the power consumption of a chiller is influenced by many factors. However, there are some influencing factors that are inconvenient to measure directly and accurately in many on-field engineering. Therefore, in order to make ENN model have good predictive accuracy and is easily applied to online prediction in many on-field engineering, those influencing factors that are not easily directly and accurately measured cannot be used as the model input parameters. As shown in [7], the researchers found that the power consumption of chiller is mainly related to the *PLR* (Partial Load Ratio), the chilled water flow rate, the inlet and outlet chilled water temperature, the cooling water flow rate, and the inlet and outlet cooling water temperature. As mentioned above, all influencing factors can be measured directly by all kinds of sensors except the *PLR*, which is difficult to measure directly in many on-field engineering. In addition, although the water flow rate can be measured directly, it is often difficult to measure accurately. Equation (1) shows that the water flow rate is related to the water pressure difference of chillers that can be measured directly and accurately. Therefore, this paper takes the chilled water pressure difference, the inlet and outlet chilled water temperature, the cooling water pressure difference, and the inlet and outlet cooling water temperature as input parameters.

$$\Delta p_{\text{chw}} = SL_{\text{chw}}^2 \quad (1)$$

4 Description of Measurement System and Experimental Data

Using the datasets from the on-field engineering of a hospital located in Zhejiang (China), this paper develops the ENN model for chiller. Figure 1 shows the diagram of air-conditioning remote monitoring and control system where we can obtain the experimental data. In this research work, in order to measure the real-time operational parameters, the electrical power transducers, the temperature sensors, and the pressure difference sensors were installed. The measurement accuracy for the electrical power transducers is $\pm 1\%$. The measurement accuracy for the temperature sensors is $\pm 0.1\text{ }^\circ\text{C}$, and for the pressure difference sensors are $\pm 0.002\text{ MPa}$. All the parameters were measured and stored continuously in a web-based data acquisition with a 5 min sampling period.

Table 1 provides the details of the number of data sets and parameters ranges for the chiller.

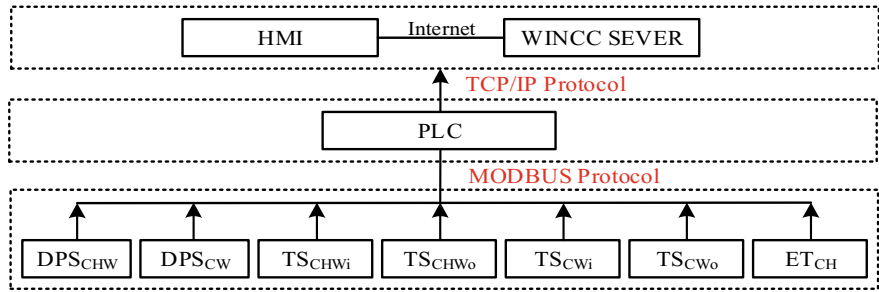


Fig. 1 The diagram of air-conditioning remote monitoring and control system

Table 1 Parameter ranges and total data sets

Parameters	Training data		Testing data	
	1800 Data sets		466 Data sets	
	Min	Max	Min	Max
Δp_{chw} (kPa)	39.90	117.20	46.60	99.00
Δp_{cw} (kPa)	28.30	186.80	160.90	174.90
T_{chwi} (°C)	10.63	14.91	11.02	13.87
T_{chwo} (°C)	7.52	12.40	7.70	9.72
T_{cwi} (°C)	26.05	32.18	28.22	31.54
T_{cwo} (°C)	27.68	38.39	30.43	35.08
N_e (kW)	138.51	293.71	167.54	279.54

5 Development of the ENN Model

Figure 2 provides the main procedures for developing the ENN model. The ENN model is implemented under MATLAB2015b and consists of four distinctive parts: (1) Date preprocessing; (2) The network architecture of the ENN model; (3) ENN model training; (4) ENN model testing. The detailed descriptions of the four distinctive parts are shown in Sects. 5.2, 5.3, 5.4, and 5.5.

5.1 Evaluating Indicators of ENN Model

This paper uses relative error (RE) and root-mean-square error (CV) to indicate how well the ENN model can satisfy the prediction. The RE value and CV value can be calculated as shown in Eqs. (2) and (3). It has been proven that the empirical model with 3–5% CV for performance prediction is acceptable in practical applications [8].

$$RE = (y_{mi} - y_{ei})/y_{ei} \times 100\% \tag{2}$$

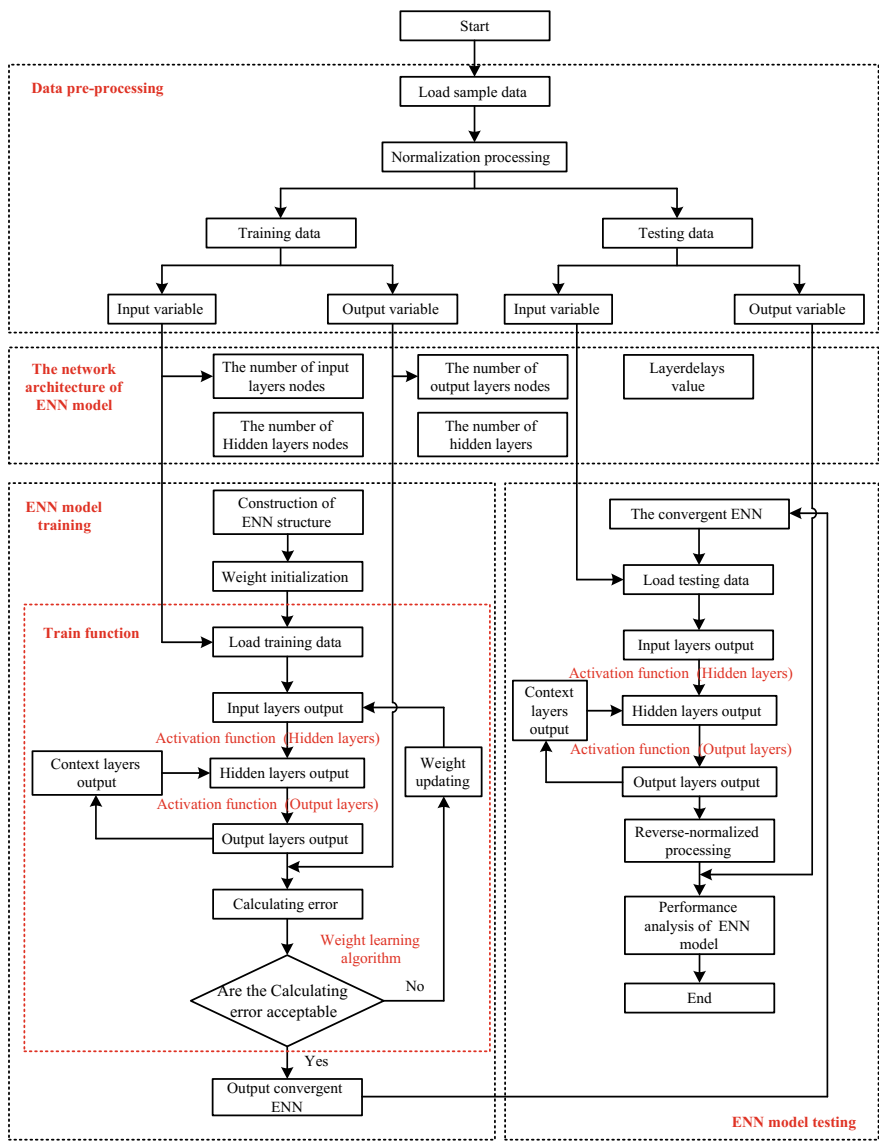


Fig. 2 Procedure of developing the ENN model

$$CV = RMSE / ((\sum y_{ei}) / n) \times 100\% \quad (3)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (y_{mi} - y_{ei})^2 \right]^{\frac{1}{2}} \quad (4)$$

where y_{mi} and y_{ei} are the modeling and the experimental values, respectively, n is the number of the testing (or training) data sets, and the subscript i denotes the corresponding value from 1 to n sets.

5.2 Data Preprocessing

Because of the different dimensions of the input and output parameters, the training time will increase when the ENN model is trained. Consequently, in order to eliminate this unfavorable influence, the “mapminmax” function in MATLAB 2015b is used to normalize the data sets before the ENN is started. After this procedure, the data sets are limited in $[-1, 1]$. Simulation result shows that the ENN model has higher prediction precision after the normalization processing with “mapminmax” function.

5.3 The Network Architecture of ENN Model

The network architecture of ENN model is depicted in Fig. 3. It consists of one input layer, one hidden layer, one context layer, and one output layer. Because each input and output node corresponds to an element of an input and output parameter, the input and output layer consists of 6 nodes and 1 node, respectively. In addition, the hidden layer consists of 11 nodes based on trial and error. For the number of context nodes, it is the same as the number of hidden nodes.

5.4 ENN Model Training

Figure 2 (ENN model training) shows the detailed procedure for training the ENN model

Step 1: This step establishes 6-11-1 ENN architecture based on the “elmannet” function in MATLAB2015b.

Step 2: The “init” function in MATLAB2015b is used to initialize the weights of each layer.

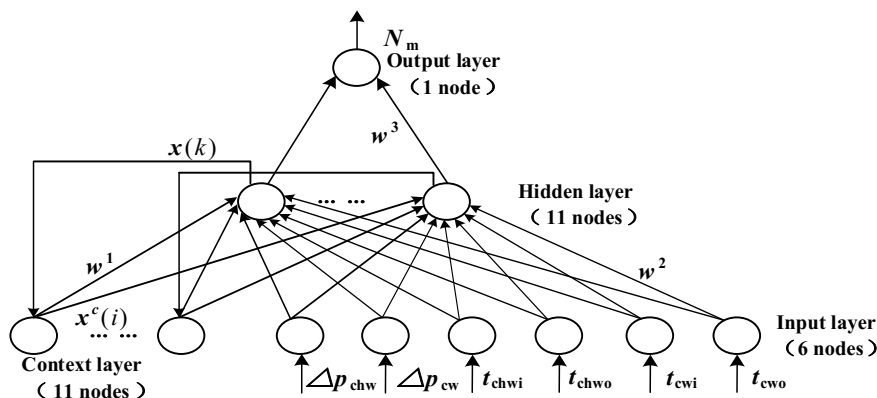


Fig. 3 The network architecture of ENN model

Step 3: The training data are put into the ENN. Then, using the “train” function in MATLAB2015b trains the ENN model. In this procedure, the training algorithm for the ENN is “traingdx”.

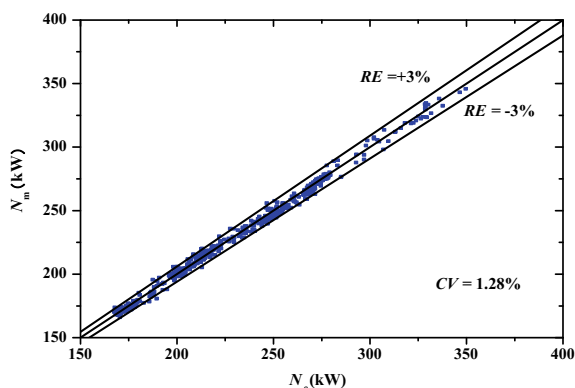
Step 4: The training will complete after the ENN reach to the best training performance that the error between actual and target output approach to 0.007. Then, the ENN weights are saved.

5.5 ENN Model Testing

The ENN model is tested as illustrated in Fig. 2 (ENN model testing). Putting the testing data sets into the convergent ENN model after the procedure of ENN model training, the modeling power consumption will be obtained.

Figure 4 depicts a comparison between the results of the modeling and experimental power consumption where 446 steady-state data points were realized for the chiller. As shown in Fig. 4, the horizontal axis represents the experimental values, while the vertical axis represents the modeling values. The central solid line represents that the experimental values are equal to the modeling values. It can be seen that the results of the model are in good agreement with the experimental values with the majority of the RE values ranging from -3 to 3% . In addition, the CV value is 1.28% less than 3% . Consequently, the ENN model has satisfactory accuracy.

Fig. 4 Comparison of modeling and experimental results of the power consumption



6 Application of the ENN Model

Figure 5 depicts the method of how we use ENN model to predict the power consumption of chiller in on-field engineering. As shown in Fig. 5, we need to use OPC technology to establish the real-time data link between MATLAB and WINCC, where WINCC serves as OPC server (MATLAB serves as OPC Client). The main procedures of the method are shown as follows:

Step 1: Based on the air-conditioning remote monitoring and control system (Fig. 1), we can obtain the real-time operational data of chiller in HMI and WINCC.

Step 2: The real-time operational data can be passed to MATLAB workspace based on OPC technology. Then, we can complete the ENN model training in MATLAB.

Step 3: The developed ENN model can calculate the power consumption of chiller based on the real-time operational data from WINCC, and then pass the calculation results to the WINCC.

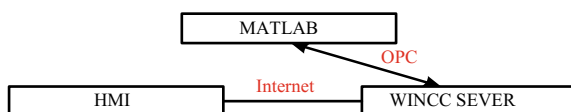
7 Conclusions

Based on the results above, the main conclusions are the following:

The modeling method for the power consumption of chiller proposed in this paper is feasible.

The proposed ENN model has a good predictive precision with the RE below 3% for the majority of the testing data sets, and can complete the training process within

Fig. 5 The diagram of application of the ENN model



3 s. Furthermore, the ENN model was developed by the water pressure difference of chillers without measuring the water flow rate, which is difficult to measure accurately in many on-field engineering. Therefore, the ENN model is suitable for the online prediction of the power consumption of chiller in on-field engineering.

Acknowledgements The work is supported by National Key Research and Development Project of China (Grant No. 2017YFC0704100, entitled “New Generation Intelligent Building Platform Techniques”) and “the Fundamental Research Funds for the Central Universities” (Grant No. DUT17ZD232).

References

1. Ng, K.C., Chua, H.T., Ong, W., Lee, S.S., Gordon, J.M.: Diagnostics and optimization of reciprocating chillers: theory and experiment. *Appl. Therm. Eng.* **17**(3), 263–276 (1997)
2. Lee, T.S.: Thermodynamic modeling and experimental validation of screw liquid chillers. *Ashrae Trans.* **110**, 206–216 (2004)
3. Ma, Z., Wang, S.: Supervisory and optimal control of central chiller plants using simplified adaptive models and genetic algorithm. *Appl. Energy* **88**(1), 198–211 (2011)
4. Misenheimer, C., Terry, S.D.: The development of a dynamic single effect, lithium bromide absorption chiller model with enhanced generator fidelity. *Energy Convers. Manag.* **150**, 574–587 (2017)
5. Liu, Z., Tan, H., Luo, D., et al.: Optimal chiller sequencing control in an office building considering the variation of chiller maximum cooling capacity. *Energy Build.* **140**, 430–442 (2017)
6. Swider, D.J., Browne, M.W., Bansal, P.K., et al.: Modelling of vapour-compression liquid chillers with neural networks. *Appl. Therm. Eng.* **21**(3), 311–329 (2001)
7. Zhou, X., Cai, P., Lian, S., et al.: Research on COP prediction model of chiller based on PSO-SVR. *J. Refrig.* (2015)
8. Hydeman, M., Sreedharan, P., Webb, N., Blanc, S.: Development and testing of a reformulated regression-based electric chiller model. *Ashrae Trans.* **108**(2), 1118–1127 (2002)

Part V
Information Technology and Intelligent
Transportation Systems

Research on Driving Decisions in Winter and Summer Based on Survey Date



Mingxia Huang, Haiqiang Zhang and Zhu Bai

Abstract It is a fact that comparing the effect of signal change (especially changing from green to yellow) on driving decisions in different seasons is a significant way to promote traffic safety. In this study, by using video recording, a variety of data through observing vehicles on four intersections in Shenyang were obtained to analyze drivers' driving decisions during green countdown signal change in summer and winter. Based on data including speed of a vehicle, the remaining time of countdown timers, and the Electronic Police (E-Police), and a Logistic model was established to analyze the influence of signalized intersections on drivers' decision-making in summer and winter. The result shows that the stop decision is only impacted when the countdown signal is coexistent with E-Police in summer. Also, both the countdown signal and the E-Police have a great influence on the stopping decision in winter.

Keywords Drivers' decision · Countdown signal · E-Police · Inter-green period

1 Introduction

The rapid socioeconomic development has led to a substantial increase in the number of vehicles in cities. This result in a large number of problems such as traffic congestion and an increasing number of traffic accidents, which negatively affect the smooth operation of cities. If there is no information provided at intersections which causes drivers' difficulties to decide to stop or through the intersection. However, a wrong decision may lead to a variety of traffic accidents, such as rear-ending with other vehicles, collision, scratching, and so on. In order to prevent traffic accidents and improve the capacity of the road network, many countries have introduced the countdown signal. Countdown signal is considered as one of the best measures, which provided the proper timing for drivers' decisions to avoid these traffic accidents.

M. Huang · H. Zhang (✉) · Z. Bai
School of Transportation Engineering, Shenyang Jianzhu University, Shenyang, China
e-mail: 1079360134@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_27

In order to assist the driver in making a correct driving decision during the phase conversion process, various countries have tried to make changes in traffic signals, such as taking the blinking signal and the countdown signal that play the role of an early warning indicator for the driver. A relatively new development is the countdown signal that shows the remaining time, which is displayed as a digital signal in a few seconds with the signal light until the next signal switches. The countdown signal can be divided into two categories: green signal countdown display (GSCD) and red signal countdown display (RSCD). At present, scholars from different countries have different attitude toward using countdown signals. Peng and his colleagues [1] indicated that the red signal countdown can significantly reduce departure time and delay time of vehicles at the intersection; Islam et al. [2] insisted that the green signal countdown is able to improve the safety of intersections of America through the development and verification of predictive models. Some scholars have disputed about the positive impact of the signal countdown, especially the positive impact of GSCD. Yang and his collaborates [3] insisted that the current control effect of signal countdown still needs further exploration and research; Cao and his colleagues [4] concluded that dangerous driving behaviors, such as red lights running and illegal lane changing, occur at the end of the green light period; Girish et al. [5] found that with the countdown signal application, irregularities are increasing at the end of red signal countdown.

Consequently, there are a large number of studies related to countdown signal across the world, but few of them are involved in the comparative analysis of the countdown signal effect under road conditions in different seasons. Pan et al. [6] considered the value of the vehicle and applied it to the impact analysis of the driving behavior at the countdown signalized intersections. The results showed that the value of the vehicle influences driving behavior at the countdown signalized intersections. Based on this study, the paper introduced the road conditions of different seasons and analyzed its impact on driving decisions. Due to the driver's different driving decisions under the road condition in winter, the impact of the countdown signal on the driver will be different from that in summer.

2 The Survey Data

The data surveyed in this paper was conducted on August, 5 2017 (summer data) and January, 20 2018 (winter data), and four intersections in Shenyang—the capital city of Liaoning Province in China. The range from the stop line up until 90 m upstream was studied in this research because this distance can be read by the driver. The studied vehicles will be called “subject vehicles” in this paper. Also, we used the video survey method to obtain the data of the subject vehicle. The surveys were conducted on weekdays during an off-peak hour when the traffic volume is neither too high nor too low. This is because valid and appropriate data could be obtained during this period. Most intersections have no all-red period in Shenyang. Similar to most cities, the signal cycle starts from green to yellow, and then to red. The

Table 1 Intersections characteristics and traffic conditions

Survey date	Intersection name	#Lanes	Number of cycles	Direction (through traffic)	Green (sec)	Yellow (sec)	Traffic volume (pcu/h)
5 Aug 2017	Wenhui	3	4	Southbound	45	3	184
	Wenhudong	4	4	Westbound	53	3	175
10:00–11:00 (Summer)	Jinfandong	2	3	Southbound	32	2	157
	Tiancheng	3	3	Southbound	42	3	168
20 Jan 2018	Wenhui	3	4	Southbound	45	3	165
	Wenhudong	4	4	Southbound	50	3	166
10:00–11:00 (Winter)	Jinfandong	2	3	Southbound	32	2	168
	Tiancheng	3	3	Southbound	42	3	125

existence of E-Police is also taken into account. Table 1 describes the features of surveyed intersections and traffic conditions on the survey date.

3 Based on Survey Date Analyze Vehicle Driving Characteristics

3.1 Speed Profile of Stopping Vehicles

The subject vehicle in this analysis was observed when they enter our observing area (0–90 m far away from the stop line). The speed of the subject vehicle is shown in Fig. 1a, b for summer and winter, respectively. T, NT, P, and NP indicate the intersection with countdown signal, without countdown signal, with E-Police, and without E-Police separately. The dotted line indicates the speed at the intersection with the countdown signal. In summer, the deceleration at the intersection with the countdown signal and E-Police (Wenhui) was not so abrupt, and the speed gradually decreased between 40 and 90 m before the stop line. The deceleration at the intersection only with E-Police (Wenhudong) also relatively became smooth. At the intersection only with the countdown signal (Jinfandong), a large deceleration occurred, the speed was basically unchanged before the yellow but suddenly decreased after the yellow started. In the winter, the three intersections had the same pattern, and the deceleration was dramatic at intersections where there are neither countdown signals nor E-Police (Tiancheng). It is worth noting that data about the subject vehicle at the Tiancheng intersection was not obtained successfully obtained in summer and therefore, it was not included in the analysis.

At 90 m before the intersection, comparing between summer and winter, we can find that the speed of vehicles in summer was generally higher than that in winter in the figure. In addition, regardless of the season, in the absence of the countdown

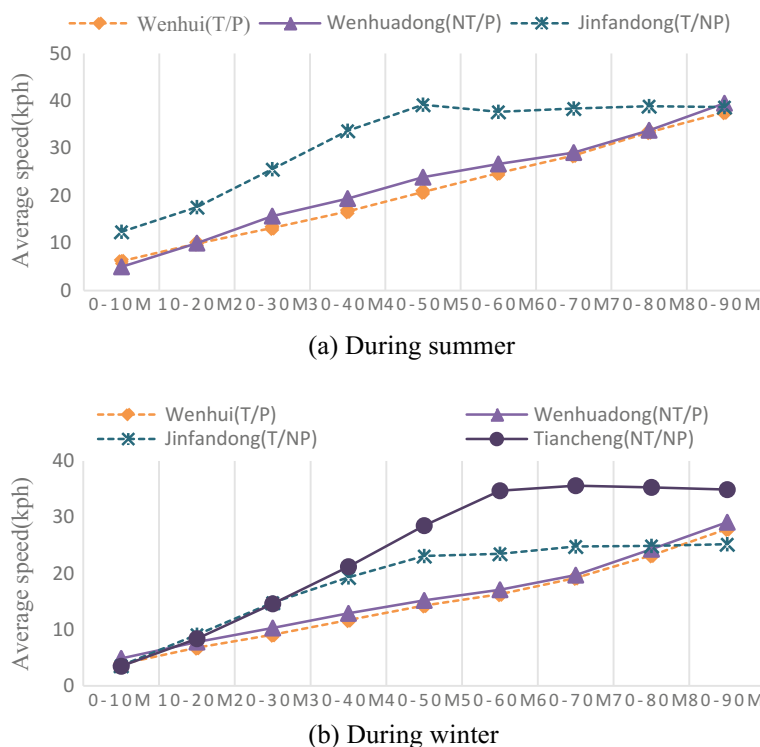


Fig. 1 The average speed of stopping vehicles

signal, the subject vehicle would begin to decelerate 30–50 m before the intersection. At the intersection with a countdown signal, the deceleration was low, and the speed change is not so abrupt, it usually starts between 50 and 90 m before the stop line. Comparing summer with winter, in winter, roads were more likely to be slippery, therefore, whether countdown signals exist or not, the deceleration was more stable, and the maximum speed is lower. In short, a sudden deceleration phenomenon occurs without a countdown signal, while a smoother deceleration can be found with a countdown signal.

3.2 Proportion of Stopping Vehicles

Figure 2a, b shows the percentage of drivers who decided to stop during the red stage, by their distance from the intersection at the start of the yellow phase, for winter and summer, respectively. In the summer, at intersections with countdown signal or E-Police (Wenhui, Wenhudong), most vehicles will decide to stop between 20 and 40 m before the stop line at the beginning of yellow stage. At the intersection without

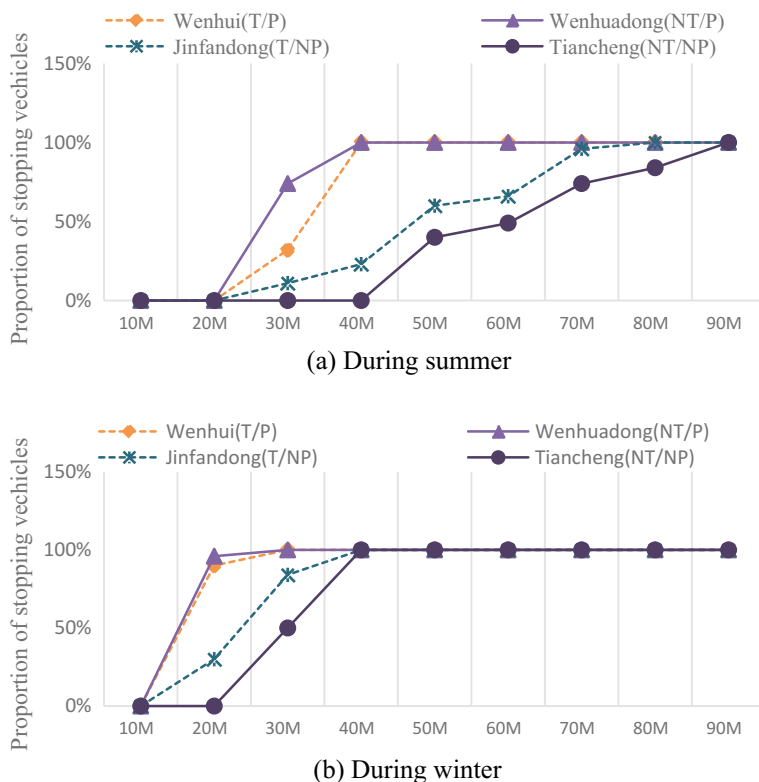


Fig. 2 The proportion of stopping vehicles

E-Police (Jinfandong) and at intersections where there are neither countdown signal nor E-Police (Tiancheng), most of the vehicles decide to stop between 40 and 70 m before the stop line at the beginning of the yellow stage. On the other hand, in winter, many vehicles make stop making decisions during 10–20 m before the stop line when the yellow stage is on. In addition, whether in the summer or in winter, at intersections where there are neither countdown signal nor E-Police (Tiancheng). The proportion of traversing vehicles is quite high even when the yellow phase started, which means that many vehicles have red light running behavior at this type of intersection.

4 Analytical Model

Through the above data analysis, we infer that the speed, the remaining countdown time, the distance from the stop line, and the existence of E-Police will affect the decision-driving at the countdown signalized intersections during the green stage. These factors are used as explanatory variables for the model construction. Since the

driver noticed a countdown signal and began to change driving behavior, this study also used the speed of the 80–90 m before the stop line and the remaining green time of the 80 before the stop line as explanatory variables. The used variables in the model are shown in Table 2. The subject vehicle included in the analysis must be the vehicle that actually made the decision. Therefore, in each signal cycle, the first vehicle stopped at the red stage and the last vehicle passed through the intersection just before are within the analysis range. Then, based on the collected data, we established a Logistic model that stopping or traversing decision-making during the two seasons of the green stage, and analyzed the influence of the countdown signal on driving behavior. After repeated attempts, Table 3 shows the best model for stopping or traversing decisions in summer and winter. The compliance rate of these two models is better, reaching more than 91%.

(1) The Logistic model for stopping or traversing decision in summer is

$$\log\left(\frac{p_1}{p_0}\right) = 2.287x_1 - 1.386x_4 - 0.233x_5 + 13.879 \quad (1)$$

where x_1 is the dependent variables of T-P, x_4 is the dependent variables of Speed_80–90 m, and x_5 is the dependent variables of Remaining Green Time_80 m. The Chi-square deviation of the model is 538.577, with a significant level of 0.001, indicated that the model is acceptable. From Eq. 1, it can be seen that in the summer, the stopping decision is related to the existence of both devices (T-P) compared

Table 2 Definitions of variables

	Variables	Definition
Dependent variables	Stopping or traversing decision	A dummy variable for stopping or traversing decision, if stopped = 1, otherwise = 0
Explanatory variables	Speed_80–90 m	Average speed at the 80–90 m section, which is located at 80–90 m from stop line (kph)
	Remaining green time_80 m	Remaining green time at 80 m before stop line
	T-P	A dummy variable for countdown timer and E-Police, if installed = 1, otherwise = 0
	T-NP	A dummy variable for countdown timer and no E-Police, if installed = 1, otherwise = 0
	NT-P	A dummy variable for E-Police and no countdown timer, if installed = 1, otherwise = 0

Table 3 Stopping or traversing decision model

Dependent variables	Stopping decision (P_1) Traversing decision (P_0)	Summer	Winter
		b-value	b-value
Explanatory variables	T-P (x_1)	2.287	8.265
	T-NP (x_2)	–	6.892
	NT-P (x_3)	–	6.302
	Speed_80–90 m (x_4)	–1.386	–2.189
	Remaining green time_80 m (x_5)	–0.233	–
	Constant	13.879	–
Number of samples		537	217
Hit ratio (%)		92.9	91.7
Chi-square deviation		538.577	192.351
Significance level		0.001	0.001

to traversing decisions. And the regression coefficient is positive, it can be concluded that both the countdown signal and the E-Police can increase the stopping tendency during the phase conversion process, and the existence of both devices has a greater impact on the stopping decision. The speed of 80–90 m phases before the stop line (Speed_80–90 m) and the remaining green time for 80 m before the stop line (Remaining Green Time_80 m) are all negative. This means that if the original speed in the 80–90 m phase is high and the countdown remaining time is large, the driver tends to go through the intersection.

(2) The Logistic model for stopping or traversing decision in winter is

$$\log\left(\frac{p_1}{p_0}\right) = 8.265x_1 + 6.892x_2 + 6.302x_3 - 2.189x_4 \quad (2)$$

where x_1 is the dependent variables of T-P, x_2 is the dependent variables of T-NP, x_3 is the dependent variables of NT-P, and x_4 is the dependent variables of Speed_80–90 m. The Chi-square deviation of the model is 192.351, with a significant level of 0.001, indicated that the model fits well. From Eq. 2, it can be seen that in the winter, compared with the traversing decision, the two devices (T-P), the countdown signal alone (T-NP), and the E-Police alone (NT-P) all have a greater impact on the drivers' stopping decision. And the regression coefficients for three patterns are positive, indicating that installing either of these two devices increases the driver's tendency to stop during the phase conversion process in winter. The regression coefficient for the speed of 80–90 m before the stop line (Speed_80–90 m) is negative. This means that if the original speed is higher at 80–90 m before the stop line, the driver tends to go through the intersection.

5 Conclusions

Through the analysis of the survey data of summer and winter stopping and traversing decisions from the green to the yellow at the four intersections provided by the Liaoning Provincial Transportation Bureau, China, the observation range of 0–90 m before the stop line was used as the observation scope. The driving characteristics of the vehicle were first analyzed through survey data, and then the influence of factors such as countdown signal and the electronic police (E-Police) on stopping or traversing decisions were also estimated through establishing models. In conclusion, through data analysis and modeling analysis, the driver's stopping decision depends on the coexistence of the countdown signal and E-Police in summer. The speed of the 80–90 m before the stop line is high and the countdown remaining time is also long, the driver tends to traverse the intersection if both countdown lights and E-Police exist in summer. In winter, the driver's speed at the intersection is not high, and installing either of the two devices can greatly increase the driver's stopping tendency during the transition of the signal.

Acknowledgements and Statements The work was supported by the Social Science Fund of Liaoning Province of China (L18BGL026) and Basic Scientific Research Project in the University of Liaoning Province of China (Z2218043).

The investigation of this paper was conducted with the consent of Shenyang Transportation Bureau (Shenyang, China), and participants (drivers) were informed of the use of relevant information in this investigation.

References

1. Peng, H.H., Zhang, C.B., Shi, Y.H., et al.: Research on influence of red countdown on vehicle departure time at intersection. *J. Wuhan Univ. Technol.* **40**(3), 540–549 (2016). (in Chinese)
2. Islam, M.R., Wyman, A.A., Hurwitz, D.S.: Safer driver responses at intersections with green signal countdown timers. *Transp. Res. Part F: Traffic Psychol. Behav.* **51**, 1–13 (2017)
3. Yang, X.B., Wang, K.F.: Impact analysis and prospect of traffic signal countdown display. *Compr. Transp.* **37**(10), 61–67 (2015). (in Chinese)
4. Cao, Y., Yang, Z.Z., Zuo, Z.Y.: Influence of countdown signal of green light on driving behavior. *China Saf. Sci. J.* **25**(2), 77–82 (2015). (in Chinese)
5. Girish, V.V., Asha, A., Vanajakshi, L.: Effect of countdown timers on red light violations under heterogeneous traffic conditions. *J. Inst. Eng. (India): Civ. Eng. Div.* **92**, 37–43 (2011)
6. Pan, F.Q., Liu, T., Zhang, L.X., et al.: Modeling of driving behaviors at countdown signalized intersections considering the value of car. *J. Transp. Syst. Eng. Inf. Technol.* **16**(2), 64–69 (2016). (in Chinese)

The Performance Evaluation and Improvement of Urban Taxi Firms Using Data Envelopment Analysis and Benchmarking Approach



Zhu Bai and Shuai Bian

Abstract This paper aims to evaluate the performance of taxi firms and discuss its improvement for those underperforming taxi firms. The performance measured from the operational efficiency and service effectiveness perspectives has a significant impact on the input–output balance in the taxi market and economic benefits and social welfare. Based on this, we propose Data Envelopment Analysis (DEA) method to achieve the measurement of performance and adopt the benchmarking approach to complete its improvement. In doing so, this paper first discusses how to make the appropriate selections in terms of decision-making units, DEA models, and input–output indicators. The input-oriented DEA models are then applied to assess different types of performance indicators. And moreover, to provide the practical improvement recommendations, the targets assigned by the benchmarking approach and the gaps between targets and current values are also presented and the future effort directions are thus pointed out. Our results suggest that assessing and improving the performance of taxi firms are conducive to make and carry out management strategies in the taxi market.

Keywords Data · Envelopment · Analysis · Bench marking · Taxi performance evaluation

1 Introduction

It is well known that taxi is an important component of urban public transport systems. This trip mode has become more attractive not only because of the comfortable and speedy characteristic, but mainly because of its door-to-door personalized service. However, many researchers have found taxi's air pollution emissions are higher than those of other trip modes [1, 2]. Especially in China, there is a serious problem associated with taxi's air pollution emissions in several cities [3–5]. This is likely

Z. Bai (✉) · S. Bian

School of Transportation Engineering, Shenyang Jianzhu University, Shenyang, China
e-mail: 84005291@qq.com

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_28

301

to be attributed to higher vacant ratio and inefficient management and operation in the taxi market. Under the circumstances, much research interest about how to improve the operational performance of the taxi market has been simulated. The majority of existing research on the public transport systems focuses on evaluating the performance of public transit [6–11]. Little attention has been given to research on the performance of the taxi market. Consequently, it is necessary to reveal practical situations and further put forward constructive proposals for the taxi market.

In this paper, our idea is to integrate Data Envelopment Analysis (DEA) and benchmarking method, considering both evaluation and improvement of performance at the level of a taxi firm. DEA, which is a nonparametric frontier approach, is applied to determine the relative performance of entities by measuring multiple attributes such as efficiency and effectiveness. The benchmarking method is a quantitative management technique that provides powerful support for achieving the advancement of performance. These two analytical techniques are devoted to better understanding of the situations of the taxi market and providing proposals on planning strategies for taxi policymakers.

In this paper, a practical approach which not only measures system performance but also provides powerful support for identifying the source of inefficiency will be developed in terms of taxi firms. The remaining of this paper is structured as follows. Section 1 presents previous researches on how to measure performance for public transport systems. Section 2 specifies the DEA models and introduces the benchmarking method for this study. Section 3 identifies input–output indicators for this study and analyzes the statistic characteristic of data used. Section 4 illustrates how to apply the methodology and discusses its corresponding results. Section 5 summarizes several conclusions.

2 Methodology

To assess the performance of a taxi firm from both service provider and consumer perspectives, we measure operational efficiency and service effectiveness. Operational efficiency is similar to technical efficiency that is mainly used to identify physical productivity. Its score describes how to obtain maximal output from a given set of inputs in the service production process for taxi firms in our case [12]. The score of service effectiveness indicates the optimal benefits received from a given combination of inputs in the service consumption process for taxi firms in our case. DEA is a mathematical method for calculating the relative performance of decision-making units (DMUs). DMUs consist of a homogeneous set of observed data.

In general, a DEA model is established with either an input or output orientation under the underlying assumptions of constant returns to scale (CRS) or variable returns to scale (VRS) [13]. In our case, the input-oriented DEA models are adopted, as their objective is to seek the minimal ratio of contraction for all inputs when keeping all outputs unchanged [14]. More importantly, compared with those outputs of DEA models, taxi firms can exert a greater impact on their inputs. This is mainly because

outputs of taxi firms are substantially influenced by road network conditions and different regulations of the taxi market. In terms of returns to scale, the assumption of CRS is that all observed DMUs operate at the optimal scale while the one of VRS is that efficiency may increase or decrease with a change of size in input or output. The CCR model (being established by A. Charnes, W. W. Cooper, and E. Rhodes) explicitly assumes that DMU is operating at its optimal scale, namely assuming CRS, while BCC model (being established by Banker, Charnes, and Cooper) is used for analyzing VRS cases. In addition, the ratio of CCR score to BCC score is called as the scale efficiency score (SE). Thus, we calculate both CCR and BCC models for the purpose of better reflecting the performance of taxi firms.

3 Data

Harbin, the capital city of Heilongjiang Province, is located in the northeast of China. In this paper, it analyzed the total data of 20 taxi firms obtained from Harbin taxi management department over the full 12 months from May 2012 to April 2013. The data was involved in total number of taxis, operating revenues, travel kilometers of taxis, and occupied kilometers of taxis, etc. Taxi firms are mainly consisted of Beihuan Tongyong (BHTY), Beixin Tongyong (BXTY), Chaoyun Tongyong (CYTY), Dazhong Tongyong (DZTY), Dazhong Yaxun (DZYX), Feiyu Tongyong (FYTY), Fengxiang Tongyong (FXTY), Fengxiang Yaxun (FXYX), Huaqiao Tongyong (HQTY), Jihua Tongyong (JHTY), Longjiang Lvyou Tongyong (LLTY), Longyun Tongyong (LYTY), Longyun Yaxunjian (LYYXJ), Lvyou Qiche Yaxun (LQYX), Rongshida Tongyong (RSDTY), Shiyun Tongyong (SYTY), Taikesi Yaxun (TKSYX), Tiane Tongyong (TETY), Tiane Yaxun (TEYX), and Zhongshun Tongyong (ZSTY). According to the principle of DEA method, each taxi firm is treated as a decision-making unit in each of the models.

There are many considerations in specifying input–output factors for evaluating the performance of public transport in previous DEA studies. In general, labor, capital, and energy are used as inputs while vehicle-km, seat-km, or passenger-km as outputs [15]. Generally, capital and energy in our case are considered as input factors. More specifically, we choose the total number of taxis (TN) as the measure of capital and daily average fuel consumption per taxi (TF) as the measure of energy.

In general, vehicle-miles that are viewed as the output of produced type is usually selected to assess operational efficiency or technical efficiency, while passenger-miles and ridership that are the output of consumed type are usually adopted to measure service effectiveness [16]. More specifically, vehicle-miles that is defined as daily average travel kilometers per taxi (km) and trip distance, that is, defined as daily average one-way travel distance per taxi (km) are applied to act as output indicators for measuring operational efficiency, while passenger-miles that are defined as daily average occupied distance per taxi (km) and ridership, which is defined as daily average number of passenger traveled per taxi are adopted to act as output indicators for identifying service effectiveness.

4 Results

To evaluate the performance of the taxi firms using different models, these models were first solved by Lingo 11.0, which is a software used for optimizing the operational problems. The results of DEA models are shown in Table 1.

Table 1 shows that the relative scores of performance are determined by using DEA models for all of the taxi firms. First, LYTY and TETY are identified as achieving relatively overall technical efficiencies. They are the best-performing firms and thus act as the benchmarks. The overall technical efficiencies of other 18 taxi firms are relatively inefficient, of which 4 firms are scale inefficient but purely technically efficient and 14 firms are both and scale inefficient. Second, in terms of service effectiveness, a total of three taxi firms are evaluated as achieving relatively overall service effectiveness, including HQTY, LYXJ, and RSDTY. These three taxi firms are thus used as the benchmarks for adjusting service effectiveness. Other 17 taxi firms are done as relatively overall service ineffectiveness, of which 3 are scale

Table 1 The evaluation results of performance based on CCR and BCC models

DMU _i	Operational efficiency			Service effectiveness		
	CCR	BCC	SE	CCR	BCC	SE
BHTY	0.76198	0.95502	0.79786	0.96174	0.98096	0.98040
BXTY	0.82059	0.91021	0.90153	0.96690	1	0.96690
CYTY	0.84254	0.87892	0.95860	0.99232	1	0.99232
DZTY	0.75004	0.95649	0.78415	0.89224	0.95866	0.93071
DZYX	0.71310	0.89200	0.79943	0.94501	0.95635	0.98814
FYTY	0.70918	0.87002	0.81513	0.99231	0.99319	0.99911
FXTY	0.71982	0.91562	0.78615	0.98951	0.99179	0.99770
FXYX	0.78107	0.83731	0.93283	0.80784	0.80897	0.99860
HQTY	0.95755	1	0.95755	1	1	1
JHTY	0.77371	0.95561	0.80965	0.97776	0.98816	0.98947
LLTY	0.99047	1	0.99047	0.92933	0.93280	0.99628
LYTY	1	1	1	0.90930	0.93520	0.97230
LYXJ	0.71345	0.86586	0.82397	1	1	1
LQYX	0.67144	0.84062	0.79874	0.92852	0.93248	0.99575
RSDTY	0.75700	0.90834	0.83338	1	1	1
SYTY	0.68045	0.81772	0.83213	0.87986	0.88206	0.99750
TKSYX	0.82080	1	0.82080	0.97230	1	0.97230
TETY	1	1	1	0.95063	0.96390	0.98623
TEYX	0.79817	0.93699	0.85184	0.93222	0.93746	0.99441
ZSTY	0.70349	1	0.70349	0.65602	0.65738	0.99793

ineffective but purely technically effective and 14 are both purely technically and scale ineffective.

To further explore some valuable improvements for those inefficient taxi firms, we introduce the benchmarking approach based on the results of the dual CCR model. One hand, LYTY and TETY are the benchmarks obtained in the evaluation process of operational efficiency and their dual weights make up the reference sets, which are displayed in the second column of Table 3. On the other hand, HQTY, LYYXJ, and RSDTY are the benchmarks and their dual weights make up the reference sets, which are presented in the third column of Table 2.

First, Table 2 shows that the reference sets for each inefficient taxi firm solely consist of dual weights of LYTY and TETY in the operational efficiency model, while the reference sets in the service effectiveness model are comprised of dual weights of HQTY, LYYXJ, and RSDTY.

Second, the results also reveal that in these two types of models, not all of the benchmarks are involved in the advancement of the performance, as several reference

Table 2 The reference sets for improving operational efficiency and service effectiveness

DMU _i	Reference sets for operational efficiency	Reference sets for service effectiveness
	{LYTY; TETY}	{HQTY; LYYXJ; RSDTY}
BHTY	{0; 0.7309}	{0; 0; 0.8591}
BXTY	{0.7499; 0}	{0.9431; 0; 0.1132}
CYTY	{0.0102; 0.9300}	{0; 0; 1.0317}
DZTY	{0.1759; 0.5263}	{0; 0; 0.7859}
DZYX	{0.0721; 0.6556}	{0; 0; 0.9018}
FYTY	{0.1731; 0.5765}	{0; 0; 0.9860}
FXTY	{0.6413; 0.0852}	{0; 0; 0.9637}
FXYX	{0; 1.0154}	{0; 0; 0.9779}
HQTY	{0.7814; 0}	{1; 0; 0}
JHTY	{0.5331; 0.2131}	{0; 0; 0.9037}
LLTY	{0.9813; 0.0592}	{0; 0; 0.9441}
LYTY	{1; 0}	{0; 0; 0.8814}
LYYXJ	{0; 0.7631}	{0; 1; 0}
LQYX	{0; 0.7320}	{0; 0; 0.9427}
RSDTY	{0.7614; 0}	{0; 0; 1}
SYTY	{0.5901; 0.1842}	{0; 0; 0.9613}
TKSYX	{0.5480; 0.1722}	{0; 0; 0.8190}
TETY	{0; 1}	{0; 0; 0.8852}
TEYX	{0.7339; 0}	{0.5141; 0; 0.4351}
ZSTY	{0.1410; 0.9673}	{0; 0; 0.9675}

sets with zero dual weights arise in the models. More specifically, LYTY acts as a benchmark firm for other 14 inefficient firms while TETY also for 14 of inefficient firms within this group of operational efficiency. HQTY acts as a benchmark firm for only 2 ineffective firms, such as BXTY and TEYX, while RSDTY does for 14 of ineffective firms within this group of service effectiveness.

Finally, taking DZTY as an example, the dual weight of TETY is much greater than that of LYTY, showing that TETY makes much larger contribution to advancing the operational efficiency of DZTY. Within the group of service effectiveness, taking BXTY as an example, the dual weight of HQTY is the largest, and the one of LYXXJ is the smallest, and the one of RSDTY is centered in this reference set, which implies that the largest contributor is HQTY and the smallest one is LYXXJ in improving the service effectiveness of BXTY.

Table 3 gives the improvements and effect examinations of input indicators.

From Table 3, those wasted input indicators are presented in the fourth and seventh columns, respectively. Specifically, the maximal gap between target TN and current TN is 521 vehicles while the minimal one is 2 vehicles. The ratio of firms whose improvement gap is larger than 100 vehicles is 17%, and that the ratio of ones whose adjustment gap is less than 50 vehicles is 39%, and that the ratio of others is 44%.

Table 3 The improvements and effect examinations of input indicators

DMU _{<i>i</i>}	Target TN	TN	Gap	Target TF	TF	Gap	New score
BHTY	154	218	64	19.54	25.64	6.10	1
BXTY	130	159	29	20.86	29.68	8.82	1
CYTY	198	235	37	25.14	29.84	4.70	0.9992
DZTY	142	199	57	18.96	25.28	6.32	0.99997
DZYX	151	343	192	19.53	27.39	7.86	1
FYTY	152	214	62	20.23	28.52	8.29	0.99827
FXTY	130	180	50	20.12	27.95	7.83	0.99749
FXYX	214	280	66	27.14	34.75	7.61	1
HQTY	136	142	6	21.74	26.98	5.24	0.99862
JHTY	138	178	40	20.53	26.53	6.00	0.99823
LLTY	183	185	2	28.88	29.16	0.28	0.99921
LYTY	174	174	521	27.82	27.82	8.19	1
LYXXJ	161	682	284	20.40	28.59	9.57	0.99987
LQYX	154	438	43	19.57	29.14	7.52	1
RSDTY	132	175	66	21.18	28.70	10.02	0.99998
SYTY	142	208	56	21.34	31.36	4.33	0.99753
TKSYX	132	188	32	19.85	24.18	7.86	0.99982
TETY	211	211	96	26.73	26.73	12.55	1
TEYX	128	160	64	20.42	28.28	6.10	0.99784
ZSTY	229	325	29	29.78	42.33	8.82	0.99826

The maximal gap between target TF and current TF is 12.55 L while the minimal one is 0.28 L. Those firms whose adjustment gap is much larger than 10 L account for 11%, and that the ones whose adjustment gap is less than 5 L account for 17%, and that the others do for 72%.

5 Conclusions

In this study, the performance of the taxi firms in Harbin was measured by introducing DEA method and improved by adopting the benchmarking approach. In doing so, three aspects were taken into account. First, operational efficiency was proposed to reveal the productivity of a taxi firm while service effectiveness was used to reflect the consumption situations of service provided by the taxi firm. Second, the input-oriented CCR model was applied to measure the overall performance score while the input-oriented BCC model was presented to determine the pure performance score. The full picture of the performance could be then identified, considering not only the attributes of model output but also the characteristics of performance. Based on this, those appropriate benchmarks were also pointed out. Third, the benchmarking approach was further conducted to explore the adjustment degree of input indicators for those underperforming firms.

The aforementioned analysis, which considers both evaluation and improvement of performance simultaneously, provides deeper insights and practical recommendation for taxi policymakers and operators. More specifically, using DEA models, the performance of each firm is measured from both service provider and service consumer perspectives. By inspecting the operational efficiency and service effectiveness, we are able to obtain best performers as well as inefficient or ineffective firms. The findings contribute to clarifying current operation situations and determining great challenges which need to be completed in Harbin taxi market. Moreover, each reference set indicates the extent to the contribution of different benchmarks in adjusting the inefficient or ineffective firms and the targets based on this reference set enable taxi operators to recognize the gap with the benchmarks. In addition, examining the effect of the improvement approach allows us to present more feasible operational policies for taxi operators.

References

1. Stead, D.: Relationships between transport emissions and travel patterns in Britain. *Transp. Policy* **6**, 247–258 (1999)
2. Reddy, A.K.N., Anand, Y.P., D'Sa, A.: Energy for a sustainable road/rail transport system in India. *Energy Sustain. Dev.* **4**(1), 29–44 (2000)
3. Sun, C.W., Luo, Y., Li, J.L.: Urban traffic infrastructure investment and air pollution: evidence from the 83 cities in China. *J. Clean. Prod.* **172** (2018)

4. Oliver, H.H., Li, M., Qian, G., Qin, K., Zhang, J., He, K., Liu, H., Gallagher, K.S., Davis, N., Lents, J.: In-use vehicle emissions in China: Tianjin study (2008)
5. Wang, H., Fu, L., Zhou, Y., Du, X., Ge, W.: Trends in vehicular emissions in China's mega cities from 1995 to 2005. *Environ. Pollut.* **158**, 394–400 (2010)
6. Qiu, F., Shen, J.X., Zhang, X.C., An, C.C.: Demi-flexible operating policies to promote the performance of public transit in low-demand areas. *Transp. Res. Part A* **80** (2015)
7. Triantis, K.P.: Engineering applications of DEA. In: Cooper, W.W., Seifor, L.M., Zhu, J. (eds.) *Handbook on Data Envelopment Analysis*, pp. 401–441. Kluwer Academic Publications (2004)
8. Sheth, C., Triantis, K., Teodorovic, D.: Performance evaluation of bus routes: a provider and passenger perspective. *Transp. Res. Part E* **43**, 453–478 (2007)
9. De Borger, B., Kerstens, K.: The performance of bus-transit operators. In: Hensher, D.A., Button, K.J. (eds.) *Handbook of Transport Modelling*, 2nd edn, pp. 693–714. Elsevier, Amsterdam (2008)
10. Lao, Y., Liu, L.: Performance evaluation of bus lines with data envelopment analysis and geographic information systems. *Comput. Environ. Urban Syst.* **33**, 247–255 (2009)
11. Chiu, Y.-H., Huang, C.-W., Ma, C.-M.: Assessment of China transit and economic efficiencies in a modified value-chains DEA model. *Eur. J. Oper. Res.* **209**, 95–103 (2011)
12. James, O., Svein, B.: A meta-analysis of DEA and SFA studies of the technical efficiency of seaports: a comparison of fixed and random-effects regression models. *Transp. Res. Part A* **46**, 1574–1585 (2012)
13. Rico, M., Hensher, D.A.: The impact of strategic management and fleet planning on airline efficiency-A random effects Tobit model based on DEA efficiency scores. *Transp. Res. Part A* **45**, 686–695 (2011)
14. Lin, E.T.: Route-based performance evaluation of Taiwanese domestic airlines using data envelopment analysis: a comment. *Transp. Res. Part E* **44**, 894–899 (2008)
15. De Borger, B., Kerstens, K., Costa, A.: Public transit performance: what does one learn from frontier studies. *Transp. Rev.* **22**(1), 1–38 (2002)
16. Fielding, G.J.: *Managing Public Transit Strategically: A Comprehensive Approach to Strengthen in Service and Monitoring Performance*. Jossey-Bass Publishers, San Francisco (1987)

Research on the DV-Hop Location Algorithm Based on the Particle Swarm Optimization for the Automatic Driving Vehicle



Pei Huang, Xinjian Xiang and Bingqiang Huang

Abstract Localization is the foundation of automatic driving and it has always been a topic of research hotspot and difficult to deal with. The DV-Hop algorithm is most widely used in node localization research, and it will be used together with particle swarm optimization algorithm. In this paper, an NJPDH algorithm is proposed to compensate for the inferior accuracy of DV-Hop localization algorithm. Based on the DV-Hop location algorithm, the algorithm is added to the weight of each beacon node, the average distance is weighted, the particle swarm optimization is optimized from two aspects of the inertia weight and the active factor to avoid the particles trapped in the local optimization, and then get the location of the unknown nodes better. The simulation results show that under the same hardware conditions, compared with the particle swarm based PSO-DV-Hop algorithm, it can effectively reduce the impact of the jump distance, increase the coverage of the nodes, improve the positioning accuracy and robustness of the positioning process, and have better applicability.

Keywords Automatic driving vehicle location · DV-Hop algorithm · Particle swarm optimization · Hop distance and inertia weighting

1 Introduction

Automatic driving vehicle is the high-tech combination product of the automotive industry, the Internet of things, high-performance computing, and AI [1]. It has many advantages such as greatly improving the efficiency and safety of traffic system, reducing the cost of human travel, and is the main direction of intelligent development

P. Huang

School of Mechanical and Energy Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China

X. Xiang (✉) · B. Huang

School of Automation and Electrical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China

e-mail: 2263934848@qq.com

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_29

of the global automobile and traffic field, and has become the strategic commanding point of all countries. Automatic driving is based on precise positioning technology, so node localization technology is the primary support technology for autonomous driving vehicle [2]. DV-Hop location algorithm is widely used in large-scale location because of low hardware requirements and uncomplicated calculation, but the positioning accuracy in irregular topology is not high. It uses hop count and average hop distance to estimate the distance between unknown nodes and beacon nodes, and then uses mathematical equations to locate unknown nodes. There are two ways to improve it: the first method is to use the jump distance weighting method to improve the average jump distance accuracy; the second is to optimize the positioning results with the group intelligence algorithm [3]. In these two ideas, many scholars have proposed a variety of optimization algorithms, such as in [4], it introduced the jump number threshold, selected the optimal jump distance instead of the average jump distance to locate the unknown node, thus improving the positioning accuracy, but the improvement is not obvious. In Guo and Dong [5], it used particle swarm optimization to replace the simple equation of the third phase of DV-Hop algorithm to optimize the results, but it is easy to fall into the “precocious” predicament, and the optimization effect is not ideal. In Miao et al. [6], it proposed a PSO-DV-Hop method, which is based on particle swarm optimization (PSO) and uses the learning factor perturbation strategy to help PSO to jump out of the “precocious” dilemma. Although its positioning accuracy is better than that of the first two documents and the computational complexity is increased, the positioning effect is not very good.

Based on the above analysis, an NJPDH algorithm is proposed in this paper. The algorithm weights the beacon nodes and jumps, and corrects the positioning results by using the improved particle swarm optimization algorithm. The inertia weight is improved and the active factor is added to make the particle swarm algorithm jump out of the “precocious” predicament. The simulation results show that the NJPDH algorithm has better robustness and better location performance compared with the PSO-DV-Hop algorithm in [6].

2 DV-Hop Algorithm Induction

DV-Hop location algorithm is the most widely used algorithm in the range-free location algorithm. It collects hops and hops information of multiple nodes, and uses three edge ranging or maximum likelihood to locate unknown nodes. Its positioning principle and process can be divided into the following three steps [7]:

- (1) The beacon nodes share the information of their own. The hop count is one of the information. After receiving the information sent by other nodes, the receiving node will save the minimum hop number to the corresponding node, and the information of the larger hop number of the same beacon node will not be recorded. Then, the number of hops corresponding to each beacon node is added to 1, and the hops information is continued to be shared. In this way,

the minimum hop values between every two beacon nodes are obtained and preserved for these two nodes.

- (2) After each node has information from other nodes, the average distance per hop can be calculated using Eq. (1).

$$D_i = \frac{\sum_{j \neq i} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{\sum_{j \neq i} h_{ij}} \quad (1)$$

The coordinates of beacon nodes i and j are represented by (x_i, y_i) and (x_j, y_j) . The number of hops between beacon node i and j ($i \neq j$) is expressed by h_{ij} . Next, the beacon nodes begin to share their average hop distance within the communication range.

- (3) The number of hops and the corresponding hops of other beacon nodes in the unknown node has been obtained in the shared information, and the location information is estimated by using the three edge distance method or maximum likelihood method.

Suppose that, there are m ($m > 3$) known as nodes, the coordinates are (x_m, y_m) , and the distance between the known nodes m and the unknown nodes is r_m . As a result, it can be obtained as

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 = r_1^2 \\ \vdots \\ (x_m - x)^2 + (y_m - y)^2 = r_m^2 \end{cases} \quad (2)$$

In the form, (x, y) is the coordinate of the unknown node. The first $M - 1$ equation is used to reduce the first m equation as

$$\begin{cases} x_1^2 - x_m^2 - 2x(x_1 - x_m) + y_1^2 - y_m^2 - 2y(y_1 - y_m) = r_1^2 - r_m^2 \\ \vdots \\ x_{m-1}^2 - x_m^2 - 2x(x_{m-1} - x_m) + y_{m-1}^2 - y_m^2 - 2y(y_{m-1} - y_m) = r_{m-1}^2 - r_m^2 \end{cases} \quad (3)$$

The linear equation is used to express the $M - 1$ equations as follows:

$$A = 2 \times \begin{bmatrix} x_1 - x_m & y_1 - y_m \\ \vdots & \vdots \\ x_{m-1} - x_m & y_{m-1} - y_m \end{bmatrix}$$

$$d = \begin{bmatrix} r_m^2 - r_1^2 + x_m^2 - x_1^2 + y_m^2 - y_1^2 \\ \vdots \\ r_m^2 - r_{m-1}^2 + x_m^2 - x_{m-1}^2 + y_m^2 - y_{m-1}^2 \end{bmatrix}, \quad x = \begin{bmatrix} x \\ y \end{bmatrix}$$

The least square method is used to solve the problem. We get $\hat{X} = (A^T A)^{-1} A^T d$.

From the above equation, we can see that the coordinates of the unknown nodes are intersection of some circles which take $(x_1, y_1), \dots, (x_m, y_m)$ as the center and take $d_1, \dots, d_m$ as the radius, respectively. But the m circle in the actual situation is likely to intersect not only a point, but a small area with many points, which is to be improved for the DV-Hop algorithm. There is a certain degree of error in the location result obtained. In this paper, the NJPDH algorithm is used to solve the localization problem by solving the minimum location error. The fitness function for positioning error is as follows:

$$fitness(x, y) = \frac{1}{ART} \sum_{t=k}^T \sum_{k=1}^A [\sqrt{(x_k - x_{tk}^*)^2 + (y_k - y_{tk}^*)^2}] \quad (4)$$

In the form, the actual coordinates of the unknown node k are represented by (x_k, y_k) ; the coordinates of the unknown node k for the t estimation are expressed by (x_{tk}^*, y_{tk}^*) ; R is the communication radius; T is the number of iterations; A is the total amount of the unknown node.

3 NJPDH Algorithm Based on Adaptive Particle Swarm Optimization

In the second stage of the DV-Hop algorithm, the traditional idea is not refined when calculating the distance, and the average jump distance is used, and the spatial relationship with other reference nodes is ignored, and this neglect will lead to large positioning errors [8]. In the third stage, the traditional way of thinking uses simple mathematical equations to solve the coordinates, and the number of selected nodes is few. At this time, if the average distance of the selected node is larger, the degree of the positioning result will be greatly affected [9].

In this paper, the weight of each beacon node is added to the NJPDH algorithm, the average distance is weighted, the inertia weight is improved, and the active factor is added, which can make the particle swarm algorithm get rid of the “precocious” predicament and improve the location performance of the algorithm.

3.1 Beacon Node Improvement

This paper improves the average hop distance of the global beacon nodes by adding weights to each beacon node. The weight is determined by the ratio of the error of each beacon node in the total error. So first, we get the distance error between every two beacon nodes and express it by coefficient θ_{ij} . The number of hops between beacon nodes i and j is represented by h_{ij} , and the communication radius is represented by

R . In ideal state, R is hopping distance, and the ideal distance is the product of R and h_{ij} . The actual distance between beacon node i and beacon node j is expressed by d_{ij} , so θ_{ij} can be obtained as follows.

$$\theta_{ij} = |R \times h_{ij} - d_{ij}| \quad (5)$$

The weighted coefficient ω_{ij} of the weighted calculation is

$$\omega_{ij} = \frac{\theta_{ij}}{\sum_{j \neq i}^m \theta_{ij}} \quad (6)$$

The average hop size of the modified beacon node is

$$\bar{D}_i = \sum_{j \neq i}^n \omega_{ij} \times D_{ij} \quad (7)$$

3.2 Beacon Node Improvements

The number of hops used in the traditional algorithm is only the main node, but the hops information used by the improved algorithm comes from each beacon node, which greatly reduces the impact of the hop number on the location results. The estimated distance between the main node i and the unknown node K is β_{ik} .

$$\beta_{ik} = \bar{D}_i \times h_{ik} \quad (8)$$

Based on the above idea, the unknown node hop coefficient is defined as

$$\alpha_k = \frac{h_{jk}}{h_{ik}} \quad (9)$$

where α_k is the global deviation of the unknown node k .

The local deviation coefficient of the unknown node k is

$$\lambda_k = \frac{h_{ij}}{h_{ik} + h_{jk}} \quad (10)$$

So after improvement, the hop distance between unknown node k and beacon node j is

$$\bar{D}_{jk} = \frac{\alpha_k}{\alpha_k + \lambda_k} \bar{D}_i + \frac{\lambda_k}{\alpha_k + \lambda_k} \bar{D}_{ij} \quad (11)$$

The estimated distance between the unknown node k and the beacon node j is

$$\beta_{jk} = \bar{D}_{jk} \times h_{jk} \quad (12)$$

4 Introduction of Particle Swarm Optimization Algorithm

The PSO algorithm mimics the group foraging of natural creatures, and each member of the group constantly changes its foraging way by mastering the latest information of itself and the group, so as to find the food location faster and more accurately. PSO algorithm has been widely applied to various practical optimization problems due to its advantages of less adjustment parameters, faster convergence speed, and higher search accuracy [10].

PSO algorithm finds the best solution by iteration. It first produces some particles in a certain range, representing random solutions. The mathematical model of PSO algorithm can be expressed as in n -dimensional space, a population composed of m particles is expressed as $X = (x_1, x_2, \dots, x_m)$, mimic the biological species of the natural world. The position of the i particle is $X = (x_1, x_2, \dots, x_m)$, representing the location of each member of a population. The best location of the i particle is $pbest_i = (pbest_{i1}, pbest_{i2}, \dots, pbest_{in})$, simulating the best position of the i group member, and the global optimal position is expressed as $P_{gbest} = (p_{gbest1}, p_{gbest2}, \dots, p_{gbestn})$, to simulate the optimal location of the entire population. The velocity of the i particle is $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$, simulate the speed of the i member. In each iteration, the position and speed of each particle is adjusted with the two extremums ($pbest$, $gbest$), until the number of iterations has reached the maximum or the particle has gathered around the optimal location. The equation for adjusting speed and position is as follows:

Displayed equations are centered and set on a separate line.

$$v_{i+1} = v_i + c_1 \times r_1 \times (pbest_i - x_i) + c_2 \times r_2 \times (gbest - x_i) \quad (13)$$

$$x_{i+1} = x_i + v_{i+1} \quad (14)$$

where $i = 1, 2, \dots, M$, M is the total number of particles. v_i is the velocity of particle i , the maximum is v_{max} , if $v_i > v_{max}$, then v_{max} is assigned to v_i , r_1 and r_2 are both (0,1) random numbers, x_i is the current position of particles, c_1 and c_2 are learning factors, and generally $c_1 = c_2 = 2$.

In the process of optimizing the PSO algorithm, the particle is easily converged to the optimal position of the global and its own history, and it is difficult to jump out of the local optimal, so it will fall into "precocious", and the precision of the algorithm search is not high in the end, and the accuracy is not high [8]. In this paper, we improve the inertia weight and increase the active factors, respectively, so as to expand the search scope and improve the search accuracy of the algorithm.

4.1 Improvement of Inertia Weight

In the early stage of iteration, the larger inertia weight will make the algorithm skip the optimal position and lead to low positioning accuracy. In the later period of the iteration, a small inertia weight will cause the particle to fall into the optimization dilemma and affect the positioning accuracy as well. For this reason, we add linear decrement factor ω , and the fluctuation amplitude decreases with the number of iterations. It has the maximum value ω_{max} and the minimum value ω_{min} . The calculation equation of ω is as follows:

$$\omega = \omega_{max} - (\omega_{max} - \omega_{min})t/T + (2 - 1 \times e^{(-t/T)}) \times rand \quad (15)$$

where $rand$ is the number generated in random (0, 1) range; T is the largest iteration number; t is current iterations. After many practices, it can be concluded that ω_{max} takes 0.9 and ω_{min} takes 0.4, and the positioning effect is better.

4.2 Active Factor Center Learning Particle Swarm Optimization Algorithm

This paper introduces a learning strategy with active factors. In order to make the particle jump out of the “precocious” predicament and add the early maturity identification mark flagmt, if it is in the “precocious” state, the learning strategy is converted to increase the diversity of the particles so as to jump out of the plight of the local optimal. The speed updating equation is as follows:

$$v_i^{k+i} = \omega \cdot v_i^k + c_1 \cdot r_1(pbest_i^k - x_i^k) + c_2 \cdot r_2(gbest_i^k - x_i^k) + c_3 \cdot r_3(c_{end}^k - x_i^k) \quad (16)$$

where $c_3 = 2$ is the definition, c_{end} is an active factor. The other parameters are explained as above.

$$c_{end} = \frac{r_1 \times pbest_i + r_2 \times gbest_i}{2} \quad (17)$$

4.3 Algorithm Procedure

The algorithm flow in this paper is as follows:

- (1) At initialization stage, several nodes are randomly set in the target area, and the beacon nodes send their own information to the network.

- (2) Calculating the average hop distance weighted by the beacon node according to the Eq. (5)–(7).
- (3) Using Eqs. (8)–(12) to correct the hop distance between each unknown node and each beacon node.
- (4) Initialization of the algorithm: the size of the particle swarm is M , the maximum number of iterations is T_{max} , the maximum value M_{max} and the minimum value M_{min} , the maximum velocity v_{max} of the particle, the learning factor c_1 , c_2 , and c_3 , the maximum ω_{max} and the minimum value ω_{min} of the inertia weight.
- (5) Initializing the particle swarm. $t = 0$, the random generation of particles, including their own historical optimal position $pbest_i$, is the initialized position x_i , the global historical optimal position gbest is the optimal position of the initial particle population, and the velocity v_i and position x_i of each particle, and the fitness function Eq. (4) is calculated.
- (6) Let $t = t + 1$ and calculate the inertia weight according to (15) and adjust the speed and position of each particle according to (14, 16 and 17).
- (7) Updating the individual extremum $pbest_i$ and global extreme value gbest of each particle.
- (8) Judging whether the condition of the termination of the iteration is satisfied, and the global optimal solution is obtained if the algorithm has been satisfied; if not, return to step (5), and the algorithm continues to execute.

5 Simulation Studies and Result Analysis

In order to verify the localization performance of the algorithm, simulation comparison between PSO-DV-Hop algorithm and NJPDH algorithm is carried out. The MATLAB R2014b platform is used to test the 200 sensor nodes randomly in the area of 100×100 . The communication radius is 20 m, the beacon node is 25, the number of particle population is 30, the maximum velocity of the particle is 10 m/s, and the maximum iteration number is 100. The results and analysis of the simulation are shown in Figs. 1–3.

- (1) The influence of the total number of network nodes on the average positioning error.

As shown in Fig. 1, the number of nodes from 100 to 500 has an obvious effect. This is due to the increase in the number of nodes which will reduce the distance between nodes, and the network nodes are denser. Thus, the estimation of the jump distance is more accurate, and the route the node determines is closer to the best route, finally improving the positioning effect. However, the average positioning error of this paper is 6.4% lower than that of PSO-DV-Hop algorithm, so it has better location effect.

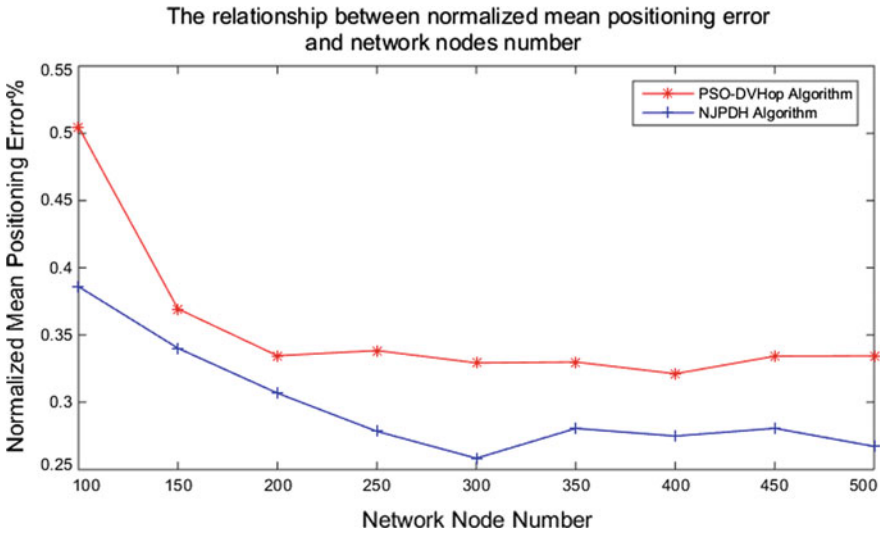


Fig. 1 Relationship between normalized mean positioning error and network node number

(2) Influence of beacon nodes on average positioning error

As shown in Fig. 2, the error of the two algorithms is obviously reduced when the number of beacon nodes is from 10 to 50, and the change of the positioning error gradually becomes smaller after the beacon node to 15. This is mainly due to the increase of the beacon nodes that make the unknown node get more accurate information on the hops and jumps of the beacon nodes, and thus the location error is reduced. By drawing, the average positioning error of this algorithm is reduced by 5.3% than that of the PSO-DV-Hop algorithm. It shows that the location of this algorithm can be exactly as accurate in the case of relatively small number of beacon nodes, and it can effectively reduce the cost in practice.

(3) The influence of the communication radius on the average positioning error

As shown in Fig. 3, when the communication radius is from 15 to 50, the error of the two algorithms is very small and the error varies little when the number of nodes is 20. This is because the network coverage of the beacon nodes is more extensive as the communication radius becomes larger. The number of hops and the average jumps are also available. It will be closer to the ideal value. The average error of this paper is 5.2% less than that of PSO-DV-Hop algorithm, which indicates that the algorithm has better localization performance in network coverage.

Therefore, this algorithm is more adaptive, and the algorithm has higher search efficiency and better stability.

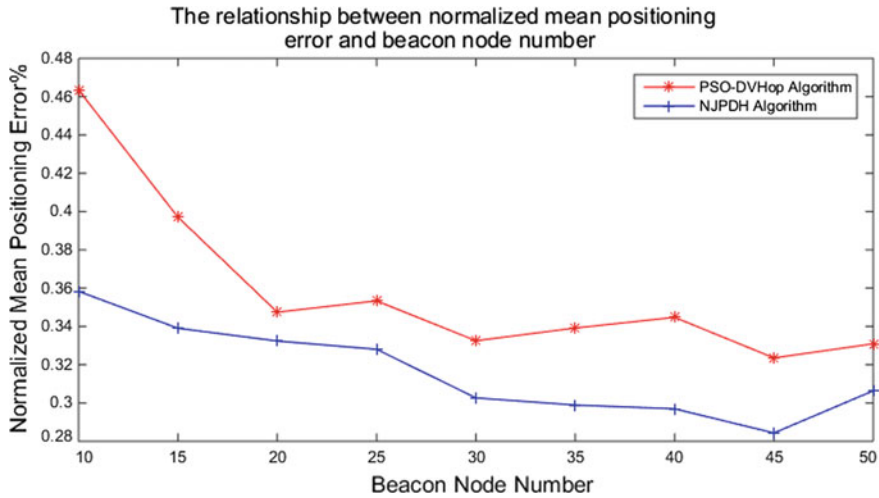


Fig. 2 Relationship between normalized mean positioning error and beacon node number

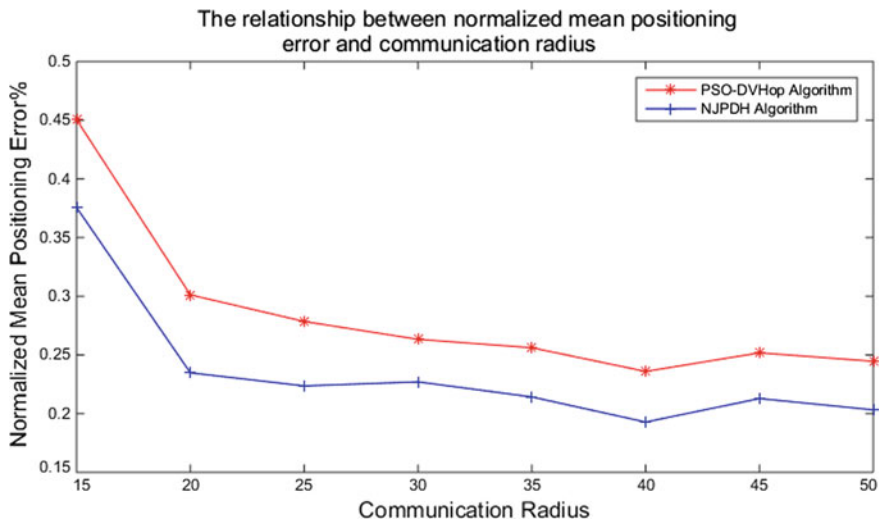


Fig. 3 Relationship between normalized mean positioning error and communication radius

6 Conclusions

This paper analyzes the error generated by the second and third stage of the DV-Hop positioning. By adding the weight of each beacon node and the weighted calculation of the average distance, it solves the problem that the average error is too large in the second stage, and the particle swarm optimization is optimized by increasing the inertia weight and the jump factor. In order to get more accuracy positioning

results, the third stage is improved and directly use the average distance and hop count to solve the location error of the unknown nodes. Simulation results show that the improved algorithm has a higher positioning accuracy under the same equipment condition, and is a better algorithm.

References

1. Zhao, J.F.: The key technology toward the self-driving car. *Int. J. Intell. Unmanned Syst.* **10**(1108), 2–20 (2018)
2. He, J., Li, L.Y.: Research on particle swarm optimization-based node positioning algorithm for wireless sensor networks. *Appl. Mech. Mater.* **10**(5028), 1415–1419 (2014)
3. Liu, S.L.: Optimization analysis of WSN location process based on hybrid PSO algorithm. In: *IEEE International Conference on Unmanned Systems*, pp. 1–3. IEEE BEIJING SECTION (2017)
4. Walid, M.I., Najah, A.A., Hossam, S.H.: Characterizing multi—hop localization for internet of things. *Wirel. Commun. Mob. Comput.* **10**(1002), 3316–3331
5. Guo, W., Dong, D.: The research on a kind of optimization of the DV - Hop localization algorithm. *J. Conver. Inf. Technol.* **9**(1205), 346–353 (2013)
6. Miao, F., Park, H.S., Kim, C., Ahn, S.: Swarm intelligence based on modified PSO algorithm for the optimization of axial-flow pump impeller. *J. Mech. Sci. Technol.* **10**(1007), 4867–4876 (2015)
7. Liu, B., Pan, H.X.: An expert system for fault diagnosis in diesel engine based on wavelet packet analysis and hybrid PSO-DV based neural network. In: *International Conference on Intelligent Computing and Cognitive Informatics*, pp. 1–4; Intelligent information technology application society (2010)
8. Kawakami, K.J., Meng, Z.Q.: Improvement of PSO and hybridization. *ITE Technical Report* **10**(11485), 1–4 (2009)
9. Zhou, H.X., Chen, C., Jia, J.Y.: Research on Wireless Sensor Networks with an Improved DV-Hop Localization Algorithm Based on Weighted Processing and Error Modification **10**(4028), 396–401 (2014)
10. Fernández, J.L., Martínez, E.G.: The PSO family: deduction, stochastic analysis and comparison. *Swarm Intell.* **10**(1007), 245–273 (2009)

Multi-objective Optimization Coordination for Urban Arterial Roadway Based on Operational-Features



Feng Qiao, Haochen Sun, Lingzhong Guo, Haolin Song and Zhaoyan Wang

Abstract In this paper, a new coordinated control model is proposed based on the vehicular operational features, and a multi-objective optimization algorithm NSGA-II is employed to the model for the operation of the vehicle traveling on an urban arterial road taking three evaluation indexes into consideration as the average vehicle delay, the queue length, and the vehicle exhaust emission. A numerical experiment was made in an urban arterial road with three intersections on VISSIM for the proposed strategy, and the simulation results were compared with two commonly used pre-timed methods: Webster's method and MAXBAND coordinated control method to verify the effectiveness of the proposed strategy in dealing with the unbalanced traffic volume condition, and it proved its advantages in designing and managing traffic systems more efficiently.

Keywords Multi-objective optimization · Coordinated control · Urban arterial road · Vehicle delay · Queue length · Exhaust emission

1 Introduction

In the modern city, urban transportation is the basis for the development of the urban economy. In recent years, with the rapid development of social economy and urbanization, it has been paid more and more attention to the transportation problems to improve the urban traffic system [1]. While the more developed the city always means the more vehicles the city has. So, some serious traffic problems are presented, such as the traffic congestion and the environmental pollution. In this case, it is the most important work for the urban traffic managers and researchers that how to solve

F. Qiao (✉) · H. Sun · H. Song · Z. Wang

Faculty of Information and Control Engineering, Shenyang Jianzhu University, Shenyang, China
e-mail: fengqiao@sjzu.edu.cn

L. Guo

Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_30

321

the traffic congestion and environmental pollution due to the increasing number of vehicles. At first, all the transportation engineers focus on the improvement in the traffic infrastructure and the vehicle technology, which have no apparent effect on dealing with the traffic congestion and the environmental pollution. Then, it is required that all the researchers have to find the other efficient way to solve such problems, for example, using the method of traffic management and control.

With further research on the traffic management and control, it is clearly recognized that the traffic control strategy should be made to improve the traffic efficiency, throughput, and capacity, and diminish the vehicle delay and travel time [2–4]. And, the indexes for evaluating the intersection operational efficiency always include average vehicle delay, stop times, queue length, traffic capacity, fuel consumption, exhaust emission, etc., all of which have been influenced directly by the signal timing at intersections [5–7]. So, only by making a proper signal timing plan, can the operational efficiency be increased, and the vehicle exhaust emission be reduced, which means a more environmentally friendly transportation system is constructed.

In this paper, both the intersection operational efficiency and environmental benefit are taken into consideration, i.e., there are three operational features selected as the evaluation indexes employed comprehensively to optimize the intersection signal timing in this research, which is the average vehicle delay, queue length, and vehicle exhaust emission. But, it is bothersome that there exist some certain relations among these evaluation indexes, some of which even are contradictory. So, to handle such a contradiction, it is necessary to settle it as a multi-objective optimization problem, which means, in this study, the modeling and simulation of signal timing at intersections based on multi-objective optimization theory are researched. In this paper, a new delay model is proposed, which divides the vehicle delay into two parts: inbound and outbound, by which the problem of the coordinated control for the unbalanced traffic volume on each way of the urban arterial road is easy to be handled. And, a multi-objective optimization tool NSGA-II is used to solve such a problem and figure out the optimal signal timing. It is desirable to apply the coordinated control method on the urban artery to decrease the average vehicle delay, queue length, and vehicle exhaust emission by constructing a multi-objective optimization model.

The remaining part of this paper is organized as follows. In Sect. 2, some necessary background knowledge is introduced. In Sect. 3, the multi-objective optimization model is constructed. Then, the numerical experiments of optimizing the arterial road signal coordinated control based on Webster's method, MAXBAND model, and the proposed model are carried out on MATLAB for an arterial roadway with three-intersection, respectively, and in Sect. 4, the optimal results are applied on VISSIM for simulation with analysis and discussion. The conclusions are drawn in Sect. 5.

2 Background Knowledge

2.1 Multi-objective Optimization Problem

As the name suggests, multi-objective optimization means optimizing multiple objectives simultaneously [8]. But, the challenge is when the objectives are of contradiction to each other, i.e., the optimal solution of one of the objective functions is different from that of the other. In multi-objective optimization problems, the objective functions are to be either minimized or maximized. As in a single-objective optimization problem, the multi-objective optimization problem may involve a number of constraints, which any feasible solution must satisfy. According to the optimization theory, the objective which is maximized can be transformed into that which is minimized, so in this paper, we state the multi-objective optimization problem in its general form as

$$\left. \begin{array}{ll} \text{minimize } f_m(\mathbf{x}), & m = 1, 2, \dots, M; \\ \text{subject to } g_j(\mathbf{x}) \geq 0, & j = 1, 2, \dots, J; \\ & h_k(\mathbf{x}) = 0, \quad k = 1, 2, \dots, K; \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)} \quad i = 1, 2, \dots, n \end{array} \right\} \quad (1)$$

where $\mathbf{x} \in \mathbf{R}^n$ is a solution vector of n decision variables, written as $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$; $f_m(\mathbf{x})$ is the m th objective function; $g_j(\mathbf{x})$, and $h_k(\mathbf{x})$ mean the j th inequality constraint and the k th equality constraint, respectively; $x_i^{(L)}$ and $x_i^{(U)}$ represent the lower bound and the upper bound of the variable x_i , respectively.

2.2 Non-dominated Sorting Genetic Algorithm: NSGA-II

The traditional multi-objective evolutionary algorithms (MOEAs) have some deficiencies as they use non-dominated sorting and sharing have been criticized primarily for them: (1) the high-cost computational complexity $O(MN^3)$; (2) the non-elitism approach; and (3) the requirement for specifying a sharing parameter. So, in 2002, an improved multi-objective genetic algorithm NSGA-II is presented, which is a non-dominated sorting-based MOEA, alleviating all the three principal difficulties [9]. And, the simulation results on some complex test problems show that the proposed NSGA-II is able to find a much better spread of solutions and better convergence near the true Pareto-optimal front compared to other elitist MOEAs in most cases. In the field of traffic engineering, up to now, there are many researchers employing the NSGA-II or its improved versions for optimizing the traffic signal timing plan at urban intersections [10, 11].

2.3 Selection of the Evaluation Indexes

In this paper, three evaluation indexes are taken into consideration: the average vehicle delay, queue length, and vehicle exhaust emission. For the urban arterial road signal coordinated control, there exist some conditions that if the artery satisfies to these conditions, employing the signal coordinated control will improve the overall operational efficiency; but if not, employing the coordinated control strategy arbitrarily will get an opposite result. There is much research work which shows that the signal coordinated control is preferred to be set on a one-way road in undersaturated conditions.

Average Vehicle Delay Model. According to the conditions, a new vehicle delay model is proposed. This new model has two parts: inbound delay and outbound delay, which are different from the delay model employed on the existing previous researches. Through the proposed delay model, the coordinated control of the urban arterial road that has unbalanced traffic volume on each way can be accomplished. And, the outbound delay model is expressed as

$$d_{(i+1)d} = 0.5t_r q_d(t_d + t_r) = \frac{t_r^2 q_d u}{2(u - q_d)} \quad (2)$$

$$d'_{(i+1)d} = 0.5t_{ed} q_d(t_d + t_{ed}) = \frac{t_{ed}^2 q_d u}{2(u - q_d)} \quad (3)$$

$$D_d = \sum_{i=2}^n [\alpha_i d_{id} + (1 - \alpha_i) d'_{id}] \quad (4)$$

where $d_{(i+1)d}$ is the delay of vehicles when arriving at red at the $(i + 1)$ th intersection and $d'_{(i+1)d}$ is the delay of vehicles when arriving after red at the $(i + 1)$ th intersection; t_r , t_d represent the red time in a cycle, and the outbound evanescing time, respectively; u is the maximum traffic capacity of the intersection during green; q_d is the outbound traffic volume; t_{ed} is given as $t_{ed} = \varphi_{i,i+1} - \left\lfloor \frac{l_i}{v_d} \right\rfloor (\text{mod } T)$; and α_i is a Boolean function.

The inbound delay model is expressed as

$$d_{iu} = 0.5t_r q_u(t_u + t_r) = \frac{t_r^2 q_u u}{2(u - q_u)} \quad (5)$$

$$d'_{iu} = 0.5t_{ed} q_u(t_u + t_{ed}) = \frac{t_{ed}^2 q_u u}{2(u - q_u)} \quad (6)$$

$$D_u = \sum_{i=2}^n [\beta_i d_{iu} + (1 - \beta_i) d'_{iu}] \quad (7)$$

where d_{iu} is the delay of vehicles when arriving at red at the i th intersection, and d'_{iu} is the delay of vehicles when arriving after red at the i th intersection; q_u and t_u

are the inbound traffic volume and the inbound evanescing time, respectively; t_{ed} is given as $t_{ed} = T - \varphi_{i,i+1} - \left\lceil \frac{l_i}{v_u} \right\rceil (\text{mod } T)$; and β_i is a Boolean function.

Then, the integrated vehicle delay at all intersections in all directions is expressed as

$$\begin{aligned} D &= D_u + D_d \\ &= \sum_{i=2}^n [\beta_i d_{iu} + (1 - \beta_i) d'_{iu}] + \sum_{i=2}^n [\alpha_i d_{id} + (1 - \alpha_i) d'_{id}] \\ s.t. & 0 \leq \varphi_{i,i+1} \leq T \end{aligned} \quad (8)$$

Queue Length Model. As the intersection is in an undersaturated condition, the queue length during the n th cycle can be expressed as

$$l_n = \frac{\exp\left[-\frac{4}{3}\sqrt{\lambda C q_s} \times \frac{1-x}{x}\right]}{2(1-x)} + q_n C (1 - \lambda) \quad (9)$$

where λ is the green split; C means the cycle length; q_s is the saturated vehicle flow rate; x represents the degree of saturation; q_n represents the vehicle arrival rate in the n th cycle.

Vehicle Exhaust Emission Model.

$$E = \sum_j [EF^{PCU} \times q_j \times (L_j - l_j)] + \frac{1}{3600} \sum_j (EFI^{PCU} \times q_j \times D_j) \quad (10)$$

where E is the total quantity of the exhaust emissions; EF^{PCU} represents the emission factor; EFI^{PCU} is the idling emission factor; q_j means the traffic flow volume on the j th segment of the roadway; L_j and l_j represent the length of the j th segment and of the queue on the j th segment, respectively; and D_j is the average vehicle delay at the intersection.

According to the above, a general expression of the multi-objective signal coordinated control for the arterial road can be written as:

$$\min(D, l_n, E) \quad (11)$$

3 Modeling the Multi-objective Optimization Problem on the Urban Arterial Roadway

The basic consideration of the arterial road coordinated control is that regarding all the intersections on the artery as an integrated system, every two adjacent intersections have a spatio-temporal relationship admitting that as the first vehicle of a platoon just

arrives at the downstream signal and the light turns green, i.e., the platoon can pass through the arterial roadway without stop. And in doing so, it can be guaranteed that the vehicle running on the arterial road is able to maintain a best high speed during its travel, and the platoon can pass through intersections freely as much as possible. This approach improves the operational efficiency of the whole traffic system.

3.1 Some Main Parameters for Describing the Arterial Traffic Coordinated Control

It is required that all the signal intersections on the arterial road should have a **critical cycle** (always select the maximum cycle length among all intersections as the critical cycle) to make the signal coordinated control among the intersections easier. **Split** is calculated based on the intersection phasing and expected demand, so it is a portion of the time allocated to each phase at an intersection. And in traditional coordination logic, the split for the non-coordinated phases defines the minimum amount of green in the coordinated phases. In the research on the arterial traffic coordinated control, all the intersections employ a critical cycle, but apply different splits due to the different traffic operational conditions. **Offset** is a very important control parameter in the green-wave-coordinated control. This term defines the time relationship between coordinated phases at subsequent traffic signals, and is the time difference between the beginning of green phases for a continuous traffic movement at successive intersections that may give rise to a green wave along the arterial road.

3.2 Webster's Arterial Traffic Coordinated Control Model

The Webster coordinated control model is one kind of pre-timing models, which means all the intersections on the arterial road have the same predetermined cycle length as the critical cycle [11]. And in the Webster model, the critical cycle is the maximum cycle length among all the cycle lengths of intersections, obtained by the Webster's optimal cycle length method. The optimal cycle length of each intersection is obtained as

$$C_o = \frac{1.5L + 5}{1 - Y} \quad (12)$$

where C_o means the optimal cycle length; L is the total loss time at the intersection, given as $L = \sum(l + I - A)$, l being the start loss and always being set as equal to 3 s, I being the green interval set as equal to 7 s, and A being the amber time set as equal to 3 s; and Y is the sum of the maximum flow rate in each phase, given as $Y = \sum \max[y_1, y_2, \dots]$.

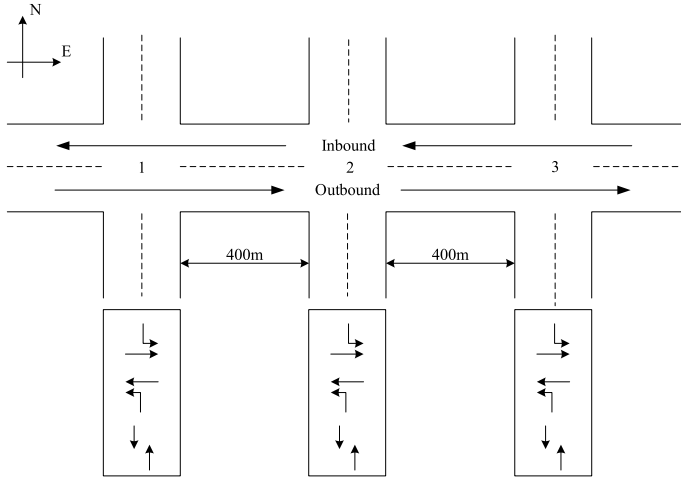


Fig. 1 The diagram of the arterial road and the signal phases

Then, select a proper offset by defining the time relationship between coordinated phases at subsequent traffic signals, and the time difference between the beginning of green phases for a continuous traffic movement at successive intersections which may give rise to a green wave along the arterial road. For a two-way coordinated control, the offset is expressed as

$$\Delta p = \text{mod} \left(\frac{T}{C_c} \right) \quad (13)$$

where Δp represents the offset between the adjacent intersections; T means the average travel time along the section between the two adjacent intersections; and C_c is the critical cycle length.

3.3 An Improved MAXBAND Arterial Traffic Coordinated Control Model

The MAXBAND model is proposed by Little to maximize the width of the green wave [12]. Give names to the n intersections along the arterial road, such as $I_1, I_2, \dots, I_n$, and assume the direction that from I_1 to I_n is **outbound**, and the opposite direction is the **inbound** (see Fig. 1 as the details).

The expression of MAXBAND model is shown as

$$\begin{aligned}
 Z_2 &= \max(b + \bar{b} - K_P |K_R b - \bar{b}|) \\
 \text{s.t. } w_i + b' &\leq 1 - r_i, \bar{w}_i + \bar{b}' \leq 1 - \bar{r}_i & i = 1, 2, \dots, n \\
 w_i + \bar{w}_i - (w_{i+1} + \bar{w}_{i+1}) + t_{(i,i+1)} + \bar{t}_{(i,i+1)} \\
 + \Delta_i - \Delta_{i-1} &= -0.5(r_i + \bar{r}_i) + 0.5(r_{i+1} + \bar{r}_{i+1}) \\
 + m_{(i,i+1)} & & i = 1, 2, \dots, (n-1) \\
 m_{(i,i+1)} &\in \mathbf{Z} & i = 1, 2, \dots, (n-1) \\
 w_i, \bar{w}_i &\geq 0 & i = 1, 2, \dots, n \\
 b = \varepsilon(b')b', \bar{b} &= \varepsilon(\bar{b}')\bar{b}' & (14)
 \end{aligned}$$

where $b(\bar{b})$ means the ratio of the outbound (inbound) bandwidth to the critical cycle; K_P is an influencing factor for distributing the green-wave bandwidth of inbound and outbound; K_R is a proportionality coefficient of the demand of the inbound and the outbound bandwidths; $r_i(\bar{r}_i)$ means the ratio of the outbound (inbound) red time to the critical cycle at I_i ; $w_i(\bar{w}_i)$ represents the ratio of the time from the right (left) side of red to the left (right) edge of outbound (inbound) green band to the critical cycle at I_i ; $t_{(h,i)}(\bar{t}_{(h,i)})$ represents the ratio of the travel time from I_h to I_i outbound (I_i to I_h inbound) to the critical cycle; $b'(\bar{b}')$ represents an intermediate variable; Δ_i is the ratio of the time from the center of $\bar{r}_i$ to the nearest center of r_i to the critical cycle; $m_{(i,i+1)}$ is the loop integer in recognition of the more general case of networks, given as $m_{(i,i+1)} = \Phi_{(i,i+1)} + \bar{\Phi}_{(i,i+1)} + \Delta_i - \Delta_{i+1}\hat{A}$, $\Phi_{(h,i)}(\bar{\Phi}_{(h,i)})$ being the ratio of the time from the center of an outbound (inbound) red at I_h to the center of a particular outbound (inbound) red at I_i to the critical cycle; and $\varepsilon(x)$ is a step function.

3.4 The Operational-Feature-Based Urban Arterial Traffic Coordinated Control Model

The original method for optimizing the arterial traffic signal is to maximize the green wave band which only takes the vehicles on the major road into account and ignores the traffic operational condition on the minor road, i.e., the proposed method can guarantees the maximization of the green-wave bandwidth on the major road, which will also reduce the average vehicle delay on the minor road. So, in this paper, a new optimization model is expressed as

$$\begin{cases} F_1 = \min \left(\sum_{i=1}^n d_{1i}, \sum_{i=1}^n l_{1i}, \sum_{i=1}^n e_{1i} \right) \\ F_2 = \min \left(\sum_{i=1}^n d_{2i}, \sum_{i=1}^n l_{2i}, \sum_{i=1}^n e_{2i} \right) \end{cases} \quad (15)$$

$$s.t. \begin{cases} g_{ij\min} \leq g_{ij} \leq g_{ij\max} \\ C_{i\min} \leq C_i \leq C_{i\max} \\ x_{i\min} \leq x_i \leq x_{i\max} \\ \varphi_{i,i+1\min} \leq \varphi_{i,i+1} \leq \varphi_{i,i+1\max} \end{cases}$$

where $F_1(F_2)$ means the multi-objective optimization function in the inbound (outbound) direction; $d_{1i}(d_{2i})$ is the average vehicle delay at I_i in the inbound (outbound) direction; $l_{1i}(l_{2i})$ represents the vehicle queue length at I_i in the inbound (outbound) direction; $e_{1i}(e_{2i})$ is the quantity of the vehicle exhaust emissions at I_i in the inbound (outbound) direction; C_i is the cycle length at I_i ; $C_{i\min}$ and $C_{i\max}$ represent the lower bound and the upper bound of the cycle length at I_i , respectively; g_{ij} is effective green time at I_i in the j th phase; $g_{ij\min}$ and $g_{ij\max}$ are the lower bound and the upper bound of the effective green time at I_i in the j th phase, respectively; x_i is the degree of saturation at I_i ; $\varphi_{i,i+1}$ is the phase difference between I_i and I_{i+1} ; $\varphi_{i,i+1\min}$ and $\varphi_{i,i+1\max}$ are the lower bound and the upper bound of the phase difference between I_i and I_{i+1} , respectively; and $x_{i\min}$ and $x_{i\max}$ mean the lower bound and the upper bound of the degree of saturation at I_i , respectively. In this paper, $x_{i\min}$ and $x_{i\max}$ are equal to 0.7 and 0.9, respectively.

4 Numerical Experiment

It is assumed that there is an artery (major) road (east-west) of dual-course six-lane with no-left-turn, there are with three successive intersections along the road, and a dual-course four-lane minor road (south-north). And, the major road has a right turn lane in each direction at each intersection. At each intersection, the signal is set with three phases. The experimental arterial road and the corresponding signal phases are shown in Fig. 1. The traffic data of each approach at each intersection is listed in Table 1.

Table 1 The experimental traffic data of each approach

Intersection		Phase 1	Phase 2	Phase 3
Intersection 1	Saturation flow (veh/h)	6660	6660	2160
	Flow ratio	0.259	0.282	0.164
Intersection 2	Saturation flow (veh/h)	6660	6660	2160
	Flow ratio	0.294	0.269	0.199
Intersection 3	Saturation flow (veh/h)	6660	6660	2160
	Flow ratio	0.313	0.301	0.183

According to the cycle length at each intersection, some basic traffic parameters can be determined: the critical cycle is 179s; the initial phase differences are $\varphi_{1,2} = 50s$, and $\varphi_{2,3} = 40s$. The optimal signal timing results with 10 successive cycles of the three mentioned models via MATLAB are obtained and listed in Table 2. And, the results are carried on the microscopic traffic simulation software VISSIM, and the simulation values of evaluation indexes (average vehicle delay, queue length, and vehicle exhaust emission) of the three models are shown in Fig. 2.

From Table 2 and Fig. 2, it is obviously known that: (1) the general cycle-by-cycle trends obtained by three control models are similar approximately, and the Webster model has high average vehicle delay, queue length, and exhaust emission, because it does not use the idea of the coordinated control method, while by employing the idea of coordination, the MAXBAND model plays a good performance that, in each cycle, in general, the values of three evaluation indexes all are lower compared with those obtained by Webster’s method; (2) as the model proposed in this paper is used to solve this problem, one can get more perfect results, which are significantly lower than those calculated by the other two pre-timed methods; (3) the average vehicle delay obtained by the proposed model declined by an average of 28.6% with respect to Webster’s method, and 21.8% with respect to the MAXBAND model, and the queue length computed by the proposed model decreased by an average of 16.4% with respect to the Webster method, and 11.6% with respect to the MAXBAND model,

Table 2 The results via MATLAB with three models

Webster’s coordinated control			Original coordinated control ($\varphi_{1,2} = 29s, \varphi_{2,3} = 28s$)			Proposed optimization model ($\varphi_{1,2} = 20s, \varphi_{2,3} = 12s$)		
Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
72s	68s	38s	59s	60s	24s	38s	38s	23s
cycle length = 179s			cycle length = 155s			cycle length = 112s		

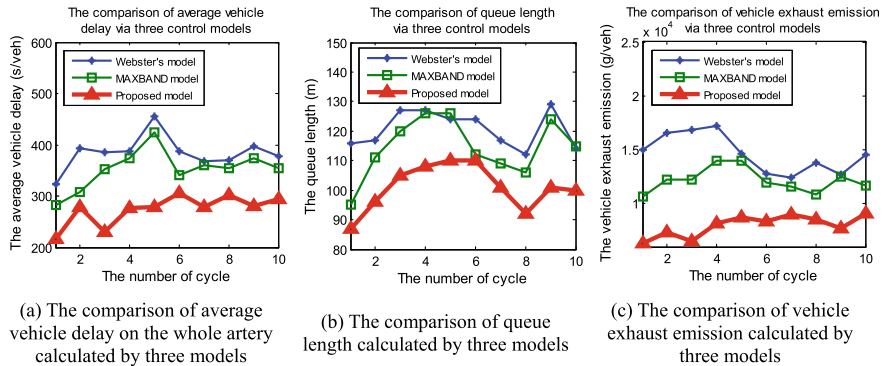


Fig. 2 The comparison of the evaluation indexes via three control models

and a laudable thing is the vehicle exhaust emission calculated by the proposed model dropped by an average of 44.5% with respect to the Webster method, and 34.1% with respect to the MAXBAND model, which is more environmentally friendly.

5 Conclusions

In this paper, some work has been done. First, an introduction of some basic knowledge is presented as the base of the whole research work in this paper. Then, the multi-objective optimization problem a three-intersection arterial road in an urban area has been discussed adequately, and some key parameters of it are suggested, including the critical cycle, green split, and offset. Thereafter, the parameter offset has been discussed further that divide the offset into two parts: one-way offset and two-way offset. And, a commonly used coordinated control model MAXBAND is introduced. After this, a new model an operational-feature-based coordinated control model for the arterial roadway is proposed in this paper, and in this model, three vehicular operational features are selected as the evaluation indexes: the average vehicle delay, the queue length, and the vehicle exhaust emission. Additionally, this new proposed model separates the average vehicle delay into two segments: inbound vehicle delay, and outbound vehicle delay, by which, the problem that the coordinated control for the unbalanced traffic volume on each way of the arterial road can be solved perfectly to make the operational efficiency of the whole traffic system optimal. Then, a numerical experiment is suggested, in this experiment, the signal phase has three parts, by which the minor road and the major road are both taken into consideration that the traditional models did not. And in the simulation experiment, first, three models: Webster's model, improved MAXBAND model, and proposed model (solved by a multi-objective optimization tool NSGA-II), are employed on MATLAB to output the corresponding optimal results; second, the obtained results are put in a traffic simulation software VISSIM. From the resulting simulation, it is clearly known that the proposed model can solve the coordinated control problem better than the other two previous models, especially, the proposed model has a strong ability to decrease the vehicle exhaust emission, which is becoming an important goal in the management and control of the modern intelligent transportation system. From this paper, the transportation engineers and professionals will get a preferable tool for managing the traffic system.

Acknowledgements This research is partially supported by the National Natural Science Foundation of China (No. 61703288).

References

1. Pisarski, A.E.: Prescriptions for research: reviewing the history of TRB's critical issues in transportation. *Tr News* (2003)
2. Zhang, X., Yu, L., Liu, M.: Traffic demand-based traffic congestion measurement models for road networks. *Modern Transp. Technol.* **5**(6), 71–75 (2008) (in Chinese)
3. Vreeswijk, J.D., Mahmood, M.K.M., Van Arem, B.: Energy efficient traffic management and control—the eCoMove approach and expected benefits. In: *International IEEE Conference on Intelligent Transportation Systems* (pp. 955–961). IEEE (2010)
4. Kirkby, P., Kadengal, R.: Traffic management and control using a single 'congestion price' like variable across multiple layers of network hierarchy. *Control Next Gener. Netw.* **67**, 4/1–4/6. IET (2002)
5. Samadi, S., Rad, A.P., Kazemi, F.M., Jafarian, H.: Performance evaluation of intelligent adaptive traffic control systems: a case study. *J. Transp. Technol.* **2**(3), 248–259 (2012)
6. Tang, T., Lu, J., Zhang, C., Zhang, W., Zhang, Y.: Optimal speed limit investigation on highway for emission reductions. *Modern Transp. Technol.* **15**(1), 82–86 (2018) (in Chinese)
7. Grote, M., Williams, I., Preston, J., Kemp, S.: Including congestion effects in urban road traffic CO₂ emissions modelling: do local government authorities have the right options? *Transp. Res. Part D Transp. Environ.* **43**, 95–106 (2016)
8. Xiao, X.W., Xiao, D., Lin, J.G., Xiao, Y.F.: Overview on multi-objective optimization problem research. *Appl. Res. Comput.* **28**(3), 805–167 (2011)
9. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: nsga-ii. *IEEE Trans. Evol. Comput.* **6**(2), 182–197 (2002)
10. Cui, M.Y., Liu, H.Z., Liu, W., Jiang, H.L., Wang, Z.P., Huang, Y.D.: Multi-objective optimization based on improved nsga-II algorithm. *J. Nanyang Normal Univ.* **13**(12), 8–12 (2014)
11. Qiao, F., Sun, H., Wang, Z., Tobi, F.A.: NSQGA-based optimization of traffic signal in isolated intersection with multiple objectives. *Innov. Tech. Appl. Modell. Ident. Control*, 291–305 (2018)
12. Little, J.D.C., Kelson, M.D., Gartner, N.H.: Maxband: a program for setting signal on arteries and triangular network. *Transp. Res. Rec. J. Trans. Res. Board* **795**, 40–46 (1981)

Applicability Analytic of Closed Intersection Along Tramway Based on Simulation



Weidong Liu, Xingquan Guan, Yixiu Wu and Shanshan Fan

Abstract Whether the intersections of the tramway are closed have a very significant influence on the operating efficiency of trams and vehicles. How to coordinate the right-of-way relationship between trams and vehicles is of vital importance improving the overall traffic efficiency of the corresponding intersection. Based on the historical data of the intersections along Hunnan Modern Tram No. 5 line, we studied the relationship between the ratio of traffic flow to non-winding traffic flow and the delay. A single-factor analysis of the applicability of closed conditions of the intersections was carried out by simulation, and the optimal applicability conditions were achieved. The conclusions of this study provide a scientific basis for intersection signal timing optimization of the tramway and traffic system.

Keywords Closed intersection · Tramway · Road right conflict · Simulation

1 Introduction

As a relatively new development of public traffic tool, tramcar has a series of irreplaceable advantages, such as high resource utilization, low environmental pollution level, moderate traffic volume, and comfortable riding. These advantages make the vigorous development of the tramcar a new trend. In order to ensure the efficient operation of a tramcar, it is necessary to provide control strategies of priority when the tramcar is in conflict with other vehicles. Traffic signal control at the intersection along the tramway has a great impact on traffic delay and queue length. Therefore, many scholars from various countries make efforts in signal control optimization of tramcar intersections.

W. Liu · X. Guan (✉) · S. Fan

School of Transportation Engineering, Shenyang Jianzhu University, Shenyang 110168, China
e-mail: 44667276@qq.com

Y. Wu

School of Civil Engineering, Shenzhen University, Shenzhen 518067, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_31

Considering the signal timing will affect the operation of tramcar along the intersection, Currie and Shalaby [1] adjusted the operation schedule of tramcar to realize tramcar has more priority than buses according to the traffic signal timing. Skabardonis [2] divided signal priority into passive priority, active priority, and real-time priority. A method of trunk line coordination control is studied, and find that the passive priority strategy has a good control effect compare the active priority when the tram traffic is large and running stably. Based on the loss calculation of all intersections, Luiz and Werner [3] analyzed the way of modern tramcar leaving and approaching the platform, proposed a method of selecting priority according to specific priority requests, and demonstrated the advantage of the controlled departure of tramcar. Yu and Yang [4] proposed a partial priority control strategy applying to the peak direction, moreover, the applicable situation is also analyzed to ensure that this control strategy is stopped when the benefit of the public transport system is maximized. At the same time, a heuristic algorithm is defined to evaluate the reliability of the service. Motivated by the backpressure routing, Wu et al. [5] proposed delay-based traffic signal control algorithm in transportation networks. The proposed delay-based control achieves optimal throughput performance.

In general, most of the scholars study the coordination of the conflict between tramcar and other vehicles, and focus on the signal optimization of intersections, however, there are few researches on the sealing applicability of the intersections along the tramcar.

2 The Applicability of Closed Intersection

2.1 Traffic Organization After Closing Intersection

After the intersection is closed, a large number of bypass vehicles enter the bypass intersection, and the traffic volume of the bypass intersection has changed. In order to allow vehicles to pass through the intersection smoothly and reduce the delay of vehicles passing through the intersection, the traffic timing of the bypass intersection should be adjusted so that it can still be relatively smooth after entering the bypass vehicle, without traffic jams and excessive delays.

From Fig. 1, we can see that the most important indicator of the distance of the detour route is the distance between the intersection and the adjacent intersection on both sides. The shorter the distance, the shorter the overall detour route, and the corresponding delay will be reduced. And, it is also better to analyze its overall indicators from the perspective of fuel economy and automobile carbon emissions.

The passenger capacity of trams passing through the intersections and the direct traffic flow in the same direction as the trams account for a large proportion of the traffic at the intersections [6]. When the intersection is closed, the greatest benefit from the intersection's traffic efficiency is to go straight in the same direction as the tram. When the intersection is closed, the traffic lights will be canceled. The delay



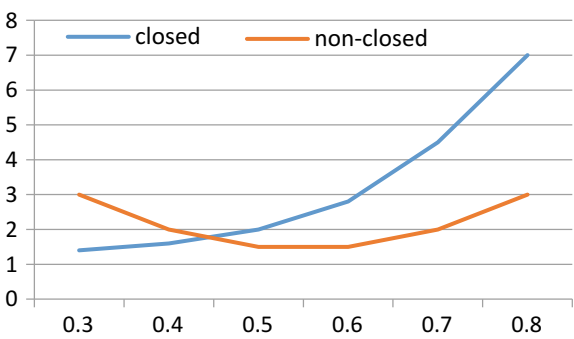
Fig. 1 The vehicle bypass route

of above two types of vehicles pass through the intersection will be greatly reduced, and the greater the carrying capacity of these two types of vehicles, the higher the utilization rate of this type of less delayed traffic is, and the higher the effectiveness of the closure of the entire intersection, that is, after the intersection is closed. The ratio of bypass traffic to non-bypass traffic is relatively low, and this ratio is called the traffic flow ratio of the intersection. The lower the ratio, the more suitable the intersection is closed.

The relationship between traffic flow ratios and intersection delays under the ideal conditions is shown in Fig. 2.

The ordinate of Fig. 2 is the delay level of the intersection group, and the abscissa is the traffic flow ratio of the intersection. In the unclosed state, since the set total traffic volume is constant, when the traffic flow ratio is too large or too small, there will be excessive traffic flow in one direction or several directions leading to an overall increase in the delay of intersections. However, if the intersection is closed, traffic flow is relatively small, vehicles and trams in the direction of trams that have a relatively large amount of traffic have almost no delay through the intersection,

Fig. 2 Relationship between flow ratio and delay level in ideal condition



and there are fewer bypass vehicles. The delay caused by the bypass increases not obviously, so the overall level of delay is relatively low. When the traffic volume is relatively large, the number of bypass vehicles increases, which causes a great impact on the intersections passed by. And when the ratio exceeds a certain value, the bypass intersection will not be able to withstand the influx of so many vehicles, causing great delays, even traffic jams. When we set the intersection point of the two curves, the traffic flow ratio is A . If we only consider the influence of this factor, then when the traffic flow ratio is greater than A , it is not suitable for closure; when the traffic flow ratio is less than A , it is suitable for closure.

2.2 Intersection Surrounding Environment Reference Conditions

The bypass distance directly affects the delay of the bypass vehicle passing through the area. From the economical point of view, if the bypass distance is too large, the intersection is not suitable for closure [4]. The geographical information at intersections and their adjacent intersections will directly affect the adaptability of intersections.

There are no important institutions near the intersections such as schools, hospitals, gas stations, government service agencies, and other important educational and medical service institutions [7]. Such institutions will increase traffic at the intersection and pedestrians passing through the intersection, which greatly increases the complexity of the traffic at the intersection, and medical institution such as hospitals, whether ambulances, or patients or their families driving, detour into the hospital are unacceptable, medical service-related vehicles should have the greatest priority in traffic [8]. However, closed intersections pose great risks for them, so the surrounding environment at the intersections is very important [9].

3 Simulation Modeling

This paper evaluates the overall traffic efficiency by calculating the per capita delay. In this survey, the delay before closure can be obtained from the following equation:

$$T = \frac{X_1(\sum d_1 t_1 + \sum d_2 t_2 + \sum d_3 t_3) + X_2(\sum y_1 t_4 + \sum y_2 t_5 + \sum y_3 t_6)}{X_1 \sum d + X_2 \sum y} \quad (1)$$

where

$d_1 - d_3$ d_1 represents the hourly traffic flow at the intersection of Hunnan Middle Road and Langyue Street; d_2 represents the hourly traffic flow at the inter-

- section of Hunnan Middle Road and Langri Street; d_3 represents the hourly traffic flow at the intersection of Hunnan Middle Road and Langyun Street;
- $y_1 - y_3$ They represent the number of trams that pass each hour at three intersections. The intersection sequence corresponds to the above.
- $t_1 - t_3$ The delays at the intersections of the three intersections are, respectively indicated, and the intersection order corresponds to the above (trams are not included).
- $t_4 - t_6$ The delays at the intersections of the trams at the three intersections are, respectively indicated, and the intersection sequence corresponds to the above.
- d The sum of cars driving through these three intersections within an hour.
- y The sum of the number of tram vehicles passing through the three intersections within an hour.
- X_1, X_2 Represent the average passenger capacity of single cars and trams, respectively. Based on the passenger capacity of the road vehicles and the passenger capacity of trams, X_1 and X_2 take 2 and 120, respectively.

The per capita delay after closure can be obtained in (2):

$$T = \frac{X_1(\sum d_2t_2 + \sum d_3t_3 + T_1) + X_2(\sum y_2t_4 + \sum y_3t_5)}{X_1 \sum d + X_2 \sum y}, \quad (2)$$

where T_1 in (2) is the vehicle detour time, can be calculated by the following equation:

$$T_1 = \left(\sum d_4t_7 + d_4 \frac{L_1}{V} + \sum d_5t_7 + d_5 \frac{L_1}{V} + \sum d_6t_8 + d_6 \frac{L_2}{V} + \sum d_7t_8 + d_7 \frac{L_2}{V} \right) \quad (3)$$

where

- $d_4 - d_7$ d_4 represents the number of vehicles left to operate at the east entrance of the junction before closure; d_5 represents the number of vehicles in the north of the junction before the closure of the road to the north for direct or left-turn operations; d_6 indicates the number of vehicles that were to be operated on the left at the entrance of the junction before closure; d_7 indicates the number of vehicles in the south of the intersection that were to be used for direct or left-turn operations before closing.
- $t_7 - t_8$ t_7 indicates the delay at the intersection of the east entrance turning vehicle at the intersection of Hunnan Middle Road and Langyun Street; t_8 indicates the delay at the intersection of the west entrance of the intersection of Hunnan Middle Road.
- L_1, L_2 L_1 represents the round-trip distance of the vehicles passing through the intersection of Hunnan Middle Road and Langyun Street, and the previous investigation of L_1 takes 800 m. L_2 represents the detour distance of the vehicle passing through the intersection of Hunnan Middle Road and Langri Street, and the previous investigation L_2 takes 930 m.

Table 1 Comparison of data before and after the intersection closed

	Per capita delays/s	Numbers of stops
Unclosed intersection	28.6	0.40
Closed intersection	28.5	0.35

V indicates the average speed of the vehicle on the detour route, through an investigation of 40 km/h.

The calculated results of the substitution data are shown in Table 1. It is easy to see that the per capita delays and numbers of stops before and after the closure of the studied intersections have been reduced.

4 Experiment and Result Analysis

4.1 Data Survey and Experimental Design

This paper selects Tram Line 5 in Hunnan, Shenyang as the research object, through artificial statistical method to get the data.

According to the statistics of the survey data, the traffic flow data of the three intersections are shown in the following figures: Fig. 3 is the intersection of Hunnan Middle Road and Langyue Street, which is the survey intersection; Fig. 4 is the intersection of Hunnan Middle Road and Langri Street, which is located at the east side of the survey intersection and adjacent to the survey intersection; Fig. 5 is the intersection of Hunnan Middle Road and Langyun Street, which is located at the west side of the survey intersection and adjacent to the survey intersection.

Use the collected data, and the input traffic flow, the traffic signal timing, the traffic decision-making direction, and the other information into VISSIM software. Then check each parameter information, and record road sections and data (delay), which are set to simulate the collected data. Due to the closure of the survey intersections, the traffic flow in the direction of the left-turning direction of the adjacent intersections increases sharply, so it should be reflected in the traffic distribution. Therefore, the green time of the east–west left-turn phase of the calculation is increased before the closure, so as to ensure that the entire intersection is not congested.

Fig. 3 Vehicle flow of the intersection of Hunnan Middle Road and Langyue Street (pcu/h)

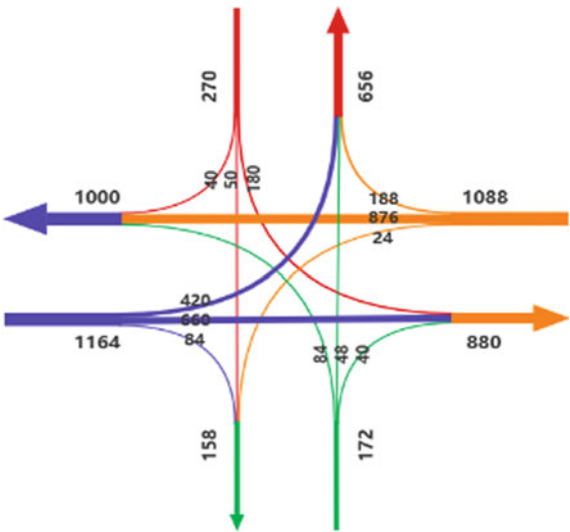


Fig. 4 Vehicle flow of the intersection of Hunnan Middle Road and Langri Street (pcu/h)

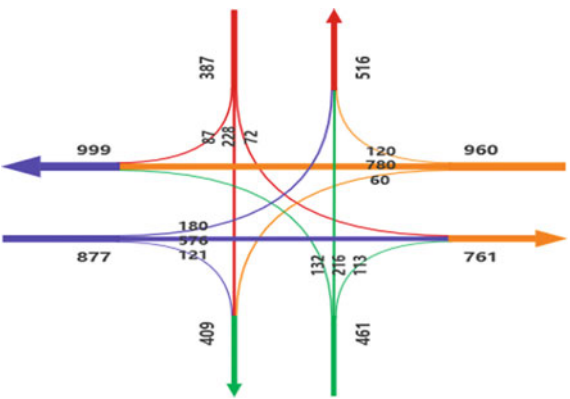


Fig. 5 Vehicle flow of the intersection of Hunnan Middle Road and Langyun Street (pcu/h)

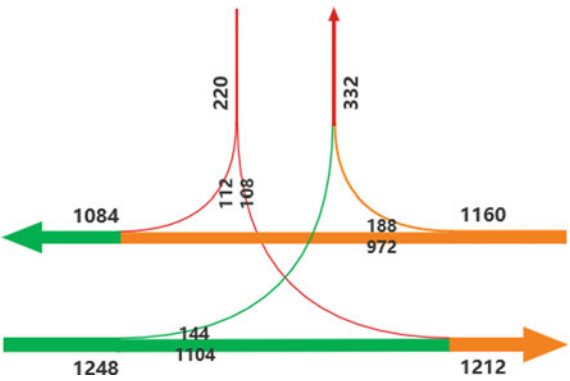


Table 2 Influence of traffic flow ratio on single factor

Bypass flow ratio		0.3:1	0.4:1	0.5:1	0.6:1	0.7:1
Delays/s	Closed	23.7	26.0	28.5	39.7	56.3
	Unclosed	35.2	30.7	28.6	29.7	29.1
Time of stops	Closed	0.34	0.37	0.35	0.29	0.34
	Unclosed	0.41	0.38	0.40	0.37	0.29

4.2 Result Analysis of Bypass Flow Ratio

In the case of other conditions unchanged, set bypass flow ratio (the ratio of bypass vehicles to non-bypass vehicles at investigated intersection) to 0.3:1, 0.4:1, 0.5:1, 0.6:1, and 0.7:1, and use the simulation software to simulate, then analyze the results, compares, and analyses the delay per capita with stops, respectively, to get Table 2.

If there is too much bypass traffic, the traffic efficiency of the whole area will be reduced. Through the single-factor analysis of the impact of its adaptability, when A value of 0.5 or so, closed and non-closed to the overall efficiency of traffic impact is not big, but when the ratio is higher than this value, the closed intersection will produce very big effect on the transportation efficiency of the whole area, delay will increase greatly, especially the number of cars than in the more than 0.6:1, delay level rise directly to the 40 s, and with the increase of the traffic delay time increases faster and faster. The service level of the entire intersection group is greatly reduced.

The results obtained from the simulation are different from those in the ideal state, not under the closed curve is flat, and not “two head high middle and low” situation, probably because in investigating the intersection is not closed, not as two round intersection survey of intersection traffic than changing, dilute the changes caused by investigating intersection delay.

4.3 Result Analysis of Left-Turn Capacity and Traffic Flow

If after adjustment of the signal timing, the number of bypass vehicles is still higher than the difference between the capacity of the bypass vehicle passing through the intersection left and the number of left-turned vehicles. Then, if the vehicles on the left-turning phase of the bypass intersection are congested, they cannot all pass. This situation is not allowed. Therefore, set the bypass traffic volume to Q_1 , the traffic volume of the bypass road before the closure is Q_2 , and the capacity of the import road is c . Then when $Q_1 > c - Q_2$, the bypass vehicle will cause congestion at the bypass intersection, so the intersection is not suitable for closure. When $Q_1 < c - Q_2$, the bypass vehicle can smoothly pass through the intersection. However, due to the change in traffic flow, the traffic timing is adjusted. Therefore, the overall traffic efficiency needs to be judged.

5 Conclusions

Based on the actual intersection data, this paper studies the applicability of the intersection of trams along the intersection, and can draw the following conclusions:

- (1) The aspect of bypass traffic flow. Traffic flow ratio cannot be too high, when the ratio is less than 0.5:1, suitable for closure, when the ratio is greater than 0.6:1, cannot be closed.
- (2) Passenger capacity of trams. When the intersection is closed, vehicles with trams and trams running in the direction of benefit from the transportation efficiency are the largest. Therefore, the increase in the number of vehicles can become traffic efficiency after the intersection is closed. The more passengers on the tram, the more suitable it is for closure.
- (3) Traffic conditions at surrounding intersections. Once the intersection is closed, many vehicles will detour and these vehicles will have a certain impact on the surrounding intersections. The first two intersections adjacent to the trams are the first ones. The adjacent intersections, especially the two intersections, are an important indicator of flow saturation.
- (4) Distribution of important departments and units around the intersection: If there are important public service units such as hospitals, schools, and gas stations, the intersections are often not suitable for closure. Once closed, the operating efficiency of vehicles arriving at these departmental units will be greatly affected.

Acknowledgements This work is supported by the Postdoctoral Science Foundation of China (No. 2016M601373). Research Topic of Economic and Social Development in Liaoning, China (2018lslktqn-058). Basic Scientific Research Projects of Higher Education, Liaoning, China (L17JBKY04). Philosophy and Social Science Planning Project of Shenyang, China (2017012). The Project of Shenyang Social Sciences, China (SYSK2018-08-09, SYSK2018-08-10).

References

1. Currie, G., Shalaby, A.: Active transit signal priority for streetcars: experience in Melbourne, Australia, and Toronto, Canada. *Transp. Res. Rec. J. Transp. Res. Board* **2042**(I), 41–49 (2008)
2. Skabardonis, A.: Control Strategies for Transit Priority [DB/CD]. TRB, National Research Council, Washington DC (2000)
3. Luiz, A.K., Werner, K.J.: Simultaneous control of traffic lights and bus departure for priority operation. *Transp. Res. Part C* **18**, 288 (2010)
4. Yu, B., Yang, Z.: Real-time partway deadheading strategy based on transit service reliability assessment. *Transp. Res. Part A: Policy Pract.* **46**(8), 1265 (2012)
5. Wu, J., Ghosal, D., Zhang, M., et al.: Delay-based traffic signal control for throughput optimality and fairness at an isolated intersection. *IEEE Trans. Veh. Technol.* **PP**(99), 1–1
6. Liu, L.L., Li, J.C.: Research on signal priority control strategy of modern tramway intersection based on VISSIM. *Highw. Motorcars* **6**, 56–58 (2014)
7. Kontos, P.C., Miller, K.L., Mitchell, G.J., et al.: Dementia care at the intersection of regulation and reflexivity: a critical realist perspective. *J. Gerontol. Series B: Psychol. Sci. Social Sci.* **66**(1) (2010)

8. Ahn, W., Lee, D.Y.: Real-time resolution of self-intersection in dynamic cylindrical free-form deformation. *IEEE Trans. Vis. Comput. Graph.* **17**(4) (2011)
9. Shi, J.G., Sun, Y.S., Schonfeld, P., Qi, J.: Joint optimization of tram timetables and signal timing adjustments at intersections. *Transp. Res. Part C*, 839 (2017)

Vehicle Scheduling Optimization of Urban Distribution Considering Traffic Control



Qianqian Shao, Huifei Li, Yunfeng Zhang and Feng Guan

Abstract In order to alleviate the internal traffic situation, the first- and second-tier cities have introduced traffic control measures to restrict the driving roads and period of freight vehicles, which makes the operation of distribution enterprises more difficult. Based on the recursive matrix of vehicle arrival time under soft time window, a distribution path optimization model considering urban freight vehicle traffic control is constructed, and a hybrid genetic algorithm based on “chain coding” and “sub-chain compensation” crossover operator is developed to solve the problem. The changing trend of total distribution cost under different control scenarios and the comparison of convergence quality under different sub-chain size thresholds show that the proposed model and algorithm are effective and practical.

Keywords City distribution · Restrictive regulation · Vehicle routing · Hybrid genetic algorithm

1 Introduction

Urban distribution accelerates the circulation of local commodities, drives urban economic development, and improves the convenience of people's lives. At the same time, it also brings external environmental problems such as environmental pollution, traffic congestion, and energy consumption. In response to this, various government departments have formulated corresponding to urban freight traffic control measures, hoping to limit the negative impact of distribution to people by limiting the driving time and road sections of freight vehicles within the city. The government's management measures have had a major impact on distribution companies: enterprises must not only complete the distribution tasks with high efficiency and low cost, but also

Q. Shao (✉) · H. Li · Y. Zhang · F. Guan
School of Transportation Engineering,
Shenyang Jianzhu University, Shenyang 110168, China
e-mail: 308977035@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_32

consider a series of practical factors such as prohibited road sections and restricted time limits.

Considering the distribution and optimization of urban freight vehicle traffic control is an extension of the vehicle routing problem (VRP), which is an NP-hard problem. Lang and Hu [1] used a hybrid genetic algorithm to solve the logistics distribution path optimization problem; Ombuki, et al. [2] combined Pareto method with genetic algorithm and verified the effectiveness of the algorithm through Solomon's classical example. In terms of cargo control, George, et al. [3] established an evaluation model for freight policy to verify the effectiveness of cargo control policy implementation; Quak and Koster [4] applied the vehicle path calculation software Shortrec 7.0 to estimate the distribution route to analyze the cargo control policy, in order to analyze the influence of logistics decision-making on the operation cost and environment of distribution companies; Wang, et al. [5] assessed the current urban freight management policies and measures based on the investigation and analysis of the rules of freight transportation in the domestic commercial center area. Hu, et al. [6] conducted a comparative analysis of the comprehensive benefits of urban distribution optimization under different city freight traffic control scenarios, but did not consider the customer's demand for delivery time windows. Huang [7] investigated in detail the traffic control policies of Beijing local licenses and foreign license vehicles, analyzed the logistics efficiency of various control policies, and proposed suggestions for improving the traffic control policies of Beijing freight vehicles. According to the influence of urban traffic restrictions and customer soft time window requirements on express delivery business, Liu, et al. [8] combined the traffic penalty cost and time penalty cost with the VRP problem, and proposed and verified the VRPTRSTW problem. In order to optimize the urban logistics distribution under the conditions of traffic control time, Zhe, et al. [9] comprehensively considers the distribution cost, environmental cost and delivery time of the distribution enterprise, and constructs a two-tier planning model.

In general, from the perspective of distribution enterprises, there are few studies that consider the problem of distribution optimization of urban freight traffic control policies.

Starting from the actual situation, based on the recursive matrix of vehicle arrival time under soft time window, this paper constructs a distribution route optimization model that considers the traffic control of urban freight vehicles, and develops a crossover operator based on "chain coding" and "sub-chain compensation". The hybrid genetic algorithm is used to compare and analyze the distribution optimization schemes under different control scenarios with examples to provide the dispatching decision support for the optimization of distribution in the forbidden road sections and regulatory periods.

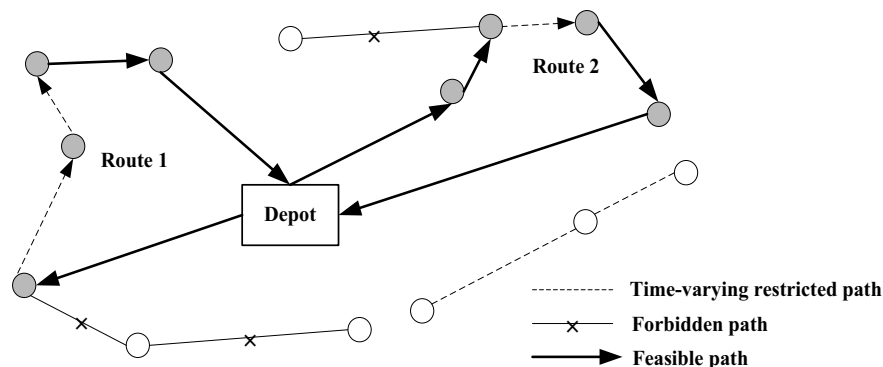


Fig. 1 The routing solution for VRPTW under traffic regulations

2 Problem Description and Modeling Conditions

Considering the distribution model of urban freight traffic control can be seen as an extension of the vehicle scheduling problem with time window constraints. Cargo delivery from multiple delivery vehicles to multiple demand points from one distribution center. Under the conditions positions, demand and time window constraints, weight per vehicle, and forbidden road sections, and the lower limit of the control period, we rationally arrange the route of the vehicle to minimize the total cost and meet the following assumptions:

- (1) The sum of demand for each distribution route does not exceed the carrying capacity of the vehicle.
- (2) The demand for each point can only be served once by one vehicle.
- (3) Each vehicle must start from the distribution center and return to it after completing the delivery.
- (4) The forbidden roads are not accessible. Vehicles are only allowed to enter the restricted-period roads during the regulatory period (Fig. 1).

3 Scheduling Optimization Model and Algorithm of Urban Distribution Considering Traffic Control

3.1 Variables and Parameters

Delivery directed graph $G = (N, M)$, the network characteristics are as follows: m_{ij} represents the arc from the demand point i to j ; the set of all arcs is represented by M ; the vertex 0 represents the distribution center; the vertex $1, 2, \dots, n$ represents the demand point; the vertex $0, 1, 2, \dots, n$ are recorded as a set

N , $N = \{0, 1, 2, \dots, n\}$; and the vertex $1, 2, \dots, n$ are recorded as a set C , $C = \{1, 2, \dots, n\}$; m_{ij} corresponds to the distance value d_{ij} and the time value t_{ij} , which t_{ij} is the running time from demand point i to j ; the demand point i corresponds to the demand q_i , s_i represents the service time; $V = \{1, 2, \dots, k, \dots, K\}$ is the vehicle collection, vehicle collection k corresponds to vehicle maximum load capacity Q_k ; E represents prohibited road collections, there is $M \cap E = E$; in case of morning and evening peak limit, $[R_{b1}, R_{e1}]$ is the morning peak traffic control time, $[R_{b2}, R_{e2}]$ is the evening peak traffic control time, R is a set of restricted roads that are subject to regulated time, with $M \cap R = R$, $R \cap E = \emptyset$. In addition, we usually use α to indicate the variable cost of the vehicle's unit transport distance, β is the fixed cost of the unit vehicle in the objective function. Here, d_{ij} , t_{ij} , q_i , s_i , Q_k , R_b , R_e , α , β are nonnegative, respectively.

In this model, there is decision variable x_{ijk} , for arcs m_{ij} and vehicle k ,

$$x_{ijk} = \begin{cases} 1, & \text{Vehicle } k \text{ from customer } i \text{ to customer } j \\ 0, & \text{Otherwise} \end{cases}$$

The total number of customers visited by the vehicle k is n_k , there is $n_k = \sum_{i \in C} \sum_{j \in N} x_{ijk}$. Collection $N_k = \{0, 1, 2, \dots, n_k\}$; the number l of the customer visited by vehicle k is O_{lk} , $O_{lk} = \sum_{j \in N} j x_{O_{(l-1)k}jk}$; vehicle k arrives at l th customer's time is TB_{lk} , the leave time is TE_{lk} , there is $TE_{lk} = TB_{lk} + s_{O_{lk}}$, TB_{lk} and TE_{lk} are recursive sequence. The time at which the vehicle k leaves the distribution center is denoted by st_k . When $m_{ij} \in R$, the expression for the time TB_{lk} to reach the first customer is as follows:

$$TB_{lk} = st_k + t_{0O_{lk}} + \begin{cases} 0 & st_k < R_{b1} \text{ or } R_{e1} < st_k < R \text{ or } st_k > R_{e2} \\ R_{e1} - st_k & R_{b1} < st_k < R_{e1} \\ R_{e2} - st_k & R_{b2} < st_k < R_{e2} \end{cases}$$

Then

$$TB_{lk} = TE_{(l-1)k} + t_{O_{(l-1)k}O_{lk}} + \begin{cases} 0 & TE_{(l-1)k} < R_{b1} \text{ or } R_{e1} < TE_{(l-1)k} < R_{b2} \text{ or } TE_{(l-1)k} > R_{e2} \\ R_{e1} - TE_{(l-1)k} & R_{b1} < TE_{(l-1)k} < R_{e1} \\ R_{e2} - TE_{(l-1)k} & R_{b2} < TE_{(l-1)k} < R_{e2} \end{cases}$$

When $m_{ij} \notin R$, $TB_{lk} = TB_{(l-1)k} + t_{O_{(l-1)k}O_{lk}} + s_{O_{(l-1)k}}$, this sequence establishes a one-to-one correspondence between variables i and variables (l, k) , so $TB_i = TB_{lk}$. The demand point i corresponds to the time window $[a_i, b_i]$, which a_i represents the earliest delivery time allowed by i . And b_i represents the latest delivery time. If the time TB_i to arrive at the demand point i is earlier than a_i or later than b_i , the penalty should be cost, p, q is the unit time penalty coefficient that the vehicle arrives early or late. The penalty cost function is expressed as follows:

$$P_i(TB_i) = \begin{cases} p(a_i - TB_i) & TB_i < a_i \\ 0 & a_i \leq TB_i \leq b_i \\ q(TB_i - b_i) & b_i < TB_i \end{cases}$$

3.2 Model Construction

$$\min z = \alpha \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} d_{ij} x_{ijk} + \beta \sum_{k \in V} \sum_{j \in C} x_{0jk} + \sum_{i \in C} P_i(TB_i) \quad (1)$$

$$\text{s.t.} \quad \sum_{k \in V} \sum_{j \in C} x_{ijk} = 1 \quad \forall i \in N \quad (2)$$

$$\sum_{i \in N} x_{ihk} - \sum_{j \in N} x_{hjk} = 0 \quad \forall h \in C, \forall k \in V \quad (3)$$

$$\sum_{j \in C} x_{0jk} = \sum_{j \in C} x_{j0k} \quad \forall k \in V \quad (4)$$

$$\prod_{l \in N_k} x_{O_k O_{(l+1)k}} = \sum_{j \in C} x_{0jk} \quad \forall k \in V \quad (5)$$

$$\sum_{k \in V} x_{ijk} = 0 \quad \forall m_{ij} \in E \quad (6)$$

$$\sum_{k \in V} x_{ijk} = 0 \quad \forall m_{ij} \in R \quad (7)$$

$$\sum_{i \in C} q_i \sum_{j \in N} x_{ijk} \leq Q_k \quad \forall k \in V \quad (8)$$

$$x_{ijk} \in \{0, 1\} \quad \forall i \in N, \forall j \in N, \forall k \in V \quad (9)$$

$$TB_i \geq 0 \quad \forall i \in N \quad (10)$$

The objective function (1) indicates that the total cost of distribution is minimal, including the vehicle distance cost, the number of vehicle activation costs, and the time delay or penalty cost ahead. Equation (2) indicates that each demand point can only be serviced once by a car; Eqs. (3) and (4) ensure that each vehicle starts from the distribution center and returns to the distribution center after completing a series of distribution tasks; Eq. (5) avoids the generation of sub-loops; Eq. (6) shows that no-forbidden roads are not traversable; Eq. (7) shows that restricted-route road sections are not traversable during the regulatory period; Eq. (8) shows the demand for each demand point on each path. The sum does not exceed the maximum load of the car; Eq. (9) is the integer constraint of the decision variable; Eq. (10) is the nonnegative constraint of the arrival time.

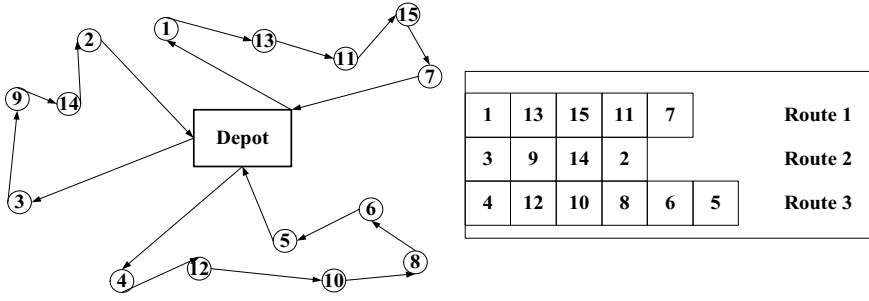


Fig. 2 The corresponding between chromosome and feasible routes

3.3 A Hybrid Genetic Algorithm Considering Urban Transport Control Constraints

Combining the structural characteristics of the feasible solution of the abovementioned optimization model and taking full account of the influence of traffic control-related constraints on genetic operators, a hybrid genetic algorithm based on “chain coding” was developed. The algorithm design mainly includes coding, generation of the initial population, determination of fitness function and design of genetic operator details as follows:

(1) Coding

In the network diagram model, the delivery path for each freight car is embodied as a chain of n_k distribution customers and $n_k + 1$ arcs. The information of the arcs can be derived from the point information in the distance (time) matrix. The encoding is set to a string of real numbers representing the order of delivery and the value is the number of the shipping customer. Several strings of real numbers form a chromosome. The length and quantity of the distribution chain are not fixed, the upper limit is n . Figure 2 (left) shows a distribution plan with three delivery routes, Fig. 2 (right) is its corresponding code.

(2) Initial population

In order to ensure the randomness of the initial solution in the feasible domain, a heuristic rule for generating the initial population was designed. Let A_k denotes the k -th distribution chain, A_{kw} denotes the w -th customer that constitutes the chain, set customer collection $B = C$ and collection M_w , and the latter denotes the reachable customers that start from w -th customer (distribution center) in collection B (including time limit line waiting for up to customers), set a variable $k = 1$, the rules can be described as:

Step 1: If $B = \emptyset$, the rule ends, otherwise Step 2;

Step 2: Select customer i randomly in the set M_0 with probability $\left| \frac{1}{M_0} \right|$ and remove them from the set B . Let $w = 1, A_{kw} = i$; ① Customer i is randomly

selected with a probability of $\left| \frac{1}{M_w} \right|$ in the set M_w and removed from the set B , let $w = w + 1, A_{kw} = i$, execute ②; ② If $\sum_{l=1}^w q_{ol_k} > Q_k$, then a natural number p from 1 to $w - 1$ is generated with probability $\frac{1}{w-1}$, and customer $A_{kw'} (w' = p + 1, p + 2, \dots w.)$ is added to set B , let $A_{kw'} = 0 (w' = p + 1, p + 2, \dots w.)$, $k = k + 1$, Perform Step 1; otherwise, execute (1); Vector A_k is the delivery path and k is the number of scheduled vehicles

(3) Fitness function

Since the value of the objective function is the sum of the utility values, the difference between the merits and the weaknesses is not significant, so the fitness function selects the reciprocal of the objective function value, i.e., $f = 1/z$.

(4) Select operator

Use a combination of best individual retention and roulette selection strategies.

(5) Crossover operation

In order to overcome the influence of traffic route control and load constraints on the operator, in this paper, it simulates the block hypothesis principle and designs a “sub-chain compensation” crossover operator to decompose each distribution chain into “sub-chains” and perform other operations. It can be guaranteed that generating children are feasible solutions, thus avoiding the mode effect caused by the adjustment of solutions without theoretical support. Let parent individuals be A and B , and the respective distribution chains are A_k and B_k . The specific operation methods are:

① Set parameter ε as the sub-chain size (length) threshold; ② Dissolve the distribution chain with a length greater than ε in parent A . The decomposition step is to generate a random number u from 1 to ε , and decompose the sub-chain of length u from A_k . Repeat this process until the remaining length of A_k is less than ε , then can get δ sub-chains, from which to choose $\left[\frac{\delta}{2} \right]$ sub-chains to compose set SA ; ③ In every distribution chain B_k of parent B , the customers contained in SA are eliminated and several sub-chains are obtained. The sub-chain whose size exceeds α is decomposed by applying the method in ② to obtain sub-chain set SB ; ④ Combining SA and SB into CB , which contains the customer at both ends of the sub-chain constitute a set CC , and the construction of a new individual rule is as follows:

Let C_k denote the $k - th$ distribution chain of the child, C_{kw} denotes the $w - th$ chain that constitutes the chain, and let M_w denote the reachable customer (including time limit line) in the set CC starting from the end customer (distribution center) of the $w - th$ child chain. Let us set the variable $k = 1$. The rule can be described as:

Step 1: If $CB = \emptyset$, the rule ends; Otherwise Step 2;

Step 2: In the set M_0 , the customer i in CC and its sub-chain I are randomly selected with probability $\left| \frac{1}{M_0} \right|$, I is removed from CB , and the endpoints i and i' of I are removed from CC . Let $w = 1, C_{kw} = I$;

① In the set M_w , the customer i in CC and its sub-chain I are randomly selected with probability $\left| \frac{1}{M_w} \right|$, I is removed from CB , and the endpoints i and i' of I are removed from CC . Let $w = w + 1$, $C_{kw} = I$, execute ②; ② If C_k breaks the load-carrying constraint, a natural number p from 1 to $w - 1$ is generated with probability $\frac{1}{w-1}$, and the child chain $C_{kw'}(w' = p + 1, p + 2, \dots w.)$. Join CB , the endpoint joins CC and let $C_{kw'} = 0(w' = p + 1, p + 2, \dots w.)$, $k = k + 1$, Perform Step 1; otherwise, execute ①; The vector C_k is the delivery path and k is the number of scheduled vehicles.

(6) Variation operation.

Select individual G with mutation probability P_m . The steps for mutation are:

- Step 1:** Customer i and customer j are randomly selected on the chromosome, and a new chromosome G' is formed at the exchange position. If G' is an infeasible solution (it does not meet the prohibited road segment constraint), it is discarded, Step 1 is performed, otherwise Step 2 is performed;
- Step 2:** Compare the fitness function values of G and G' , if $f(G) > f(G')$, then execute Step 1, otherwise accept G' , the mutation ends

4 Numerical Experiments and Result Analysis

The numerical experiment is designed to compare the experimental examples under different control scenarios. Benchmark Problems is a VRPTW standard test question bank designed by Solomon in 1953. This article selects C109 and R209, which are more suitable for real-time scenarios, as the example data. Randomly select 10, 20, and 30% of the inter-customer routes as forbidden road sections. Rules for selecting routes under regulation period: Select customer points randomly, 20% of the inter-customer path within the radial r , the radius of this point is the restricted road section.

Experiment 1: The regulatory period is the morning peak 7:00–9:00 pm peak 17:00–19:00, and the calculation results of the case C109 and R209 corresponding to different forbidden road sections are shown in Table 1.

Experiment 2: Select 20% of forbidden road sections. Table 2 shows the calculation results of calculation cases C109 and R209 corresponding to different regulatory periods

From the calculation results, it can be seen that the forbidden road section has a greater influence on the total cost of distribution than the control period. With the increase of the scope of control, the total cost of distribution of case C109 is larger than that of case R209, which is almost an exponential increase. This shows that the relatively uniform distribution of customer aggregation is more sensitive to the regulatory policy. In order to cope with urban freight traffic control, distribution

Table 1 Comparison of calculation results for different number of forbidden road sections for examples C109 and R209

Examples	C109		R209	
	Total cost	Increment (%)	Total cost	Increment (%)
Unimpeded roads	3184.82		5150.16	
10% of prohibited	3431.67	7.8	5505.40	6.9
20% of prohibited	4137.35	29.9	6089.312	18.2
30% of prohibited	5784.95	81.6	8014.20	55.6

Table 2 Computation results of calculation examples C109 and R209 corresponding to different regulatory periods

Examples	C109		R209	
	Total cost	Increment (%)	Total cost	Increment (%)
Uncontrolled period	3609.08		5904.38	
Morning 8:00–9:00	3820.46	5.9	5953.34	0.8
Evening 18:00–19:00				
Morning 7:00–9:00	4137.35	14.6	6089.31	3.1
Evening 17:00–19:00				
Morning 6:00–9:00	4396.24	21.8	6251.27	5.9
Evening 16:00–19:00				

Table 3 Basic experimental data

Number of clients	50
Ratio of prohibited road (%)	10, 20, 30
Control period	Morning peak 8:00–9:00 Evening peak 18:00–19:00 Morning peak 7:00–9:00 Evening peak 17:00–19:00 Morning peak 6:00–9:00 Evening peak 16:00–19:00
Proportion of restricted roads during the control period	Random customer point radius is 20% customer path within r range
Other parameters	$\alpha = 8, \beta = 60, p = 0.5, q = 1.5, \varepsilon = 5$

companies should develop a route optimization plan in conjunction with the distribution of customer points. When customers tend to aggregate and distribute, they can establish regional distribution centers or negotiate to change delivery time windows (Table 3).

5 Conclusions and Outlook

By establishing a distribution optimization model that takes into account the traffic control of urban freight vehicles, a hybrid genetic algorithm based on “chain coding” and “sub-chain compensation” crossover operators is developed to calculate numerical experiments. The results show that different control policies apply to distribution plans. The policy of forbidden road sections has a greater impact on distribution companies than the control period, and the calculation of customer aggregation is faster than the even distribution of total cost. Distribution companies can use this model and algorithm to calculate the optimal distribution route that takes into account the traffic control of vehicles.

Further research can quantify the advantages and disadvantages of traffic control policies for vehicles from perspectives such as the opportunity cost of distribution companies and the social responsibility of distribution companies.

Acknowledgements This work was supported by the China Science and Technology Project of Ministry of Housing and Urban–Rural development (Project Number: 2017-R2-039), the China Society of Logistics Research Project (Project Number: 2017CSLKT3-172), and the Scientific Research Innovation Project of Shenyang Jianzhu University (Project Number: CXPY2017006).

References

1. Lang, M.X., Hu, S.J.: Study on the optimization of physical distribution routing problem by using hybrid genetic algorithm. *Chin. J. Manag. Sci.* **10**(05), 52–57 (2002). (in Chinese)
2. Ombuki, B., Ross, B.J., Hanshar, F.: Multi-objective genetic algorithms for vehicle routing problem with time windows. *Appl. Intell.* **24**(1), 17–30 (2006)
3. George, Y., John, G., Constantinos, A.: Effects of urban delivery restrictions on traffic movements. *Transp. Plan. Technol.* **29**(4), 295–311 (2006)
4. Quak, H.J., Koster, M.B.M.D.: Delivering goods in urban areas how to deal with urban policy restrictions and the environment. *Transp. Sci.* **43**(2), 211–227 (2009)
5. Wang, X.K., Yang, D.Y., Hu, X.W.: Freight transportation characteristics and assessment of policy in urban central commercial district. *J. Transp. Syst. Eng. Inf. Technol.* **6**(06), 130–135 (2006)
6. Hu, Y.C., Shen, J.S., Huang, A.L.: Multi-objective optimization for city distribution under urban freight restriction. *J. Transp. Syst. Eng. Inf. Technol.* **12**(06), 119–125 (2012). (in Chinese)
7. Huang, F.Z.: Research on the Influence of Beijing Truck Traffic Control Policy on Logistics Efficiency. *Logistics Engineering and Management.* **39**(281), 116–117 (2017)
8. Liu, Wu, Wu, J., Hu, H.: Vehicle routing problem under traffic restriction and soft time window and improvement of ant colony algorithm. *Logist. Technol.* **35**(360), 98–103 (2016)
9. Zhe, X.F., Pu, Y.M., Liang, F., Qin, W.W.: Two-tier planning model for urban logistics distribution optimization under traffic control time limit. *J. Highw. Transp.* **31**(232), 149–156 (2014)

The Warrant of Slip Lane at Single-Lane Roundabout



Yuzhou Duan, Hui Li and Yabo Song

Abstract In the NCHRP Report 672, a brief introduction is given to the right-turn slip lane at a roundabout, and the lane capacity is derived. On this basis, we considered the setting mode of two types of slip lanes, such as free driving (Free Flow, FF) and yield driving (Yield, Y). The traffic flow characteristics were analyzed under the conditions of its setting. A large number of scenes were designed by VISSIM simulation, taking the average vehicle delay as the index. A comparison of operation was conducted between with and without slip lane, under the combination of different traffic flows and right-turn ratio. The critical conditions of different setting modes were preliminarily discussed. Considering the impact of slow traffic such as pedestrians and bicycles, the next research plan was put forward.

Keywords Roundabout · Slip lane · Average vehicle delay · VISSIM · Simulation · Traffic flow characteristics

1 Introduction

As special unsignalized facilities, roundabouts have been paid more and more attentions by the scholars in recent years, because of its advantages in efficiency and safety under low-to-medium volumes. Some countries, such as the United Kingdom, Australia, the United States, and others, have carried out some studies to roundabouts for years. They formed some guiding reports and manuals, such as NCHRP Report 572 [1], NCHRP Report 672 [2], and so on, which provided an effective reference to the engineers.

For the right-hand driving rule, the through and left-turn vehicles will affect the efficiency of the whole roundabout directly. Therefore, a lot of researches have been carrying out to analyze the through and left-turn flow in the inner roundabout, from

Y. Duan (✉) · H. Li · Y. Song
School of Civil Engineering and Architecture, Henan University of Technology,
Zhengzhou, China
e-mail: lbtj2010@163.com

the aspects of capacity, delay, conflicts, and so on. And, the right turns are usually not given special considerations, which may be ignored in some studies. When the right-turn ratio is low, the right turns will use the outer circulating lane to enter the roundabout, and then exit the roundabout after driving a short distance. So, the effect of right turns is small compared to the other two driving directions under these circumstances. However, when the right-turn rate is large, according to the gap-acceptance theory, the right-turning vehicle needs to enter the roundabout under the headway of circulating vehicle greater than the critical gap. So, it has a certain effect on the operation efficiency of the whole roundabout. In such a situation, we need to design the channelization of the entrance, and introduce the right-turn slip lane (also known as the right-turn bypass lane).

NCHRP Report 672 also considered the slip lane when calculating the capacity of a roundabout, and gave the recommended equation for the capacity of the slip lane [2]. With SIDRA and VISSIM software, Al-Ghandour [3–6] made a comprehensive analysis of the traffic benefits of a roundabout with the slip lanes, and predicted the conflict points with the SSAM software. On this basis, combined with the convenience of the VISSIM simulation software in the study of the roundabout [7–9], we will make a more in-depth study on the performance of two different slip lanes for free driving and yield driving.

On the basis of the above research, this paper analyzes the traffic flow characteristics under different setting conditions of slip lane. And, the detailed and comprehensive experimental designs are carried out to different types of slip lanes by VISSIM. Taking the average vehicle delay as the evaluation index, the critical condition of the warrant of the slip lane is obtained, which will propose some certain reference for engineers and technicians.

2 Slip Lane Setting Type and Traffic Flow Characteristics

2.1 Type of Slip Lane at Roundabout

Referring to the NCHRP Report 672 [2] published by TRB in the United States, there are two types of slip lanes, according to whether there is a conflict between the right turns and the exit vehicles: free driving (Free Flow, FF) and yield driving (Yield, Y). For short, FF-type and Y-type are used here, and the schematic diagram is shown in Fig. 1.

From Fig. 1a, it can be seen that for the Y-type slip lane, the entrance is channelized. The effect of the right turns on a roundabout is reduced by separating the right turns from the others in space. And, the FF-type slip lane is treated with a more thorough measure in Fig. 1b, which is not only at the entrance, along with the exit. An exclusive right-turn lane is set up, and the two lanes will be merged gradually at exit downstream. Because the FF-type takes up more road than the Y-type, the Y-type will get higher benefit under little difference in operation efficiency.

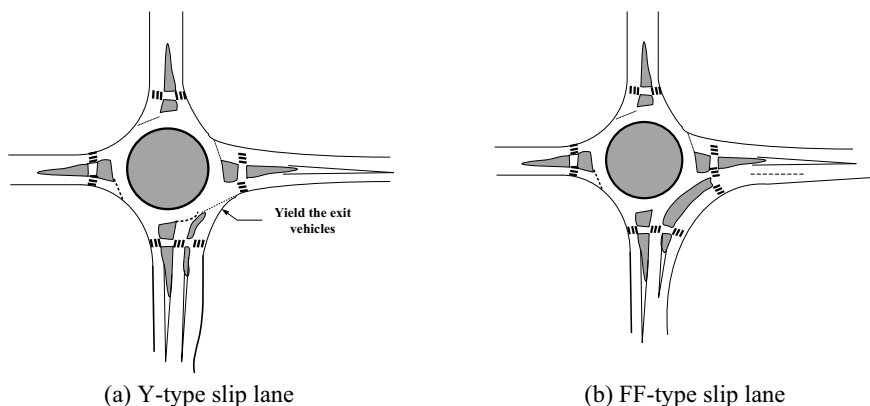


Fig. 1 A schematic diagram of two types of slip lanes

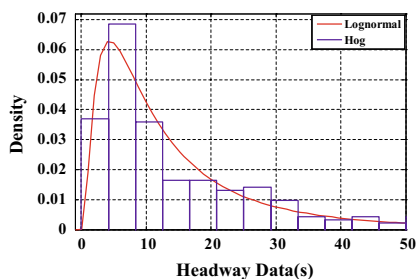
2.2 Traffic Flow Characteristics of Slip Lane

Compared with the common roundabout, the right turns with the slip lane will have different characteristics. For the two different types of slip lanes, the traffic flow characteristics are analyzed in this section.

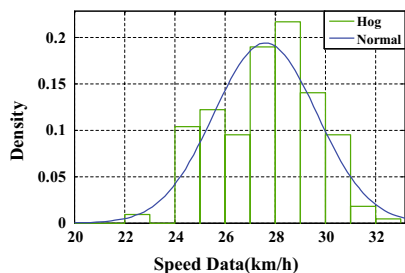
Y-type Slip Lane. Based on the repeatability characteristic of VISSIM [10], the simulation scenarios were built with different flow rates and different right-turn ratios. For the Y-type slip lane, the exit lane is equivalent to the major road, and the slip lane is equivalent to the minor road, due to the yielding. The theory of major–minor road under the unsignalized intersection could be used to study the running characteristics. The data collectors were set up at the conflict area at the exit lane and the slip lane, and then the traffic flow data are collected. For example, when the exit lane was 390 pcu/h, the traffic flow of south approach was 650 pcu/h, and the right-turn ratio was 55% (that is, the right-turn volume was 357.5 pcu/h), the headway and velocity distribution were shown in Fig. 2, respectively.

According to the gap-acceptance theory, when determining the headway distribution of the major road, the capacity of the minor could be obtained after determining the critical gap and the follow-up time. In NCHRP Report 672, the recommended capacity equation for slip lane was proposed [2], as shown in Eq. (1). The traffic capacity obtained was 765 pcu/h according to this. Based on the headway data in the simulation, the capacity of the slip lane was 1203 pcu/h, and the saturation of the traffic flow on the slip lane was 0.297. It can be seen that the Eq. (1) was based on the road conditions of the United States, resulting in a great difference between the two.

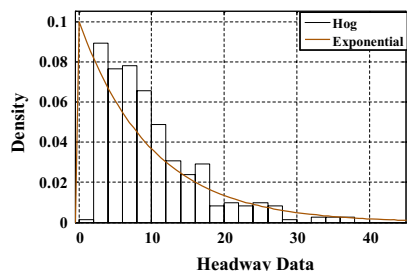
$$C_{bypass,pcu} = 1130e^{(-1.0 \times 10^{-3})v_{ex,pcu}} \quad (1)$$



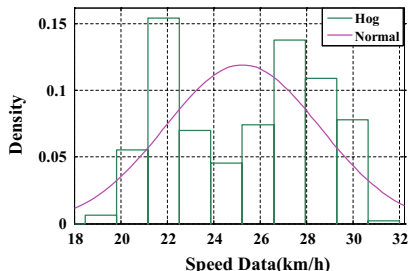
(a) Headway distribution of exit vehicles



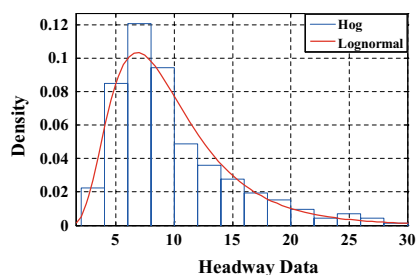
(b) Velocity distribution of exit vehicles



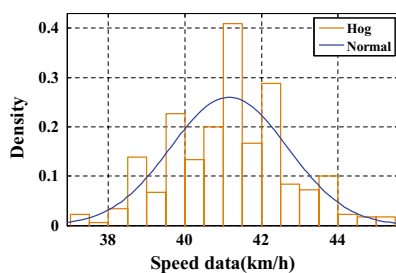
(c) Headway distribution of right turns



(d) Velocity distribution of right turns

Fig. 2 Headway and velocity distribution of Y-type slip lane

(a) Headway distribution of right turns



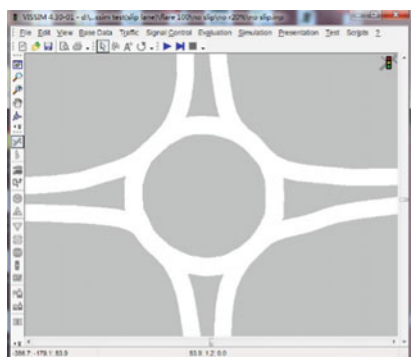
(b) Velocity distribution of right turns

Fig. 3 Headway and velocity distribution of FF-type slip lane

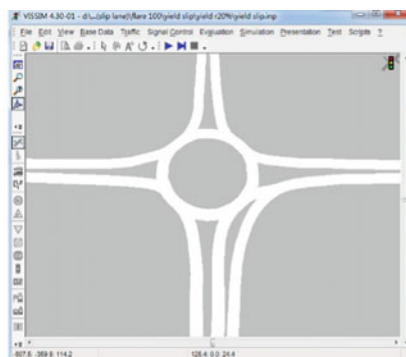
where $C_{bypass, pcu}$ is the capacity of the bypass lane, pcu/h, $v_{ex, pcu}$ is the traffic volume of exit lane, pcu/h.

FF-type Slip Lane. Similar to the Y-type slip lane, the headway and velocity distribution of the FF-type could also be obtained by using VISSIM simulation. And, the result under the same traffic ratio is shown in Fig. 3.

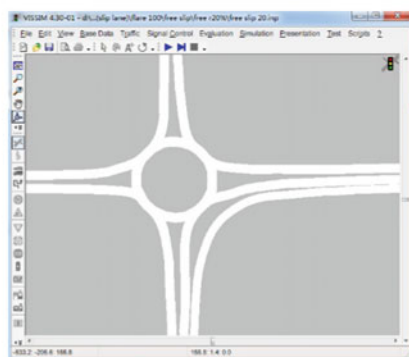
Because the FF-type set up an exclusive lane, its capacity was much larger than that of the Y-type. It is assumed that the saturated headway was 2.32 s, the traffic capacity was 1552 pcu/h, and the lane saturation was 0.23. Compared with Figs. 2 and 3, we could see that the speed of the FF-type is greater than that of the Y-type.



(a) Without Slip-lane



(b) Y-type slip lane



(c) FF-type slip lane

Fig. 4 Road network in VISSIM

3 Benefit Analysis

In order to investigate the traffic operation of different traffic volumes and different right-turn ratios, a lot of data were needed. We adopted the repeatability and flexibility of VISSIM simulation software, and obtained comprehensive traffic data by setting different simulation scenes.

3.1 Simulation Conditions

We analyzed the right-turn lane under three conditions: without slip lane, Y-type slip lane, and FF-type slip lane. The road network was constructed according to their geometric characteristics in VISSIM. The diameter of the inner island was 34.5 m, the width of the circulating lane was 4.5 m, and the road network is shown in Fig. 4.

Table 1 Right-turning flow change at south approach (pcu/h)

Right-turn ratio/ Approach volume	300	350	400	450	500	550	600	650	700	750	800
0.2	60	70	80	90	100	110	120	130	140	150	160
0.3	90	105	120	135	150	165	180	195	210	225	240
0.4	120	140	160	180	200	220	240	260	280	300	320
0.5	150	175	200	225	250	275	300	325	350	375	400
0.6	180	210	240	270	300	330	360	390	420	450	480
0.7	210	245	280	315	350	385	420	455	490	525	560

Note: the color represents the change in the flow level.

As the lane was the object of this paper, in addition to the changes in the flow of south approach, the flow of the other entrance was set to a fixed value (traffic flow was 300pcu/h, left: through: right = 3:4:3), and the right-turn volume of south approach is shown in Table 1.

In addition, in order to simulate the deceleration at the entrance, the desired speed and reduced area were combined, in which the desired speed before the entrance was (25, 30) (km/h), the exit was (48, 58) (km/h), and the reduced area at the turn was set to (25, 30) (km/h). The conflict area was used to set up the yield control. At the same time, in order to reduce the influence of randomness on the result, 20 groups of random seeds were set up in the simulation environment, and the average value was taken as the final results.

3.2 Results

In the above simulation environment, the delay of the right-turning vehicles is taken as the final evaluation index. The results of the three different slip lanes are shown in Fig. 5.

As can be seen from Fig. 5, by setting up the slip lane, the delay of the right turns was greatly reduced, especially the FF-type slip lane. Because the right turns had the exclusive right of way, and the driving behavior was unnecessary, such as stopping or yielding, which made the delay of the right turns reduce greatly.

In addition, the delay of the entire roundabout could be obtained. For example, when the south entry flow was 800 pcu/h and the right turn rate was 60%, the whole delay of the without slip lane, the Y-type slip lane and the FF-type slip lane was 18.2 s, 13.4 s, and 5.6 s, respectively. It could be seen that the setting of slip lane has played a positive role in reducing the whole delay.

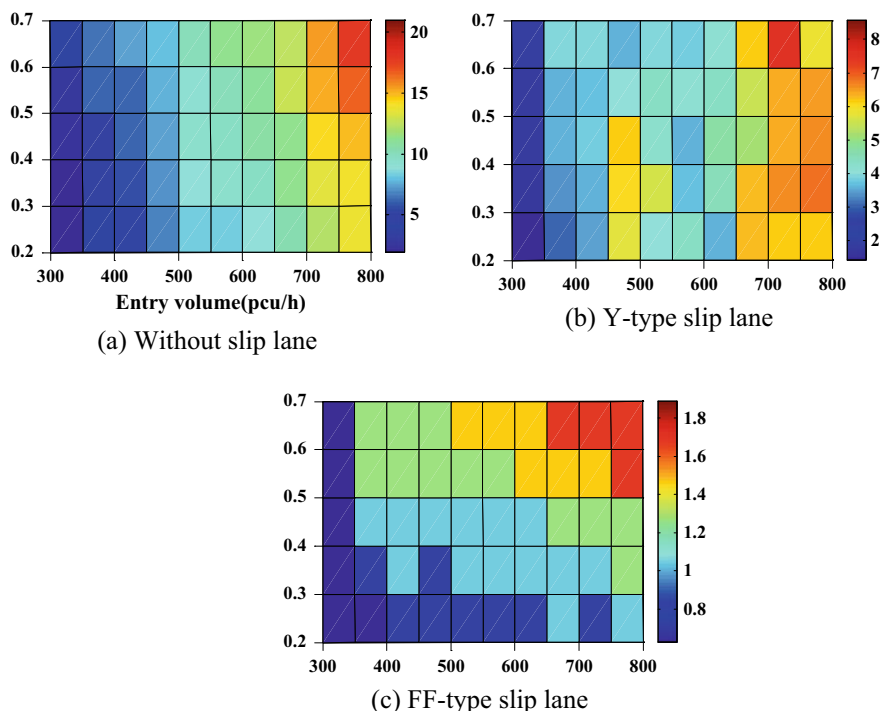


Fig. 5 The results of average vehicle delay

3.3 Critical Condition Analysis

According to the above analysis, when the total traffic flow rate and the right-turn rate were small, the roundabout without slip lane could achieve higher performance. With the increase of the right-turn volume, the delay gradually became larger. At this time, some measures should be taken to control and guide. The Y-type and FF-type slip lane were both effective control measures, which effectively reduced the delay of the right turns. Both of them, the FF-type slip lane made the right-turn vehicles get a higher level of service. However, from the geometric designs, the FF-type slip lane needs to set up the corresponding exclusive lane at the exit, which needs to spend a certain amount of road resources. The Y-type slip lane was also a kind of high-performance traffic facility, while ensuring the efficiency. Therefore, when the right-turn flow is low, it is recommended not to use the slip lane. When the flow reaches to a certain degree, the Y-type slip lane is used. Only when the Y-type slip lane could not bear the corresponding flow, it is suggested that the FF-type slip lane should be used.

Because the roundabout is a systematic system, the right turns will affect its performance in a certain range, but the performance of the entire needs to be judged

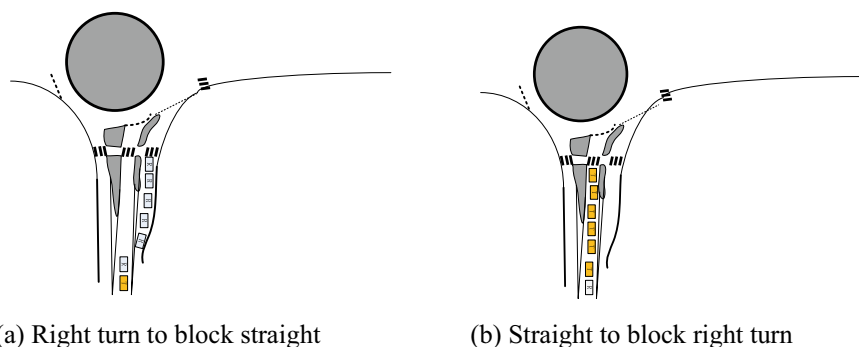


Fig. 6 Schematic diagram of congestion in Y-type slip Lane

according to the O-D distribution. So the setting of slip lane is put forward only to a certain extent. If the whole roundabout is concerned, the situation is more complex and will continue to be studied in the future.

3.4 Other Influencing Factors

In addition to the above traffic factors, some geometric conditions will also affect the operation of vehicles on the slip lane. For example, the channelization length at the entrance will affect its performance for Y-type slip lane. As shown in Fig. 6a, if more vehicles are exiting, the vehicles in the slip lane need to wait. However, when the length of the channelization section is short, the right-turn vehicles may block the other vehicles. Similarly, as shown in Fig. 6b, the through vehicles may also block the right-turn vehicles from entering the slip lane. It is necessary to analyze the situation from the microscopic perspective of probability theory, combined with the situation of downstream entry.

For FF-type slip lane, the setting length at the downstream exit is also needed to be discussed. In addition, there is no consideration of slow traffic facilities such as pedestrians and bicycles in this paper. When the flow is large, it will also have a greater impact on the motor vehicles.

4 Conclusions

Based on the VISSIM simulation, this paper made a preliminary discussion on the setting of the slip lane. The traffic flow characteristics of vehicles on the slip lane were analyzed. On this basis, a variety of scenes with different traffic volumes and right-turn ratios were designed. And the delay of without slip lane, Y-type slip lane,

and FF-type slip lane were analyzed in detail. Based on this, the critical condition of slip lane setting was proposed, which provides a certain understanding foundation for future research. Finally, combined with the impact of slow traffic such as pedestrians and bicycles, the problems that should be paid attention to in the next step are put forward.

Acknowledgements The work was supported by Key Scientific Research Projects of Higher Education in Henan Province (17A580006), High-level Research Fund Project in Henan University of Technology (31401062), and Science and Education Integration Project of Henan University of Technology (26400044).

References

1. Rodegerdts, L., Blogg, M., Wemple, E., et al.: NCHRP Report 572: Roundabouts in the United States. Transportation Research Board of the National Academies, Washington, D.C. (2007)
2. Rodegerdts, L., Bansen, J., Tiesler, C., et al.: NCHRP Report 672: Roundabouts: An Informational Guide, 2nd edn. Transportation Research Board of the National Academies, Washington, D.C. (2010)
3. Al-Ghandour, M.: Roundabout Slip Lanes: Performance and Safety Analysis. North Carolina State University, Raleigh (2011)
4. Al-Ghandour, M., et al.: Conflict models for single-lane roundabout slip lanes from microsimulation. *Transp. Res. Rec.* **2236**, 92–101 (2011)
5. Al-Ghandour, M.: Experimental Analysis in VISSIM of Single-Lane Roundabout Slip Lane Under Varying Bus Traffic Percentages. *Bridges 10*: 9780784413210-011 (2014)
6. Al-Ghandour, M., et al.: Delay analysis of single-lane roundabout with a slip lane under varying exit types, experimental balanced traffic volumes, and pedestrians, using microsimulation. *Transp. Res. Rec.* **2312**(1), 76–85 (2012)
7. Mauro, R., Marco, G.: Right-turn bypass lanes at roundabouts: Geometric schemes and functional analysis. *Mod. Appl. Sci.* **7**(1) (2012)
8. Bared, J.G., Afshar A.M.: Using simulation to plan capacity models by lane for two- and three-lane roundabouts. *Transportation Research Record*, pp. 8–15 (2009)
9. Trueblood, M., Dale, J.: Simulation roundabouts with VISSIM. Presented at 2nd Urban Street Symposium, Anaheim, Calif, July 28–30 (2003)
10. PTV AG: VISSIM User Manual 5.0. Karlsruhe, Germany (2007)

A New Method for the Minimum Concave Cost Transportation Problem in Smart Transportation



Chuan Li, Zhengtian Wu, Baochuan Fu, Chuangyin Dang and Jinjin Zheng

Abstract The minimum concave cost transportation problem is the benchmark problem in numerical computing and it has been used widely in the schedule of smart transportation. In this paper, a deterministic annealing neural network algorithm is proposed to solve the minimum concave cost transportation problem. The algorithm is derived from two neural network models and Lagrange-barrier functions. The Lagrange function is used to handle linear equality constraints and the barrier function is used to force the solution to move to the global or near-global optimal solution. The computer simulations on test problem are made and the results indicate that the proposed algorithm always generates global or near global optimal solutions.

Keywords Minimum concave cost transportation problem · Neural network algorithm · Neural network models · Lagrange-barrier functions

1 Introduction

The minimum concave cost transportation problem in this paper is about minimizing a concave quadratic function subject to transportation constraints, which is a well-known NP-hard problem. In order to compute a global optimal point of the minimum concave cost transportation problem, a number of traditional iterative algorithms

C. Li · J. Zheng

Department of Precision Machinery and Precision Instrumentation,
University of Science and Technology of China, Hefei, China

C. Li · C. Dang

Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon Tong, Hong Kong, China

Z. Wu (✉) · B. Fu

School of Electronic and Information Engineering, Suzhou University
of Science and Technology, Suzhou, China
e-mail: wzht8@mail.usts.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_34

363

have been developed in the literature such as Erickson et al. [1], Fang and Tsao [2], Gallo and Sodini [3], Pardalos and Rosen [4], and Zangwill [5]. A detailed survey of algorithms for the minimum concave cost transportation problem can be found in Guisewite [6].

Some new kinds of algorithms have been developed for optimization problems. Among them, neural network is one of the most efficient algorithms, which was originated in Hopfield and Tank [7] and has been applied to a variety of optimization problems. However, most problems tackled in the existing neural networks are linear programming problems or convex quadratic programming problems. Liu et al. [8] proposed a new recurrent neural network for linear 0–1 programming problems, Hu and Zhang [9] presented a neural network for convex quadratic programming problems, and Hu and Wang [10] proposed a dual neural network for a special class of convex quadratic programming problems. Xu et al. proposed neural nets based on the hybrid of Lagrange and transformation approaches [11–13]. Based on the integer programming, Wu [15–17] has developed a new method for airline disruption problems. In this paper, a novel neural network algorithm is proposed by integrating the annealing scheme and Lagrange-barrier function into neural network for concave quadratic transportation problems. The numerical simulations demonstrate the proposed neural network can converge to global optimal solutions.

The rest of this paper is organized as follows. We introduce the proposed neural network models and propose a neural network annealing algorithm in Sect. 2. To show that the proposed neural network annealing algorithm is effective and efficient, we conduct some numerical experiments in Sect. 3. We conclude the paper with some remarks in Sect. 4.

2 Deterministic Annealing and Neural Networks

The minimum concave cost transportation problem we consider to solve in this paper can be stated mathematically as follows: Find a global minimum point of

$$\begin{aligned}
 & \min \quad f(x) \\
 & \text{subject to} \quad \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m, \\
 & \quad \quad \quad \sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n, \\
 & \quad \quad \quad 0 \leq x_{ij}, \quad j = 1, \dots, p_i, \\
 & \quad \quad \quad 0 \leq x_{ij} \leq u_{ij}, \quad j = p_i, \dots, n, \\
 & \quad \quad \quad i = 1, 2, \dots, m,
 \end{aligned} \tag{1}$$

where $f : R^{m \times n} \rightarrow R$ is a separable and continuously differentiable concave function, $0 < a_i$, $0 < b_j$, and $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$. Assume that the problem (1) has a strictly feasible solution.

In order to solve (1), we incorporate the constraints $0 \leq x_{ij}$ and $0 \leq x_{ij} \leq u_{ij}$ into the objective function by introducing barrier terms

$$\beta x_{ij} \ln x_{ij} \text{ and } \beta(x_{ij} \ln x_{ij} + (u_{ij} - x_{ij}) \ln(u_{ij} - x_{ij})),$$

and obtain the following problem:

$$\begin{aligned} \min \quad & e(x, \beta) \\ \text{subject to} \quad & \sum_{j=1}^n x_{ij} = a_i, i = 1, \dots, m, \\ & \sum_{i=1}^m x_{ij} = b_j, j = 1, \dots, n, \end{aligned} \quad (2)$$

where

$$\begin{aligned} e(x, \beta) = \quad & f(x) + \beta \sum_{i=1}^m \left(\sum_{j=1}^{p_i} x_{ij} \ln x_{ij} \right. \\ & \left. + \sum_{j=p_i+1}^n (x_{ij} \ln x_{ij} + (u_{ij} - x_{ij}) \ln(u_{ij} - x_{ij})) \right), \end{aligned}$$

and β is a parameter and varies from a sufficiently large positive number to zero. Consider the following Lagrange function of problem (2):

$$\begin{aligned} L(x, \lambda) = \quad & f(x) + \sum_{i=1}^m \lambda_i^r \left(\sum_{j=1}^n x_{ij} - a_i \right) + \sum_{j=1}^n \lambda_j^c \left(\sum_{i=1}^m x_{ij} - b_j \right) \\ & + \beta \sum_{i=1}^m \left(\sum_{j=1}^{p_i} x_{ij} \ln x_{ij} \right. \\ & \left. + \sum_{j=p_i+1}^n (x_{ij} \ln x_{ij} + (u_{ij} - x_{ij}) \ln(u_{ij} - x_{ij})) \right), \end{aligned} \quad (3)$$

where $\lambda = (\lambda_1^r, \dots, \lambda_m^r, \lambda_1^c, \dots, \lambda_n^c)^\top$ is the so-called Lagrange multiplier vector. Computing the partial derivative of $L(x, \lambda)$ with respect to x_{ij} , we obtain

$$\frac{\partial L(x, \lambda)}{\partial x_{ij}} = \frac{\partial f(x)}{\partial x_{ij}} + g_{ij}(x_{ij}, \lambda) \quad (4)$$

where

$$g_{ij}(x_{ij}, \lambda) = \lambda_i^r + \lambda_i^c + \begin{cases} \beta(1 + \ln x_{ij}) & \text{if } 1 \leq j \leq p_i, \\ \beta \ln \frac{x_{ij}}{u_{ij} - x_{ij}} & \text{if } p_i + 1 \leq j \leq n. \end{cases}$$

Let

$$\nabla_x L(x, \lambda) = \left(\frac{\partial L(x, \lambda)}{\partial x_{11}}, \dots, \frac{\partial L(x, \lambda)}{\partial x_{mn}} \right)^\top,$$

which is the gradient of $L(x, \lambda)$ at x . The necessary optimality condition says that, if x^* is an optimal solution of the problem (2), then there exists some λ^* together with x^* satisfying that

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) &= 0, \\ \sum_{j=1}^n x_{ij}^* &= a_i, \quad i = 1, \dots, m, \\ \sum_{i=1}^n x_{ij}^* &= b_i, \quad j = 1, \dots, n. \end{aligned} \tag{5}$$

Based on all these, a neural network annealing algorithm (briefly NNAA) for problem (1) can be proposed as follows:

Algorithm (NNAA):

Step 0: Let β be a sufficiently large positive number satisfying that $e(x, \beta)$ is strictly convex. Let x^0 be an arbitrary point satisfying $0 < x_{ij}^0$, $j = 1, \dots, p_i$, and $0 < x_{ij}^0 < u_{ij}$, $j = p_i + 1, \dots, n$, $i = 1, \dots, m$. Take (r^0, c^0) to be an arbitrary positive vector. Given $x = x^0$, let $(r^*(x^0), c^*(x^0))$ be a positive approximate solution of neural network, which starts at (r^0, c^0) . Let $r^0 = r^*(x^0)$ and $c^0 = c^*(x^0)$, and compute

$$x_{ij}^1 = \begin{cases} \frac{1}{r_i^*(x^0)c_i^*(x^0)\alpha_{ij}(x^0)} & \text{if } 1 \leq j \leq p_i, \\ \frac{u_{ij}}{1+r_i^*(x^0)c_i^*(x^0)\alpha_{ij}(x^0)} & \text{if } p_i + 1 \leq j \leq n, \end{cases} \tag{6}$$

$i = 1, \dots, m$, $j = 1, \dots, n$. Let $q = 1$ and go to Step 1.

Step 1: Given $x = x^q$, let $(r^*(x^q), c^*(x^q))$ be a positive approximate solution of neural network, which starts at (r^0, c^0) . Go to Step 2.

Step 2: Compute

$$d_{ij}(x^q) = \begin{cases} \frac{1}{r_i^*(x^q)c_i^*(x^q)\alpha_{ij}(x^q)} & \text{if } 1 \leq j \leq p_i, \\ \frac{u_{ij}}{1+r_i^*(x^q)c_i^*(x^q)\alpha_{ij}(x^q)} & \text{if } p_i + 1 \leq j \leq n, \end{cases} \tag{7}$$

$i = 1, \dots, m, j = 1, \dots, n$. If $\|d(x^q) - x^q\|$ is sufficiently small, the algorithm terminates when β is small enough, or decrease β by a factor, let $x^1 = x^q, r^0 = r^*(x^q), c^0 = c^*(x^q)$, and $q = 1$, and go to Step 1. Otherwise, do as follows: Let

$$\lambda^*(x^q) = (\lambda_1^{r^*}(x^q), \dots, \lambda_m^{r^*}(x^q), \lambda_1^{c^*}(x^q), \dots, \lambda_n^{c^*}(x^q))^{\top}$$

with $\lambda_i^{r^*}(x^q) = \beta \ln r_i^*(x^q)$, $i = 1, \dots, m$, and $\lambda_j^{c^*}(x^q) = \beta \ln c_j^*(x^q)$, $j = 1, 2, \dots, n$, and compute

$$x^{q+1} = x^q + \theta_q(d(x^q) - x^q), \quad (8)$$

where θ_q is a positive number in $(0, 1]$ satisfying

$$L(x^{q+1}, \lambda^*(x^q)) < L(x^q, \lambda^*(x^q)).$$

Let $q = q + 1$, and go to Step 1.

3 Numerical Results

We have used the algorithm to solve a number of the minimum concave cost transportation problems. These test problems are from Floudas and Pardalos [14] and generated randomly. We have programmed the algorithm in MATLAB. All the problems have the following form:

$$\begin{aligned} \min \quad & f(x) = \sum_{i=1}^m \sum_{j=1}^n (c_{ij}x_{ij} + d_{ij}x_{ij}^2) \\ \text{subject to} \quad & \sum_{j=1}^n x_{ij} = a_i, i = 1, \dots, m, \\ & \sum_{i=1}^m x_{ij} = b_j, j = 1, \dots, n, \\ & 0 \leq x_{ij}, \end{aligned} \quad (9)$$

where $d_{ij} < 0$, $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$, $0 < a_i$ and $0 < b_j$. In the implementation of the algorithm, β equals 100 initially and is reduced by a factor of $\frac{8}{10}$ when $\|d(x^q) - x^q\|_1 < 0.001$. The algorithm starts with four different initial points and terminates as soon as it converges.

Example 1 This example is from Floudas and Pardalos [14] and the data is as follows: $m = 6, n = 4, a = (8, 24, 20, 24, 16, 12)^\top, b = (29, 41, 13, 21)^\top$,

$$c = (c_{ij}) = \begin{pmatrix} 300 & 270 & 460 & 800 \\ 740 & 600 & 540 & 380 \\ 300 & 490 & 380 & 760 \\ 430 & 250 & 390 & 600 \\ 210 & 830 & 470 & 680 \\ 360 & 290 & 400 & 310 \end{pmatrix},$$

and

$$d = (d_{ij}) = \begin{pmatrix} -7 & -4 & -6 & -8 \\ -12 & -9 & -14 & -7 \\ -13 & -12 & -8 & -4 \\ -7 & -9 & -16 & -8 \\ -4 & -10 & -21 & -13 \\ -17 & -9 & -8 & -4 \end{pmatrix}.$$

The algorithm converges to the global optimal solution x^* in average 61 iterations

$$x^* = (x_{ij}^*) = \begin{pmatrix} 6 & 2 & 0 & 0 \\ 0 & 3 & 0 & 21 \\ 20 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 \\ 3 & 0 & 13 & 0 \\ 0 & 12 & 0 & 0 \end{pmatrix}.$$

4 Conclusions

In this paper, to solve the minimum concave cost transportation problem efficiently, we integrate the barrier and Lagrange functions with neural network to propose a neural network annealing algorithm. The experimental results also confirm the efficiency of the proposed models and algorithm for the minimum concave cost transportation problem. In the future, the proof of theorems in this algorithm and the extended application of the algorithm will be given.

Acknowledgements This work was partially supported by NSFC under Grant No. 61672371, GRF: CityU 11302715 of Hong Kong SAR Government, Jiangsu Provincial Department of Housing and Urban-Rural Development under grants No. 2017ZD253, Ministry of housing and urban and rural construction under grants No. 2018-K1-007, the grant from Suzhou University of Science and Technology under grants No. XKZ2017011, and China scholarship council.

References

1. Erickson, R.E., Monma, C.L., Veinott Jr., A.L.: Send-and-split method for minimum-concave-cost network flows. *Math. Oper. Res.* **12**, 634–664 (1987)
2. Fang, S.C., Tsao, H.-S.J.: Linearly-constrained entropy maximization problem with quadratic cost and its applications to transportation planning problems. *Transp. Sci.* **29**, 353–365 (1995)
3. Gallo, G., Sodin, C.: Concave cost minimization on networks. *Eur. J. Oper. Res.* **3**, 239–249 (1979)
4. Pardalos, P.M., Rosen, J.B.: Constrained global optimization: algorithms and applications. In: *Lecture Notes in Computer Science*, vol. 268. Springer (1987)
5. Zangwill, W.I.: Minimum concave cost flows in certain networks. *Manag. Sci.* **14**, 429–450 (1968)
6. Guisewite, G.M.: Network problems. In: Horst, R., Pardalos, P.M. (eds.) *Handbook of Global Optimization*. Kluwer Academic Publishers (1995)
7. Hopfield, J.J., Tank, D.W.: Neural computation of decisions in optimization problems. *Biol. Cybern.* **52**, 141–152 (1985)
8. Liu, Q., Dang, C., Cao, J.: A novel recurrent neural network with one neuron and finite-time convergence for k-winners-take-all operation. *IEEE Trans. Neural Netw.* **21**(7), 1140–1148 (2010)
9. Hu, X., Zhang, B.: A new recurrent neural network for solving convex quadratic programming problems with an application to the k-winners-take-all problem. *IEEE Trans. Neural Netw.* **20**(4), 654–664 (2009)
10. Hu, X., Wang, J.: An improved dual neural network for solving a class of quadratic programming problems and its k-winners-take-all application. *IEEE Trans. Neural Netw.* **19**(12), 2022–2031 (2008)
11. Lau, K.M., Chan, S.M., Xu, L.: Comparison of the Hopfield scheme to the hybrid of Lagrange and transformation approaches for solving the traveling salesman problem. In: *Proceedings of International IEEE Symposium on Intelligence in Neural and Biological Systems*, Hendon, pp. 209–217 (1995)
12. Xu, L.: Combinatorial optimization neural nets based on a hybrid of Lagrange and transformation approaches. In: *Proceedings of World Congress on Neural Networks*, San Diego, pp. 399–404 (1994)
13. Xu, L.: On the hybrid LT combinatorial optimization: new U-shape barrier, sigmoid activation, least leaking energy and maximum entropy. In: *Proceedings of International Conference on Neural Information Processing (ICONIP'95)*, pp. 309–312. Publishing House of Electronics Industry, Beijing (1995)
14. Floudas, C.A., Pardalos, P.M.: A collection of test problems for constrained global optimization algorithms. In: *Lecture Notes in Computer Science*, vol. 455. Springer (1990)
15. Wu, Z., Li, B., Dang, C., Hu, F., Zhu, Q., Fu, B.: Solving long haul airline disruption problem caused by groundings using a distributed fixed-point computational approach to integer programming. *Neurocomputing* **269**, 232–255 (2017)
16. Wu, Z., Li, B., Dang, C.: Solving multiple fleet airline disruption problems using a distributed-computation approach to integer programming. *IEEE Access* **5**, 19116–19131 (2017)
17. Wu, Z., Gao, Q., Li, B., Dang, C., Hu, F.: A rapid solving method to large airline disruption problems caused by airports closure. *IEEE Access* **5**, 26545–26555 (2017)

Study on the Characteristics of Vehicle Lane-Changing in the Intersection



Yan Xing, Jinling Wang, Weidong Liu, Lishuang Sun and Faguang Chong

Abstract This paper analyses the behavioral characteristics of lane-changing at signalized intersections. It analyses the expected speed of a vehicle running and the benefits of lane-changing, which is based on the theory of traffic flow at first. Second, it sets up a basic model of the vehicle lane-changing in the intersection, combined with the relative speed, the relative distance and the space required for the vehicle in the lane-changing process. Finally, it analyzed the influence of vehicle lane-changing behavior on traffic flow using traffic flow data which was obtained by the actual survey. Research results show that the original lane traffic decrease by 4.6% and the target lane traffic decreases by 8.1% in 12 signal cycles, which is caused by vehicle lane-changing in the intersection.

Keywords Vehicle Lane-changing · Signal intersection · Road traffic efficiency · Traffic flow

1 Introduction

The behaviors of vehicle following and lane-changing are the two most basic driver's behaviors. Compared with the vehicle following behavior, the lane-changing behavior needs to be considered more for the vehicles. Meanwhile, the driver's decision-making process is more complex and more difficult to describe. The current scholars studying on the motor vehicle lane-changing behavior are mainly carried out as follows: In 2002, Dario D. Salvucci and others used the driving simulator to study the lane-changing behavior of the drivers on the freeway, and analyzed the relationship

Y. Xing · J. Wang · W. Liu · L. Sun · F. Chong
School of Transportation Engineering, Shenyang Jianzhu University, Shenyang,
China

Y. Xing (✉)
State Key Laboratory of Automobile Simulation and Control, Jilin University,
Changchun, China
e-mail: 359063088@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_35

between the behavior of the lane-changing and the change of viewpoint [1]. Kim et al. [2] analyzed the driver’s different lane-changing behavior in 2002, gave the impact of different sections of the road lane-changing behavior on the traffic flow. In 2003, Coifman et al. analyzed the relationship between the behavior of lane-changing occurred at the downstream of the intersection and the efficiency of road traffic [3]. Tanaka et al. in 2008, used the speed optimization model to study the effect of the lane-changing behavior on the traffic flow operation. The study found that lane-changing may lead to two models of high flow and low traffic flow [4]. Qu et al. in 2012, analyzed the vehicle’s discrete characteristics of the intersection, determined the relationship between the motor vehicle flow and the traffic efficiency of the intersection [5–7]. So far, the influence of a lane-changing on road traffic efficiency has not yet achieved a fair result.

In order to analyze the underlying relationship between the lane-changing behavior in the intersection and road traffic efficiency, this paper combined with the above research results, analyzed the benefits of the expected speed of vehicle and lane-changing based on traffic flow theory. Then, it combined with the relative speed, the relative distance and the space that required for the vehicle in the lane-changing process, it sets up the basic model of the vehicle lane-changing in the intersection.

2 Drivers’ Lane-Changing Behavior Analysis

In the traffic environment, the drivers will carry on the relevant driving behavior based on the impact of block interference caused by the surrounding objects, the impact of friction interference and their own needs. Usually, the main factors that should be considered in the vehicle lane change are the relative distance, relative angle and relative speed.

According to the provisions of the *Chinese City Road Traffic Signs and Markings Setup Guide*, different design speed corresponds to the different traffic signs and markings near the intersection. For the specific location, please refer to Table 1.

In fact, in the intersection of a city, many vehicles change lanes near the stop line. In order to pass smoothly or prepare for the road ahead, even if the traffic rules are violated, the driver will be forced to change lanes. There are two main reasons for the drivers to change lanes without considering purposeful lane-changing. One is that the current operating speed of the vehicle cannot meet the driver’s expectations. Because of the different behavior of the drivers, they have different psychological expectations of speed when they change lanes, that is to say, the vehicles are blocked by the interference. Another is the interference of other vehicles, which makes their

Table 1 The relationship between design speed and the length of the forbidden path in the intersection

Travel speed (km/h)	≥60	<60
Forbidden transformation lane length (m)	50 ~ 100	30 ~ 50

own freedom of space or comfort decreased. In order to pursue a higher efficiency, they are forced to change. In the intersection, the traffic rules do not allow to overtake. However, when the drivers do not achieve the desired speed, they often violate the rules and regulations to make lane-changing. Although this situation can meet the wishes of the drivers about lane-changing, it often forces the surrounding vehicles to change the current operating characteristics, causing traffic flow fluctuations, affecting the overall efficiency of the intersection.

The most important reason is the first mentioned above. From [8, 9], the research findings can be concluded that the expected speed of different drivers is derived from the following equation:

$$V_n = V_0 \cdot \theta \quad (1)$$

where V_n is the normal speed of the driver's vehicle, θ is the risk factor which reflects the driver type, and it is a normal distribution random number between 0.5 and 1.5 with the mean value as 1. The following method can generate random number:

$$\theta' = E\theta + \left(-7 + \sum_{i=1}^n (RN[0 \cdots 1]) \right) \cdot \sigma_\theta$$

$$\theta = \begin{cases} 0.5 \cdots \theta' < 0.5 \\ \theta' \cdots 0.5 < \theta' < 1.5 \\ 1.5 \cdots \theta' > 1.5 \end{cases} \quad (2)$$

According to the above equations, we can get desired speeds of different lane types as shown in Table 2.

Therefore, when the speed of the front vehicle cannot meet the driver's psychological expectation, the driver will choose to change lanes, or even overpass if the conditions permit.

In order to succeed in changing the lanes, it not only needs to consider the front and the rear vehicles' speed whether meet the lane-changing needs, but also needs to consider whether there is enough space for the vehicle to overtake. In the intersection, the demand of the lane-changing space is mainly reflected in the demand for lateral space. When carrying on the overtaking, the drivers turn left toward the left lane

Table 2 The expected speed of drivers in different types of lanes

Type of lane change	V = 80 km/h	V = 60 km/h	V = 40 km/h
	V_n	V_n	V_n
Rash type	94–100	72–78	48–52
Smart type	82–89	57–64	42–48
Careful type	80–84	58–65	40–47
Retarded type	70–72	50–54	36–38

or turn right toward the right lane according to their own traffic demand. When the drivers do not reach the stop line, they turn left or right to change lanes according to their own needs. In this process, the drivers no longer pursue the desired speed, but determine whether to change the road according to the distance between vehicles. Of course, in order to turn around, the driver may put his car between two cars by force.

Huang pointed out in [10] that the efficiency of vehicle lane-changing usually refers to the increase in the degree of freedom of vehicles. In the intersection, the efficiency of lane-changing not only consider the degree of freedom gained, but also consider the time efficiency, comfort, safety, etc. after taking the behavior of lane-changing to turn left or turn right. When the running speed of peripheral vehicles and space distance satisfy the lane-changing requirements for target vehicles, whether to overtake the lane-changing depends on whether it is able to increase the comprehensive benefits.

Before taking overtaking behavior in the target lane, the driver first judges the speed and distance of the front car and the rear car. When the vehicle speed is smaller and the distance is bigger, the lane-changing behavior is more advantageous. In the case of only considering the operation speed of the vehicles, the efficiency calculation model of the overtaking lane-changing is as follows:

$$V_l = [\alpha \frac{d_{na} - d_{nm}}{d_{nm}} + \beta \frac{v_a - v_n}{v_n}] \gamma_l \quad (3)$$

$$V_r = [\alpha \frac{d_{nc} - d_{nm}}{d_{nm}} + \beta \frac{v_c - v_n}{v_n}] \gamma_r \quad (4)$$

where V_r and V_l are the left and the right adjacent lanes' lane-changing benefits, respectively; α and β represent the space and speed benefit parameters, respectively; d_{nc} , d_{na} and d_{nm} represent the space distance between the target vehicle and the right front, the left front and the forward vehicles, respectively; v_c , v_a and v_n represent the speed of right front vehicle, the speed of left front vehicle and the speed of the target vehicle, respectively; γ_r and γ_l are the preference coefficient of lanes, respectively.

In the intersection, the driver has reduced the requirements of the speed and space. Under the normal conditions, before the intersection stop line, the driver has a high demand for the distance between front and rear vehicles. However, the requirement of the speed has increased after entering the intersection. Lane-changing behavior of the vehicles in the intersection can be divided into two stages:

(1) Vehicle transformation efficiency before the stop line

$$\begin{cases} V_r = \alpha \frac{d_{nc} - d_{nm}}{d_{nm}} \gamma_r \\ V_l = \alpha \frac{d_{na} - d_{nm}}{d_{nm}} \gamma_l \end{cases} \quad (5)$$

(2) Vehicle transformation efficiency after the stop line

$$\begin{cases} V_r = \beta \frac{v_{nc} - v_{nm}}{v_{nm}} \gamma_r \\ V_l = \beta \frac{v_{na} - v_{nm}}{v_{nm}} \gamma_l \end{cases} \quad (6)$$

After entering the intersection, when the driver is in the act of overtaking, drivers not only consider the factors of speed and space, but also consider the safety of such driving behavior, comfort, etc. Thus, the calculation for the driver's lane-changing behavior after the stop line is changed to:

$$\begin{cases} V_r = \beta \frac{v_{nc} - v_{nm}}{v_{nm}} \gamma_r - \sigma_r(v, d) \\ V_l = \beta \frac{v_{na} - v_{nm}}{v_{nm}} \gamma_l - \sigma_l(v, d) \end{cases} \quad (7)$$

where σ_r and σ_l are risk factors for vehicle lane-changing, respectively, mainly depending on the factors of speed and space; the other parameters are the same as above.

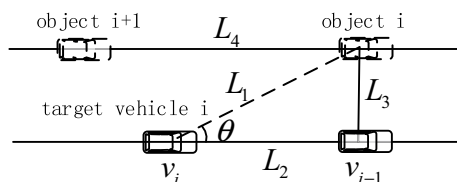
When the driver needs to change a lane in the intersection, he/she will first judge the speed and distance of the front and rear vehicles on the target lane. If the surrounding conditions permit, the driver may analyze the safety and the comfort in the process of lane-changing. When the lane-changing can not only meet the needs of the driver but also not be dangerous, the driver will take decisive action.

3 Establishing the Vehicle Lane-Changing Model in the Intersection

Under the conditions of block or friction disturbance, the basic situation of the lane-changing is shown in Fig. 1, i is the target vehicle, $i - 1$ is the vehicle that has been surpassed or the object

Because the running speed v_{i-1} of the object $i - 1$ cannot meet the expected speed $\bar{v}_i$ of the target vehicle i , the target vehicle i needs to overtake the object $i - 1$. When the target vehicle i overtakes the object $i - 1$, the actual running speed $\bar{v}_i$ of the target vehicle i will be faster than the running speed v_{i-1} of the object $i - 1$, that is $v_i > v_{i-1}$. In this point, it is needed to determine whether there is enough

Fig. 1 The basic diagram of overtaking and lane-changing



space for lane-changing. If the lane-changing space of the target lane cannot meet the conditions, the relative speed of the target vehicle i and the object $i - 1$ should meet the following Eq. (8) in order to ensure the safety of traffic.

$$\Delta v \leq 2a_s(L_2 + l_p)/(v_i + v_{i-1}) \quad (8)$$

where Δv is the relative speed of the target vehicle and the object; L_2 is the relative distance of the target vehicle and the object; a_s is the deceleration of car braking, usually is $7.5 - 8 \text{ m/s}^2$.

If the space of the target lane can meet the lane-changing requirements for target vehicles, the target vehicle just need to satisfy the longitudinal minimum distance in the lane-changing process. The lateral minimum distance is always larger than the longitudinal and the lateral minimum safe distance of the target vehicle and the leading vehicle. The detailed calculation method is shown in Eqs. (9–11):

$$t = \sqrt{\frac{v_i^2 + 4a_i L_1}{2a_i}} \quad (9)$$

$$l_p \geq \int_0^t v_i \cos \theta dt \quad (10)$$

$$l_p \geq \int_0^t v_i \sin \theta dt \quad (11)$$

where t is the time needed by the target vehicle from the start to the end of the lane-changing; θ is the angle between the connecting line from the beginning to the end of the lane-changing and the running direction before the lane-changing.

If the lane space in front of the target lane cannot meet the changing demand of the target vehicle, the target vehicle needs to continue to follow the leading vehicle and guarantee the relative speed meet the above equations. If the rear space of the target lane cannot meet the demand of lane-changing, the target vehicle needs to turn on the steering light to change the road. Until the relative speed of the target vehicle and the object in the rear of the target vehicle meets the Eq. (12), the vehicle can carry on changing lane.

$$a_i \geq \frac{a_{i+1} \cdot v_i^2}{4(L_1 - L_4 + l_p)a_{i+1} - v_{i+1}^2} \quad (12)$$

where a_i is the acceleration of the target motor vehicle's lane-changing; a_{i+1} is the acceleration variation value of object $i + 1$; L_4 is the distance between the target vehicle and the object $i + 1$ at the beginning of the lane-changing; L_1 is the distance from the start to the end of the vehicle's lane-changing; v_i and v_{i+1} are the speed of the target vehicle and the object at the beginning of lane-changing.

4 Analyze the Behavior of Overtaking Lane-Changing in the Intersection

In order to analyze the influence of the traffic flow and the traffic efficiency, which are both caused by the behavior of lane-changing in the intersection, a field survey was carried out on the intersection of People's Main Street in Changchun, China. As shown in Fig. 2, survey time is the evening peak, the investigation object is the straight lane from the west to the east, and there is a signal cycle of lane-changing. Analyze the lane cycle traffic flow, the spacing of vehicle in front and rear of vehicle lane-changing in the original lane and the spacing of vehicle in front and rear of vehicle lane-changing in the target lane in the intersection of 12 cycles. Then, it analyses the lane traffic efficiency and spacing and compares them with no vehicle lane-changing situation.

In the survey, the block lane-changing refers to the target vehicle lane-changing, under the condition of the front vehicle has a slow running speed but the adjacent lane has adequate lane-changing space. There is a certain running speed when the vehicle changes a lane. Purpose lane-changing that it refers, in this paper, the target vehicle lane-changing behavior which is currently located in the straight lane but in the downstream intersection or the current intersection needs to make a left turn and transform to the left lane in advance. In the survey, whether the purpose lane-changing or block lane-changing, only selects one car to change the lane in one cycle as the research object. Then, it compares the two with no vehicle lane-changing situation. Specific survey data is shown in Tables 3 and 4.

Analyze Table 3 and contrast the flow of block lane-changing, purpose lane-changing, and no lane-changing as shown in Fig. 3.

From Fig. 3, it can be found that the block lane-changing has a small impact on the traffic volume of the lane or target lane. The main reason is that in the condition of blocking interference, the target lane-changing is usually selected with adequate lane-changing space and has no effect on the lane. And because of the target vehicle lane-changing, the rear vehicles will accelerate to fill the vacated space caused by the lane-changing. Therefore, it will have less impact on the lane of the target vehicle; In terms of the rear vehicles of the target vehicle on the target lane, they will be affected,

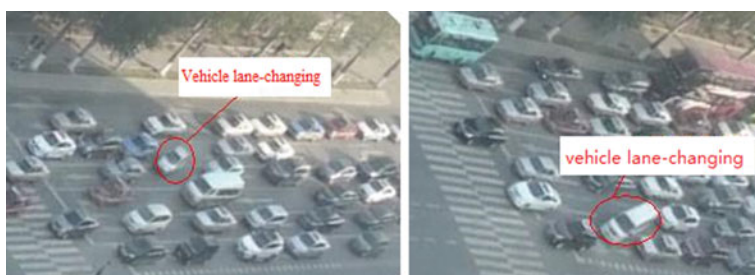


Fig. 2 Vehicle lane-changing in the intersection

Table 3 Changes of lane flow in vehicle lane-changing

Investigation of cycle number	Block lane-changing (target lane flow)	Purpose lane-changing (target lane flow)	Block lane-changing (lane flow)	Purpose lane-changing (lane flow)	No lane-changing
1	33	33	33	31	35
2	35	35	35	32	36
3	36	34	35	32	37
4	32	33	32	31	34
5	35	31	35	30	35
6	31	33	31	31	35
7	36	35	35	32	36
8	34	32	34	30	34
9	34	32	34	30	36
10	35	33	34	31	33
11	32	32	32	30	36
12	35	34	35	31	35

Table 4 Front and rear vehicles' spacing change caused by vehicle lane-changing

Investigation of cycle number	Block lane-changing (target lane flow)	Purpose lane-changing (target lane flow)	Block lane-changing (lane flow)	Purpose lane-changing (lane flow)	No lane-changing
1	12	27	22	28	14.7
2	11	28	20	28	15.3
3	14	24	21	26	16.4
4	11	27	18	30	14.2
5	13	26	25	31	14.8
6	10	25	19	29	16.2
7	12	27	22	32	15.6
8	13	24	27	29	16.3
9	14	26	21	31	13.4
10	11	25	22	27	14.5
11	12	28	19	28	15.6
12	14	29	23	30	14.7

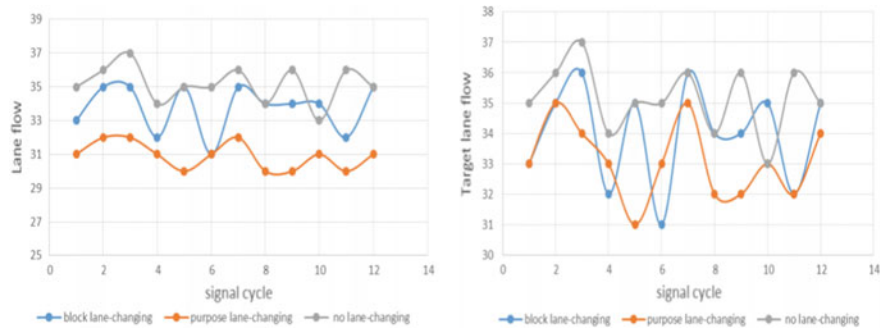


Fig. 3 Vehicle lane-changing leads to flow change in the intersection

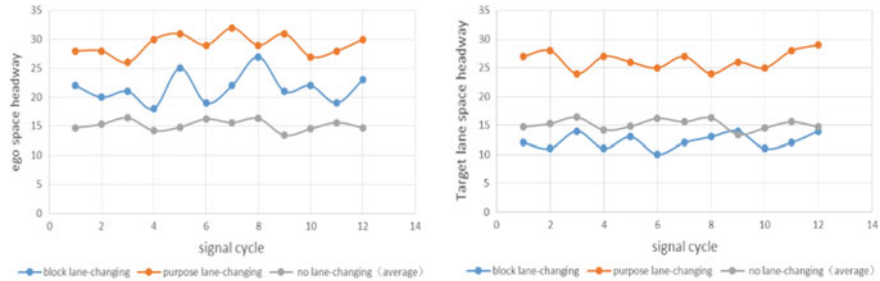


Fig. 4 Vehicle lane-changing leads to distance change between front and rear vehicles

forcing them to suddenly slow down, and even lead to traffic accidents. However, due to the addition of the target vehicle, the flow rate of the target lane is less, resulting in the lane traffic decreased by 3.3%, resulting in the target lane traffic decreased by 4.0%. Because the lane-changing space is small and the operation speed is slow, Purpose lane-changing has a more obvious influence on the flow rate of the original lane and target lane, resulting in the lane traffic decreased by 5.9%, resulting in the target lane traffic decreased by 12.1%.

Space headway refers to the distance between the two adjacent cars under the block interference after the target vehicle lane-changing. Space headway of the target lane refers to the distance between the target vehicle and the front vehicle after the target vehicle lane-changing; Space headway of purpose lane-changing refers to the distance between the two adjacent cars under the block interference after the target vehicle lane-changing. Space headway of no lane-changing refers to average headway distance in the intersection of a lane without a vehicle lane-changing. Analyze Table 4 and contrast the space headway of block lane-changing, purpose lane-changing and no lane-changing. The details are shown in Fig. 4.

From Fig. 4, it can be found the block lane-changing has a big impact on space headway. The main reason is that the target vehicle changes lane, which leads to a larger free space in the lane. Purpose lane-changing has the influence on the fluctua-

tions of the distance in the original lane. The main reason is that the target vehicle's running speed is low, which forces to impede the running speed of the motor vehicles in the original lane and cause the lane has a larger free space. Through the survey, it found that the space headway of the block lane-changing reduced relative to the no lane-changing in the target lane. This is mainly due to the block lane-changing make the vehicle has a certain speed. The lane-changing vehicle will follow the front vehicle in the target lane and lead to reduce the distance between them. Purpose lane-changing also has the influence on the fluctuations of the distance in the target lane. The reason is the same as the change the target vehicle's running speed is low, which forces to impede the running speed of the motor vehicles in the target lane and causes the lane has a larger free space.

5 Conclusions

Combined with adjacent vehicles' relative speed, relative distance, and space required for lane-changing, the paper has established the basic model of the vehicle lane-changing in the intersection. The paper also analyzed the effect of vehicle lane-changing behavior on traffic flow operation using the traffic flow data in the intersection obtained by the actual survey. The results showed that the original lane traffic decreased by 4.6% and target lane traffic decreased 8.1% in 12 signal cycles, meanwhile, the original headway reduced by 11.7% and target lane headway reduced by 82%, because of vehicle lane-changing (including block lane-changing and purpose lane-changing) in the intersection. Lane-changing and first pass have a strong impact on the traffic efficiency in the intersection. At the same time, the impact on traffic safety is also very obvious.

Acknowledgements This work is supported by the Postdoctoral Science Foundation of China (No. 2016M601373). Research Topic of Economic and Social Development in Liaoning, China (2018lsiktqn-058). Basic Scientific Research Projects of Higher Education, Liaoning, China (L17JBKY04). Philosophy and social science planning project of Shenyang, China (2017012). The project of Shenyang Social Sciences (SYSK2018-08-09, SYSK2018-08-10).

References

1. Salvucci, D.D., Liu, A.: The time course of a lane change: driver control and eye-movement behavior. *Transp. Res. Part F*, pp. 123–132 (2002)
2. Kim, S., Son, Y., Kang, J.: Behaviors and characteristics of traffic flow within the influence zone of freeway merge area in Korea. In: *Proceedings of the 3rd International Conference on Traffic and Transportation Studies*, pp. 919–926. ASCE Publication, Reston VA (2002)
3. Coifman, B., Krishnamurthy, S., Wang, X.: Lane-change maneuvers consuming freeway capacity. In: *Proceedings of 5th International Conference on Traffic and Granular Flow*, pp. 3–14. Springer, Berlin (2003)

4. Tanaka, K., Nagatani, T., Masukura, S.: Fundamental diagram in traffic flow of mixed vehicles on multi-lane highway. *Physica. Part A* **387**(5), 5583–5596 (2008)
5. Qu, Z., Xing, Y., et al.: The departure characteristics of traffic flow at the signalized intersection. *Math. Probl. Eng.* **2013**, 1–11 (2013)
6. Qu, Z., Xing, Y., et al.: A study on the coordination of urban traffic control and traffic assignment. *Discret. Dyn. Nat. Soc.* **2012**(1), 1–12 (2012)
7. Xing, Y., Gao, Z., et al.: Study on vehicle delay based on the vehicle arriving distribution at entrance lanes of intersection. *Procedia Eng.* **137**, 599–608 (2016)
8. Gipps, P.G.: A model for the structure of lane-changing decisions. *Transp. Res. Part B.* 403–414 (1986)
9. Zou, Z.J., Yang, D.Y.: Lane changing model for micro traffic simulation. *China J. Highw. Transp.* **15**(2), 105–108 (2002)
10. Huang, Q.: Study on Lane Change Behavior Character and Its Influence To Traffic Safety. Harbin Institute of Technology, Harbin (2007). [https://doi.org/10.1061/41064\(358\)17](https://doi.org/10.1061/41064(358)17)

Part VI
New Generation Intelligent Building
Platform Techniques

A P2P Algorithm for Energy Saving of a Parallel-Connected Pumps System



Qianchuan Zhao, Xuetao Wang, Yifan Wang, Ziyang Jiang
and Yunchuang Dai

Abstract A distributed optimal control algorithm is investigated, in this paper, to deal with the problem in a peer-to-peer (P2P) control setting for parallel pumps in heating, ventilation, and air-conditioning (HVAC) systems. Each pump is equipped with a controller and becomes an intelligent node (as a peer), and nodes are equal, self-organizing, and mutually coordinated. When the HVAC system provides the pressure difference and flow rate, each intelligent node applies the random generation of the speed ratio samples method. Then, the nodes coordinate with each other and constantly optimize the speed ratio of each pump, so that the total energy consumption of the system tends to be minimum. Simulation experiment results are provided to demonstrate the performance of the proposed algorithm.

Keywords Distributed · Peer-to-peer (P2P) · Pump in parallel · Speed ratio · Optimal control

1 Introduction

In China, building energy consumption holds 35% of the national energy consumption every year. With people's demands increasing for indoor environment, heating, ventilation, and air-conditioning (HVAC) system energy consumption has occupied 65% of the total building energy consumption, and has been growing [1]. Most of these energy consumptions are consumed in the water system, and pumps are the

Q. Zhao (✉) · X. Wang · Y. Wang

Tsinghua National Laboratory for Information Science and Technology,
Department of Automation, Center for Intelligent and Networked Systems,
Tsinghua University, Beijing 100084, China
e-mail: zhaoqc@tsinghua.edu.cn

Z. Jiang · Y. Dai

Building Energy Research Center, Tsinghua University, Beijing 100084, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_36

main power consumer in the water system. In 1967, the University of Cincinnati in the United States adopted the variable speed pumps for the first time in the air-conditioning system. With the increasingly serious energy crisis, the variable speed pumps have developed rapidly. Nowadays, they are commonly used in HVAC systems. It cannot only realize the stepless speed regulation of the pump, but also run at a near maximum efficiency, and can control the head and flow rate of the pump under different working conditions. In order to meet the various needs of users, public buildings provide as wide flow range as possible, and pumps are used in parallel. So, how to optimize pump working and achieve energy saving has always been a research hotspot.

Previous effort can be found in wide range methods to optimize pumps, and many of them are based on a centralized control structure. Initially, Westerlund, T. et al. proposed a mixed-integer nonlinear programming method to solve optimal pump configuration [2]. Yang, Z. et al. got static models of different pump combinations and applied Lagrangian multipliers and branch and bound methods to handle the problem [3]. Koor, M. et al. tried to maximize the total efficiency of pumps and used a Levenberg–Marquardt algorithm to realize energy efficient [4]. Jepsen, K. L. et al. proposed a solution to minimizing power consumption of pump station by two steps, first, the aim was formulated to a convex problem, and then pump choosing was solved by convex solver [5]. Yu, G. et al. studied nonlinear programming and the generalized reduced gradient method, optimized the flow to minimize energy consumption [6]. The above methods are based on traditional optimization methods. However, there are some other heuristic algorithms. Such as Wang, J. Y. et al. adopted the genetic algorithm-based pump-scheduling method to reduce energy cost [7]. Barán, B. et al. introduced a multi-objective evolutionary algorithm to solve an optimal pump-scheduling problem with four objectives [8]. Until now, most of the algorithms are based on a centralized control structure. It may limit the application of these algorithms in real because computing resources are not abundant and the implementation costs are very high.

It is well known that the centralized control structure is widely used, in which pumps in parallel connection to a central controller which collects all data and sends commands to the pumps. Pumps depend on the function of the controller, once the controller fails, the HVAC system will not work. When the system is complicated, it will be difficult to upgrade the centralized control system. In addition, most of the control algorithms are only suitable for specific buildings, and it is not easy to reuse them for other buildings. As a result, it is usually required to redesign the system for a different building, which will lead to an extra cost of resources.

Thus, some recent work starts to try to solve the problem by applying a distributed control structure as distributed control structure has better extensibility and can be easily used to any systems. There have been many applications on distributed methods, such as distributed multi-robot coordination [9], multi-agent system [10]. Zhuang, L. et al. proposed a decentralized method basing on ordinal optimization to save energy of an HVAC System [11]. Dai, Y. et al. introduced a novel decentralized method as a solution to a typical problem in HVAC systems [12]. In this paper, a distributed optimal control algorithm is investigated to deal with the problem in a

peer-to-peer (P2P) control setting for parallel pumps in HVAC systems. Each pump is equipped with a controller and becomes an intelligent node (as a peer), and nodes are equal, self-organizing, and mutually coordinated.

2 Problem Description

In general, a typical distributed parallel-connected pumps system which is used in an HVAC system as shown in Fig. 1. The parallel-connected pumps system consists of centrifugal pumps, variable-frequency drive (VFD), pipe, and controller. Each pump can be viewed as a subsystem. There is one controller for each subsystem, so every pump will become an intelligent node. It is easy to remove or add a node to the system. The problem can be expressed as in an HVAC system, pumps in parallel need to satisfy the total flow rate and the head on the main pipe demands. These constraints are equations, so every intelligent node will have a flow rate by communicating with their neighbors to achieve the lowest total energy consumption. The flow rate can be changed when speed control is used. If the demands change, the intelligent nodes will collaborate with each other to reassign the flow rate. Such a problem is common in the HVAC system, in the future, the optimization of the whole HVAC system will be studied. This paper studies the optimization of pumps in a parallel system and puts forward a control algorithm to minimize the entire energy consumption in a distributed form. There is no centralized controller, as a result, we can portion the total flow rate faster and more efficiently.

2.1 Notations

m	the number of pumps;
P_i	the energy consumption of pump i , $i \in \{1, 2, \dots, m\}$;
P^*	the upper limitation of total energy consumption;
ω_i	the speed ratio of pump i , $i \in \{1, 2, \dots, m\}$;
ω^-	the lower limitation of speed ratio;
Q_i	the flow rate of pump i , $i \in \{1, 2, \dots, m\}$;
Q_0	the total flow rate demand on the main pipe;
H_i	the head of pump i , $i \in \{1, 2, \dots, m\}$;
H_0	the total the pressure difference demand on the main pipe;
η_i	the mechanical efficiency of pump i , $i \in \{1, 2, \dots, m\}$;
n_i	the speed of pump i , $i \in \{1, 2, \dots, m\}$;
$n_{0,i}$	the rated speed of pump i , $i \in \{1, 2, \dots, m\}$;

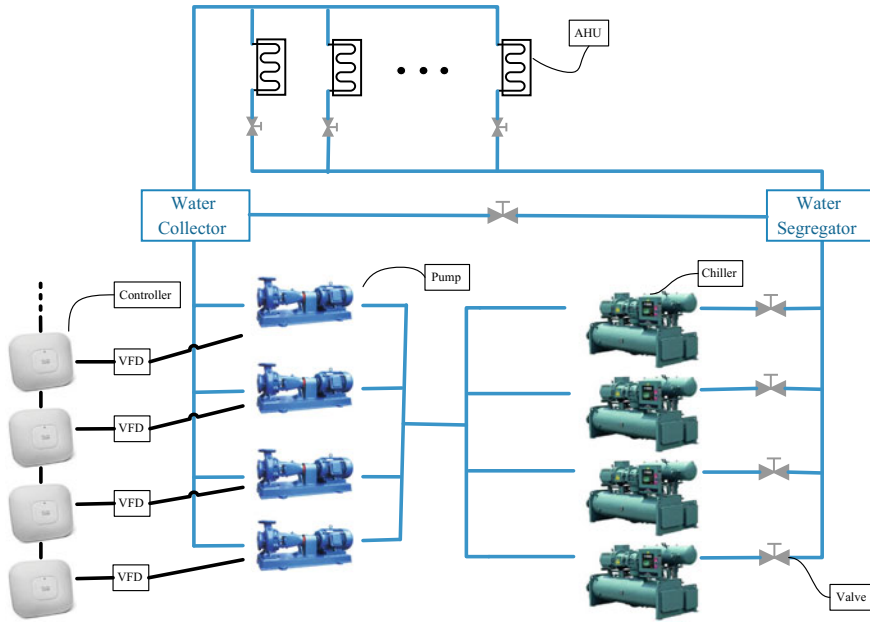


Fig. 1 A parallel-connected pumps system

2.2 Mathematical Model

Mathematical Model of Pump

In a practical application, the performance curve of a pump is assumed to be smooth, so as to often choose polynomial to represent the operation characteristics of the pump. When pumps work at the rated speed $n_{0,i}$, the pump head and flow rate, efficiency and flow rate, and pump energy consumption can be derived as below [13].

$$H_i = a_{P1,i} Q_i^2 + a_{P2,i} Q_i + a_{P3,i} \quad (1)$$

$$\eta_i = b_{P1,i} Q_i^2 + b_{P2,i} Q_i + b_{P3,i} \quad (2)$$

$$P_i = \frac{\rho g Q_i H_i}{1000 \eta_i \eta_{M,i} \eta_{VFD,i}} \quad (3)$$

where H_i is the head of pump i , Q_i is the flow rate of pump i , η_i is the mechanical efficiency of pump i , $\eta_{M,i}$, $\eta_{VFD,i}$ are motor efficiency and variable-frequency efficiency. P_i is the power consumption of pump i . $a_{P1,i}$, $a_{P2,i}$, $a_{P3,i}$ and $b_{P1,i}$, $b_{P2,i}$, $b_{P3,i}$ are pump performance parameters, which could be obtained by actual measurement data. Least square method is often used to fit the functional relation to obtain performance

parameters. ω_i is the speed ratio of pump i , which is defined as the ratio of the actual speed n_i to the rated speed $n_{0,i}$.

$$\omega_i = \frac{n_i}{n_{0,i}} \quad (4)$$

In general, the performance parameters of a pump are related to its speed. If the operating conditions do not change, one can get the performance parameters at different speeds by the pump affinity laws [14] as follows.

$$Q_i(n_{0,i}) = \frac{Q_i(n_i)}{\omega_i}, H_i(n_{0,i}) = \frac{H_i(n_i)}{\omega_i^2}, \eta_i(n_{0,i}) = \eta_i(n_i) \quad (5)$$

It is easy to obtain the equations of the pump head and flow rate, efficiency, and flow rate at a speed n_i by using (4) and (5) as below.

$$H_i(n_i) = a_{P1,i} Q_i^2 + a_{P2,i} \omega_i Q_i + a_{P3,i} \omega_i^2 \quad (6)$$

$$\eta_i(n_i) = b_{P1,i} \left(\frac{Q_i}{\omega_i} \right)^2 + b_{P2,i} \frac{Q_i}{\omega_i} + b_{P3,i} \quad (7)$$

Problem Assumptions

There are two assumptions in the parallel-connected pumps system so that more effectively the optimal control of pumps in distributed form can be solved.

In a large HVAC system, it is common to see that the lower limitation of speed ratio ω^- is set as 0.6, and the purpose is to make sure the motor stable operation. Besides, when the magnitude of the change to ω_i is relatively small, the motor efficiency and variable-frequency driver efficiency are assumed as constants.

$$\omega^- = 0.6 \quad (8)$$

When it needs to change the pump on-off status, we assume that the resistance of the entire pipe network remains a constant. So as to the total flow rate Q_0 and the head H_0 demands on the main pipe will not change.

Bernier, M. et al. pointed out that $\eta_{M,i}, \eta_{VFD,i}$ depends on ω_i [15]. If the speed ratios satisfy $0.4 \leq \omega_i \leq 1$, $\eta_{M,i}, \eta_{VFD,i}$ can be taken as constants and ω_i will have little influence on them. In this case, we can ignore the impact of $\eta_{M,i}, \eta_{VFD,i}$ and get the power consumption equation as (9).

$$P_i(\omega_i) = \frac{\rho g Q_i H_i}{1000 \eta_i} \quad (9)$$

Optimization Objective

$$\min_{\omega_i, i \in \{1, 2, \dots, m\}} \sum_{i=1}^m P_i(\omega_i) \quad (10)$$

Equation (10) shows that the optimization objective of this paper is to minimize the power consumption of all pumps in parallel. It means that we need to find the combination of ω_i to meet the total flow rate Q_0 and the pressure difference H_0 demands at the lowest energy consumption.

Constraints

$$H_0 = H_i \quad (11)$$

$$Q_0 = \sum_{i=1}^m Q_i \quad (12)$$

$$\omega^- \leq \omega_i \leq 1 \text{ or } \omega_i = 0, i \in \{1, 2, \dots, m\} \quad (13)$$

In an HVAC system, pumps in parallel should satisfy the overall demands constraint. Constraint (11) illustrates the relationship among the flow rate of pump, the speed ratio, and the pressure difference. H_0 is the total pressure difference demand on the main pipe. Constraint (12) ensures that the sum of the flow rate of each pump is equal to the total flow rate demand Q_0 on the main pipe.

3 Distributed Algorithm

In our distributed control system, each pump is equipped with a controller, a VFD. The controllers have been written an identical computing program. The program at a controller communicates with the neighbor node and calculates flow rate and power consumption. When the total flow rate and the pressure difference demands change, every intelligent node will get the information and begin to start a new process of adjustment to redesign speed ratio to meet the change and try to minimize power consumption.

In the process of adjustment, each intelligent node will perform the distributed algorithm given as below to determine the optimal speed ratio of each pump.

Algorithm 1. P2P algorithm for computing optimal pump speed ratio

- Step 1:** Initialize $P^* = \infty$, $k = 0$, r number of round, current optimal speed ratio $\omega^* = []$, starting node's calculation_flag = 1, speed ratio $\omega_i \sim U(\omega^- - \varepsilon, 1)$, ε sample adjustment parameter;
- Step 2:** Generate samples ω_i and set **trade – off_flag_i** to 1; where $\omega_i[m] \geq \omega^-$, use Eqs. (6), (7), (9) and H_0 to get $Q_i[m]$, $P_i[m]$, otherwise $\omega_i[m] = 0$, $Q_i[m] = 0$, $P_i[m] = 0$; $Q_{sum,i}^k = Q_i$, $P_{sum,i}^k = P_i$;
- Step 3:** Send calculation_flag, $Q_{sum,i}^k$, $P_{sum,i}^k$, P^* , ω_i , ω^* , **trade – off_flag_i** to neighbor, clear calculation_flag and receive calculation_flag, $Q_{sum,j}^k$, $P_{sum,j}^k$, P^* , ω_j , ω^* , **trade – off_flag_j** from neighbors;
- Step 4:** if calculation_flag == 1 and **trade – off_flag_j**[m] == 1 and num_neighbor > 1 then calculate $Q_{sum,i}^k[m] = Q_{sum,j}^k[m] + Q_i^k[m]$, $P_{sum,i}^k[m] = P_{sum,j}^k[m] + P_i^k[m]$; judge whether $Q_{sum,i}^k[m] \leq Q_0$, $P_{sum,i}^k[m] < P^*$ are satisfied, otherwise clear **trade – off_flag_j**[m]; **trade – off_flag_i** = **trade – off_flag_j**;
- Step 5:** if calculation_flag == 1 and **trade – off_flag_j**[m] == 1 and num_neighbor == 1 then calculate $Q_i^k[m] = Q_0 - Q_{sum,j}^k[m]$ and use Eq. (6) and H_0 to get $\omega_i^k[m]$; use Eqs. (7), (9) and H_0 to get $P_i^k[m]$, $P_{sum,i}^k[m] = P_{sum,j}^k[m] + P_i^k[m]$; when $\omega_i^k[m] = 0$ or $\omega^- \leq \omega_i^k[m] \leq 1$, $P_{sum,i}^k[m] < P^*$ are satisfied, update $P^* = P_{sum,i}^k[m]$ and ω^* ;
- Step 6:** $k = k + 1$ if $k > r$ then go to step 7; elsewhere calculation_flag == 1 and num_neighbor > 1, go to step 3, otherwise go to step 2;
- Step 7:** Return $\omega_i = \omega^*[i]$;

As a result, each node will obtain the speed ratio ω_i . We can find the constraint will be satisfied in the entire regulation process, and we need not discuss different combinations of pumps.

4 Experimental Results

We apply distributed simulation platform to implement the operation of distribute algorithm. The operating environment is Windows 10 × 64, Intel (R) Core (TM) i7-7700@3.60GHZ, DSP1.0, and computer memory is 8 GB.

In order to test the distributed algorithm, we simulate the pumps in parallel system that is shown in Fig. 1. Furthermore, we get the performance parameters of pump from the paper [16] for simulation. There are two kinds of pumps and the specific performance parameters as follows (Table 1).

Table 1 Performance parameters of pumps

Pump	Performance parameters					
	$a_{P1,i}$	$a_{P2,i}$	$a_{P3,i}$	$b_{P1,i}$	$b_{P2,i}$	$b_{P3,i}$
PUMP-A	−0.0046	0.0696	60.271	−0.0002	0.0254	0.0616
PUMP-B	−0.0112	0.1358	54.841	−0.0005	0.0316	0.2582

Table 2 Requirements of HVAC system

	$H_0(m)$	$Q_0(L/s)$
Case 1	26	86
Case 2	29	117
Case 3	35	173
Case 4	39	232

During the computer simulation, there are six pumps. Four of pumps are PUMP-A (big pump) and two of pumps are PUMP-B (small pump). Each pump is equipped with a controller and all controllers are connected one by one. We will use this case to illustrate the performance of the distributed algorithm and the HVAC system requirements are given in Table 2.

In this paper, in order to demonstrate the effectiveness of the proposed Distributed Algorithm (DA), we compare the result with the one from a distributed algorithm in [12] and we call it Transfer Algorithm (TA). It is possible to refer to [12] for the details of the Transfer Algorithm.

During the simulation, when we test TA, we set the flow rate of each pump randomly and meet the total flow rate constraint. Besides, the convergence threshold is set to 0.5 L/s. It is difficult to decide the step adjustment of the expected relative efficiency. At first, the step is equal to 0.005 each time. However, as for TA, the iteration of some cases is not convergence, which is shown in Fig. 2.

To show effectiveness, we define energy-saving rate ξ_P as follows.

$$\xi_P = \frac{\sum P_i^{TA} - \sum P_i^{DA}}{\sum P_i^{DA}} \times 100\% \quad (14)$$

For TA to obtain convergent results, our numerical experiments indicate that the step should be set as 0.0005 each time, which implies that TA takes more time for the operation. Table 3 shows the speed ratio and total power consumption of the pumps in the parallel system for the proposed distributed algorithm and Transfer Algorithm. Obviously, it can be found from the results that TA spends more power than DA. It is not proper for different pumps to pursue high efficiency, because extra energy will be consumed. In addition, there is no need to set the initial flow rate of each pump for the distributed algorithm to meet the total constraints; therefore it can portion the total flow rate more quickly.

Table 3 Numerical results of cases

Case	Method	ω_i								$\sum P_i$ (K W)	ξ_P	$\sum \dot{Q}_i - \dot{Q}_0$ (L/s)
		PUMP-A	PUMP-A	PUMP-A	PUMP-A	PUMP-A	PUMP-B	PUMP-B	PUMP-B			
1	DA	0.7400	0.7248	0	0	0	0	0	0	25.400	7.43	0
1	TA	0.7354	0	0	0	0	0.7269	0.7269	0.7269	27.288		0.423
2	DA	0.7937	0.7392	0	0	0	0.7427	0	0	39.501	0.29	0
2	TA	0.7790	0.7803	0	0	0	0.7710	0	0	39.614		0.358
3	DA	0.8782	0.8659	0.8993	0	0	0	0	0	68.535	2.60	0
3	TA	0.8500	0.8491	0.8482	0	0	0.8408	0	0	70.320		0.461
4	DA	0.8865	0.9441	0.9270	0.9168	0.9168	0	0	0	102.562	2.02	0
4	TA	0.8936	0.8928	0.8936	0.8936	0.8936	0.8864	0	0	104.633		0.285

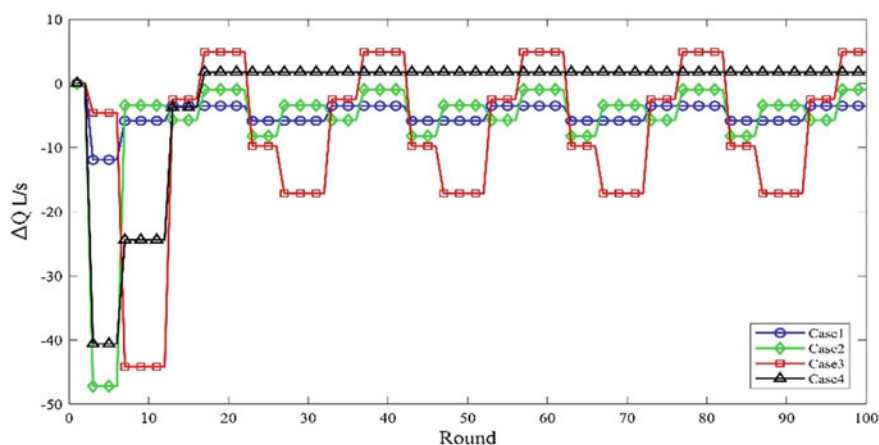


Fig. 2 The iteration of the transfer algorithm

5 Conclusions

In this paper, we proposed a distributed algorithm for parallel-connected pump systems in an HVAC system. When minimizing the total power consumption, it runs in a totally self-organized way, so that it supports the change of configurations of the pumps in the parallel system. Comparisons in numerical experiments show the effectiveness of our approach for reducing the energy consumption when satisfying the demands of the HVAC system and advantages against existing algorithms. Currently, we design and test the distributed algorithm in synchronized mode, in the future, we will try to extend it to asynchronous mode to improve its robustness and enlarge possible application scenarios.

Acknowledgements This work was supported by the National Key Research and Development Project of China (No. 2017YFC0704100 entitled New Generation Intelligent Building Platform Techniques, and 2016YFB0901900) and the National Natural Science Foundation of China (No. 61425027), the 111 International Collaboration Program of China under Grant B06002, BP2018006 and Special fund of Suzhou-Tsinghua Innovation Leading Action (Project Number: 2016SZ0202).

We appreciate the constructive and helpful comments from anonymous reviews and conference organizers.

References

1. Liu, Y., Lam, J.C., Tsang, C.L.: Energy performance of building envelopes in different climate zones in China. *Appl. Energ.* **85**(9), 800–817 (2008)
2. Westerlund, T., Pettersson, F., Grossmann, I.E.: Optimization of pump configurations as a MINLP problem. *Comput. Chem. Eng.* **18**(9), 845–858 (1994)

3. Yang, Z., Børsting, H.: Energy efficient control of a boosting system with multiple variable-speed pumps in parallel. In: 2010 49th IEEE Conference on Decision and Control (CDC) (2010)
4. Koor, M., Vassiljev, A., Koppel, T.: Optimization of pump efficiencies with different pumps characteristics working in parallel mode. *Adv. Eng. Softw.* **101**, 69–76 (2016)
5. Jepsen, K.L., Hansen, L., Mai, C., Yang, Z.: Power consumption optimization for multiple parallel centrifugal pumps. In: IEEE Conference on Control Technology and Applications (CCTA), pp. 806–811 (2017)
6. Yu, G., Powell, R.S., Sterling, M.J.H.: Optimized pump scheduling in water distribution systems. *J. Optim. Theory Appl.* **83**(3), 463–488 (1994)
7. Wang, J.Y., Chang, T.P., Chen, J.S.: An enhanced genetic algorithm for bi-objective pump scheduling in water supply. *Expert Syst. Appl.* **36**(7), 10249–10258 (2009)
8. Barán, B., von Lüken, C., Sotelo, A.: Multi-objective pump scheduling optimization using evolutionary strategies. *Adv. Eng. Softw.* **36**(1), 39–147 (2005)
9. Sheng, W., Yang, Q., Tan, J., Xi, N.: Distributed multi-robot coordination in area exploration. *Robot. Autonom. Syst.* **54**(12), 945–955 (2006)
10. Nissim, R., Brafman, R.I.: Multi-agent A* for parallel and distributed systems. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 3 (2012)
11. Zhuang, L., Chen, X., Guan, X.: A decentralized ordinal optimization for energy saving of an HVAC system. In: American Control Conference (ACC) (2016)
12. Dai, Y., Jiang, Z., Shen, Q., Chen, P., Wang, S., Jiang, Y.: A decentralized algorithm for optimal distribution in HVAC systems. *Build. Environ.* **95**, 21–31 (2016)
13. Zhao, T., Zhang, J., Ma, L.: On-line optimization control method based on extreme value analysis for parallel variable-frequency hydraulic pumps in central air-conditioning systems. *Build. Environ.* **47**, 330–338 (2012)
14. Pumps, S.: *Centrifugal Pump Handbook*, 3rd edn. Elsevier Science. (2010)
15. Bernier, M., Bourret, B.: Pumping energy and variable frequency drives. *ASHRAE J.* **41**(12), 37–40 (1999)
16. Dai, Y., Jiang, Z., Xin, S, et al.: Optimal control of variable speed parallel-connected pumps. In: Proceedings of the 13th International Conference on Indoor Air Quality and Climate, pp. 87–94 (2014)

A Distributed Algorithm for Building Space Topology Matching



Yifan Wang and Qianchuan Zhao

Abstract Nowadays, distributed systems turn to be increasingly competitive with centralized systems, especially in the construction field. Thus, it is indispensable to design new effective distributed algorithms. In this work, a distributed algorithm focusing on the building space topology matching problem is proposed. In distributed control architectures, each floor is equipped with several Computing Process Nodes (CPNs) which play roles as controllers in the system. Our goal is to make every CPN in the building system acquire its own position on a designed CAD drawing. To achieve this, we utilize the geographic relationships among nodes. In the algorithm, each node compares its local topology features to the ones in the drawing and communicates with their neighborhoods. We prove that the topology matching problem can be solved by using this algorithm and derive the upper bound of the number of iterations. The experiments show that this algorithm works successfully in real designed buildings.

Keywords Distributed algorithm · Building space topology matching · CPN · Geographic direction · Topology feature

1 Introduction

Currently accompanied with the continuous expansion of building scale and the installation of a large number of sensors, the complexity of the building structure tends to be unprecedentedly high. It becomes increasingly difficult to handle all the data processing and calculation in the traditional centralized system. As a result, new constructions controlled by distributed smart nodes have been designed. Jiang Z. Y. et al. introduced a decentralized, flat-structured building automation system con-

Y. Wang · Q. Zhao (✉)

Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Center for Intelligent and Networked Systems (CFINS), Tsinghua University, Beijing, China
e-mail: zhaoqc@tsinghua.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_37

397

trolled by smart nodes in 2017 [1]. In these architectures, distributed smart nodes, also called CPN, manage the control system. CPNs possess the same structure, a microcomputer which can store data and do calculations. Each CPN controls a subspace in the building. A CPN has 6 ports to connect to the nearby CPNs which are called its neighbors. The 6 ports mean that a node can communicate with neighbors from 6 directions of up, down, north, south, east, and west. The only way for one node to acquire the external data is to communicate with its neighbors.

To control the whole building system appropriately by distributed smart nodes, one important thing is to let all nodes find its location in the system. This is our topology matching problem, which can be converted to a two-graph correspondence problem. The previous effort has been down in many related research fields. To obtain relations between individuals in associated networks, Yao X. G. et al. set up a multi-objective optimization model and a mathematical model based on the adjacency matrix to solve the node matching problem [2, 3]. Tao S. Q. et al. proposed an algorithm based on gradient flows for subgraph matching [4]. Compared to our topology matching problem, these papers focus on the partial match in the graph, which invokes us to consider the features of a node and its adjacent subgraph. Fan Y. J. et al. introduced a method to measure graph similarity by matching global structure, considering both vertex and path feature [5]. Actually, if we just think about the graph characteristic of vertexes and edges, sometimes two nodes cannot be distinguished because they are totally symmetrical in the graph. It leads to a graph isomorphism problem discussed by McKay and Piperno [6]. Without other assumptions, the topology matching problem can only be solved when the graph's *automorphism group* only contains itself. However, Mushtaha E. S. et al. provided a thought to analyze a building from the perspective of geographic direction [7]. Learning from this idea, we hypothesize there is a direct relationship between CPNs. Under this assumption, the question is able to be solved through our algorithm. Our algorithm provides a topology model for the general Insect Intelligent Building (I²B) architecture introduced in [8].

2 Problem Description

The problem studied in this paper is based on distributed control architectures. Our objective is to make every node to obtain its location on a designed CAD drawing. The process is of great significance because sometimes, it is necessary to control several specified nodes to work simultaneously in a building. Specifically, a number of nodes belonging to one company need to do some statistical analysis at the end of each month. As a result, there will be a confusion if the company have no information about the nodes' ID and location. Since a CPN is incapable to obtain the global information and the labeling process should start after the consummation of nodes installation, it is really challenging for a CPN to locate itself.

The problem consists of n CPNs in a distributed architecture system and a CAD drawing which has the totally same topology with the system. The true IDs of nodes are labeled on the drawing. An algorithm should be designed to let all the CPNs find their true ID through local calculation and communication with their neighbors.

2.1 Notations

n	the number of CPNs;
v_i	the i th vertex in the graph, $i \in \{1, 2, \dots, n\}$;
d	the diameter of the graph;
N_i^m	the set of v_i 's m th-order neighbors, $v_j \in N_i^m$ means there is a path of length m between v_i and v_j , here and in the rest of paper, length means the number of edges in the path;
P_{ij}^m	the set of directed path of length m from v_i to v_j , the directed path is described as an array of directions;
F_i^m	the m th-order local feature of v_i on the real topology, an aggregate of all m th-order directed path between v_i and its m th-order neighborhoods;
DF_i^m	the m th-order drawing feature of v_i on the drawing topology.

2.2 Problem Assumptions

Assumption 1: The topology on the CAD drawing is exactly the same as the real topology of the corresponding building system.

This assumption ensures the problem is feasible.

Assumption 2: The number of CPNs in the building system is finite.

Absolutely, the scale and volume of a building are finite.

Assumption 3: Each node has at least one neighbor to communicate and there is a communication path between any two nodes.

This assumption means that the system cannot be divided into independent parts. There is a path for any two CPNs to communicate.

Assumption 4: There are only 6 kinds of directions from one node to its adjacent nodes: up, down, north, south, west and north.

This assumption utilizes the geographic direction according to the characteristics of architecture system and CPN. The port number of CPN is 6 and it means that there are only neighbors from 6 directions of up, down, north, south, west, and east. Actually in real CAD drawings, CPNs are connected just from these 6 directions, like a cube structure.

2.3 Mathematical Model

To describe the problem mathematically, we convert the CPN network to a special graph. In the graph, a vertex delegates a CPN while an edge represents a connection between two nodes. Only two adjacent vertexes can exchange their information.

Through the assumption, the graph is a connected graph whose number of nodes is limited. In this particular graph, the degree of each vertex is no larger than 6 and the direction set of an edge is composed of three elements: up and down, north and south, and east and west. The neighborhoods of a node can be differentiated through the direction of the edge.

Now the mathematical description of the problem can be conducted: there are two same connected graphs. One is labeled with numbers called its ID while the other is without this information. Our goal is to make nodes on the latter graph find their ID in the former graph, which delegates a real CPN searching for their specific location. For example, in Fig. 1, the true ID for every node on the drawing is shown in the left graph while the real graph without ID is the right one. The left one is the drawing topology while the right one is the real topology. Our goal is to find node-wise correspondence between these two graphs.

To achieve the goal above, we define the local feature F_i^j for each node v_i . The subscript i indicates that the feature belongs to node v_i while the superscript j represents the feature order. First, we determine that two nodes v_i, v_j are m th-order neighbors if there is a path of length m between these two nodes and we note that $v_i \in N_j^m, v_j \in N_i^m$ where N_i^m is the set of v_i 's m th-order neighbors. For example, N_i^1 is v_i 's adjacent nodes. Since the directions of edges have been determined, paths

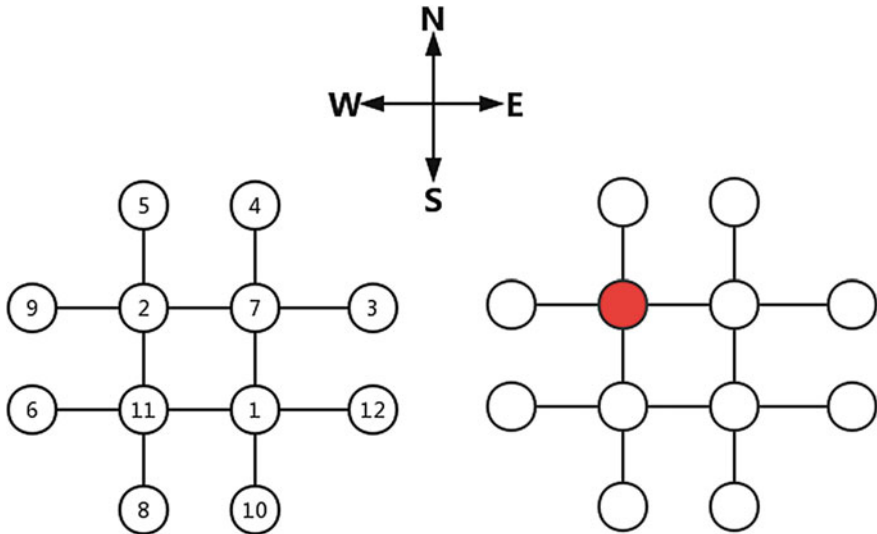


Fig. 1 Drawing topology and real topology

composed of a series of directions can be obtained between any two nodes. P_{ij}^m is defined as a set of m th-order directed paths between v_i, v_j . Each element in P_{ij}^m is an array of directions from v_i to v_j . There are only 6 directions: up, down, north, south, east, west. For instance, a directed path from node 1 to node 9 in the drawing topology is $[w, n, w]$ which means west, north, and west and $P_{19}^3 = [[w, n, w], [n, w, w]]$.

The local feature F_i^j is defined as an aggregation of all m th-order directed path between v_i and its m th-order neighborhoods. In the mathematical description, $F_i^m = \bigcup_{j \in N_i^m} P_{ij}^m$. For example, F_1^1 is defined as an array containing neighbors' direc-

tion. In Fig. 1, we assume that the red point is temporarily named node 1 in the real topology. F_1^1 can be expressed as $[[n], [s], [w], [e]]$ and F_1^2 can be expressed as $[[n, s], [s, n], [s, w], [s, s], [s, e], [w, e], [e, n], [e, s], [e, w], [e, e]]$. The drawing feature DF_i^m is determined for nodes on the drawing. It delegates the feature obtained from the CAD drawing and has the same structure as the local feature.

2.4 Objective

The objective of this paper is to design an algorithm to solve the topology matching problem described in Sect. 2.3 by comparing the local feature to drawing the feature and utilizing the information transmitted between nodes and their neighbors.

3 Building Space Topology Matching Algorithm

Now we have a distributed architecture system controlled by CPNs and a CAD drawing possessing the same topology structure. The user chooses a random CPN from the system and input the CAD drawing information. The topology matching algorithm is a synchronized distributed algorithm shown below (Fig. 2).

Step 1: Any CPN withdrawing information input from the user can be chosen as a leader.

Step 2: The leader starts to generate a spanning tree as a communication tree used to broadcast messages and collect calculation results. The specific tree generation algorithm is the layered BFS algorithm from *Distributed System* [9]. After that, the leader broadcasts the topology information of the drawing to all nodes through the communication tree.

Step 3: Now CPNs obtain the information of the designed topology on the CAD drawing. For each v_i , calculate its local feature F_i^1 and compare it to drawing features of all nodes on the drawing $DF_j^1 (j = 1, 2, \dots, n)$. A possible nodes collection of v_i is composed of drawing nodes which possesses the same feature as v_i 's. If there is only one element in v_i 's possible nodes collection, it can be determined to be identified.

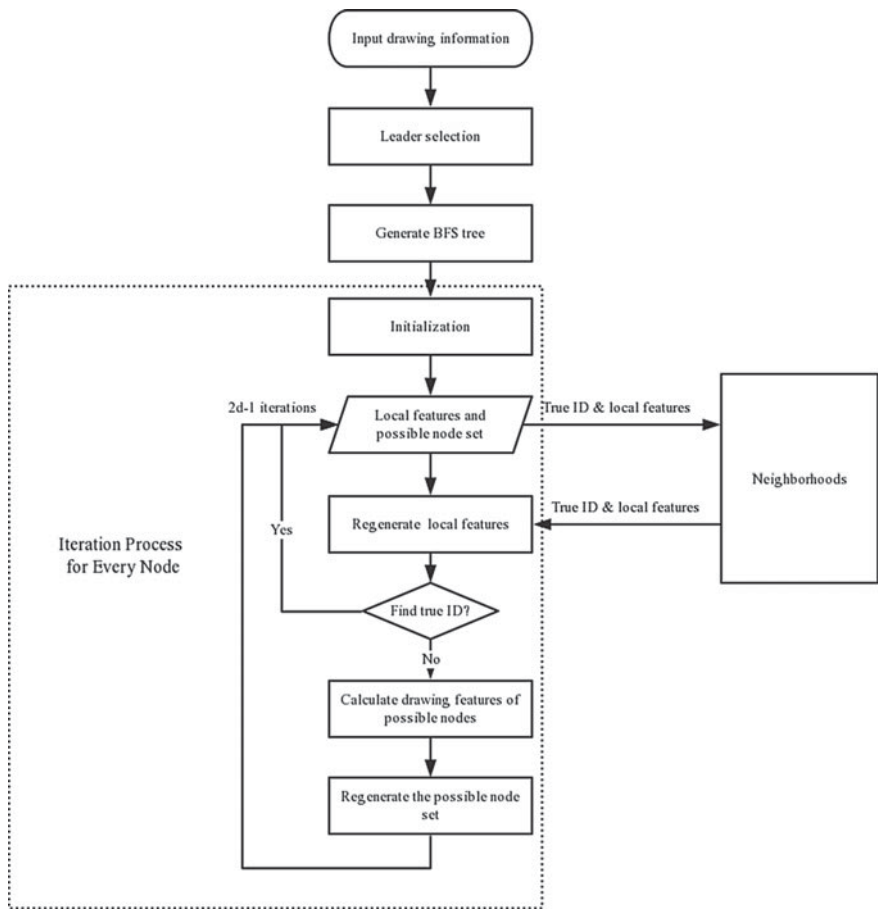


Fig. 2 Algorithm flowchart

Step 4: Identified nodes start to inform their ID to neighbors while other nodes transmit their first-order feature to their neighbors.

Step 5: If v_i receives the information from an identified node, it can be identified through the direction of the edge. Otherwise, it calculates local feature F_i^2 by using the information of its neighbors' local features $F_j^1 (j \in N_i^1)$. Meanwhile v_i can calculate the drawing features DF_k^2 of v_k in its possible node collections. Through comparison, redundant nodes can be removed from the collection. Now if v_i 's possible nodes set only has one node, it can be determined to be identified.

Step 6: Repeat step 4 and step 5 until all the nodes find its true ID. In these steps, F_i^2 , DF_j^2 should be modified to F_i^{m+1} , DF_j^{m+1} and $F_j^1(j \in N_i^1)$ should be changed to $F_j^m(j \in N_i^m)$ where m is the iteration round number. The process can be proved to finish in n iterations.

Step 7: All nodes send their true ID to the leader through the communication tree and the leader displays the results to the user.

4 Building Space Topology Matching Algorithm

In the proof section, we propose three proofs. First, one node which can be differentiated from others can be found in the graph. Second, by using the node we find in Proof 1, all the nodes can be exactly distinguished from others. Third, the algorithm can solve the problem in at most $2d - 1$ iteration rounds.

Proposition 1 In a connected graph with a limited number of nodes, there is a node whose features which are defined in Sect. 2.3 can be distinguished from others.

Proof 1 The paper uses the method of reduction to absurdity to prove it. We suppose that for every node, there is another node whose features are exactly the same. Since the number of nodes is limited, we assume that the diameter of the graph is d . We can find v_1 and v_2 that their distance is d . Through the hypothesis, v'_1 possessing the same features as v_1 can be found. A path of length d , named p_1 , can be selected between v_1 and v_2 . Similarly, v'_1 has a path named p'_1 which is the same as p_1 connecting to v'_2 . Then, we discuss in two situations.

Situation 1: p'_1 and p_1 intersect at node v_3 .

We suppose that the distance between v_1 and v_3 is d_1 and the distance between v'_1 and v_3 is d'_1 . If $d_1 = d'_1$, v_1 and v'_1 will be the same node because p_1 and p'_1 are the same. It is impossible. So according to the symmetry, we suppose that $d_1 > d'_1$. Considering the path between v_1 and v'_2 , the length is $d_1 + d - d'_1 > d$ which is contradicted to the fact that the diameter of the graph is d .

Situation 2: p'_1 and p_1 do not intersect.

Due to that the graph is connected, there is a shortest path connected p_1 and p_2 . We assume that the endpoints are node x from p_1 and node y from p_2 . The path length is $l \geq 1$. Because the path is the shortest, no other nodes from p_1 and p_2 will be on the path. We note that d_2 is the distance between v_1 and x while d'_2 is the distance between v'_1 and y . According to the symmetry, we suppose that $d_2 \geq d'_2$. Now the length of the path between v_1 and v'_2 is $d_2 + l + d - d'_2 > d$ which is also contradicted to the fact that the diameter of the graph is d .

In conclusion, the hypothesis is proved to be incorrect. We can definitely find a node whose features can be distinguished from others.

Proposition 2 After finding one node in Proof 1, all nodes can find its location on the drawing in at most $d - 1$ iteration rounds.

Proof 2 Suppose that v_1 is the node identified. For a random node v_2 in the connected graph, if the distance between v_1 and v_2 is d . Through symmetry and Proof 1, the v_2 can also be determined in the process of v_1 's distinction process. So, we assume that the distance between v_1 and v_2 is $d_{12} \leq d - 1$. Through the assumption that all edges in the graph have certified direction, v_2 can find its location after d_{12} iteration rounds.

Proposition 3 The algorithm can solve the problem in at most $2d - 1$ iteration rounds.

Proof 3 Since the d th-order local feature of the node we find in Proof 1 is unique, after d iterations, the node can definitely be identified. With the conclusion in Proof 2, the whole algorithm can end in at most $2d - 1$ iterations.

5 Experiments

The experiment is based on the distributed algorithm simulation platform written in Python. The operating environment is Mac OS X10.11.5, 2.2 GHz Intel Core i7, 16 GB 1600 MHz DDR3.

A CPN is simulated as a process in the distributed algorithm simulation platform and each process has eight different threads working. Six threads are socket servers used to simulate the communication between ports of nodes. One thread is a socket server used to receive foreign instructions and send back operation results. Meanwhile, another one is the calculation thread which is used to do data processing and calculations. We use two drawing topologies to examine the topology matching algorithm. One is designed by myself which is highly symmetrical while the other one is a part of a real conference hall with a three-floor structure.

In the first experiment, the topology which is shown in Fig. 1 is used. For the red node in Fig. 1, its F^1 is $[[n], [s], [w], [e]]$. Comparing it to $DF_i^1 (i = 1, 2, \dots, n)$, we can find that its possible nodes set is $[2, 7, 11, 1]$. However, its F^2 is $[[n, s], [s, n], [s, w], [s, s], [s, e], [w, e], [e, n], [e, s], [e, w], [e, e]]$ and its unique in the possible nodes set, which refers to node 2. So 2 iterations are needed to identify the red node, according to the symmetry, the other 3 nodes in the center also need 2 iteration rounds to be identified. The external 8 nodes can be recognized after receiving messages from the center 4 nodes. The algorithm ends after 3 iterations and the number of iterations for each node is shown in Fig. 3.

In the second experiment, we use a real CAD drawing of a conference hall of the three-floor structure. We can obtain a goal cube topology through the connection relationships of CPNs in the drawing, which is shown as the left graph in Fig. 4. There are 15 CPNs in the system. Now eleven nodes' F^1 is sole in the system and they can be identified after 1 iteration. The other 4 nodes' true ID can be determined in the next iteration. The algorithm ends after 2 iterations and the number of iterations for each node is shown in the right graph in Fig. 4.

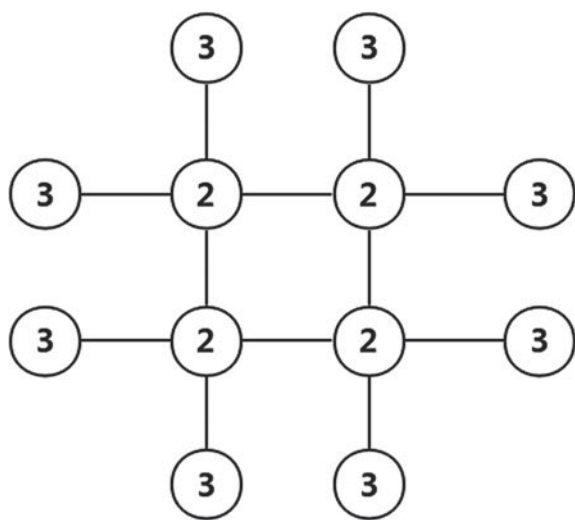


Fig. 3 Customized highly symmetrical topology

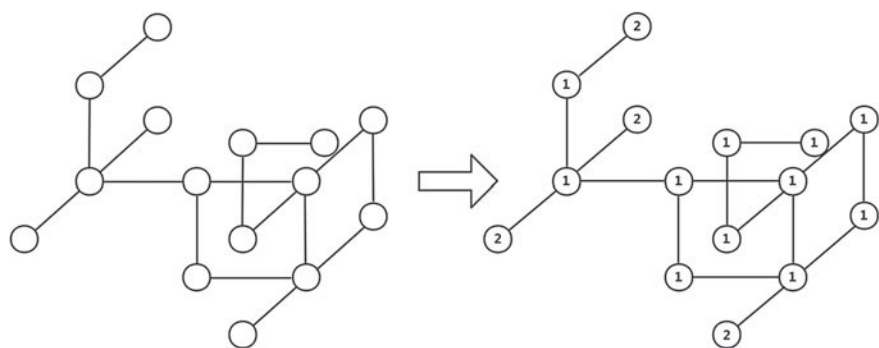


Fig. 4 Real building topology

From the above two experiments, the algorithm is examined to be effective. We can find that if the graph is lack of the symmetry, the algorithm tends to be highly efficient due to the uniqueness of most nodes' features. Furthermore, compared to the traditional methods, this algorithm can solve the topology matching problem without any extra information. As a result, this algorithm is useful in graphs with direction information, like construction field.

6 Conclusions

In this paper, we proposed a new algorithm to solve the building space topology matching problem in the distributed building system equipped with computing process nodes. By utilizing the design characteristics of CPN and the features of the architecture, we make an assumption that the edge between every two controllers has a direction and there are only three kinds of directions: up and down, north and south, east, and west. Under this hypothesis, we declare our algorithm and prove its feasibility in all cases. Furthermore, we give an upper bound of iteration rounds and use real building system to examine the algorithm. In the future, the communication complexity problem and the better design of the local features will be considered. We will also try to test our methods on larger real cases.

Acknowledgements This work was supported by the National Key Research and Development Project of China (No. 2017YFC0704100 entitled New generation intelligent building platform techniques, and 2016YFB0901900), the National Natural Science Foundation of China (No. 61425027), the 111 International Collaboration Program of China under Grant B06002, BP2018006 and Special fund of Suzhou-Tsinghua Innovation Leading Action (Project Number: 2016SZ0202).

We appreciate Dr. Ziyang Jiang and Ms. An Jiang from Tsinghua University, Beijing, China, for the helpful discussion and anonymous reviewers for constructive comments.

References

1. Jiang, Z., Dai, Y.: A decentralized, flat-structured building automation system. *Energy Procedia* **122**, 68–73 (2017)
2. Yao, X., Gong, D., Gu, Y.: Mathematic model of node matching based on adjacency matrix and evolutionary solutions. *Phys. A* **416**, 354–360 (2014)
3. Yao, X., Gong, D., Wang, P., et al.: Multi-objective optimization model and evolutionary solution of network node matching problem. *Phys. A* **483**, 495–502 (2017)
4. Tao, S., Wang, S.: An algorithm for weighted sub-graph matching based on gradient flows. *Inf. Sci.* **340–341**, 104–121 (2016)
5. Fan, Y.J., Chen, L.F., Guo, G.D.: Graph similarity measure by matching global structure. *J. Chin. Comput. Syst.* **37**, 1488 (2016)
6. McKay, B.D., Piperno, A.: Practical graph isomorphism, II. *J. Symbolic Comput.* **60**, 94–112 (2014)
7. Mushtaha, E.S., Arar, M.S.: House typology from adjacency diagram theory to space orientation theory. *Int. J. Civil Environ. Eng.* **11**(2), 60–66 (2011)
8. Zhao, Q., Jiang, Z.: Insect Intelligent Building (I^2B): A new architecture of building control systems based on Internet of Things (IoT). In: Fang, Q., et al. (eds.) *Advancements in Smart City and Intelligent Building, Advances in Intelligent Systems and Computing* 890, Chap. 42 (2019)
9. Lynch, N.A.: *Distributed Algorithms*. Morgan Kaufmann (1996)

Decentralized Differential Evolutionary Algorithm for Large-Scale Networked Systems



Guanghong Han, Xi Chen and Qianchuan Zhao

Abstract The optimization of a complex system with multiple subsystems is a tough problem. In this paper, a Decentralized differential evolutionary algorithm (DDEA) is proposed. The simulations for both DDEA and centralized DE on three benchmark functions are carried out. The numerical results show that DDEA is efficient to solve decentralized optimization problems. On these problems, the proposed DDEA outperforms centralized DE in convergence.

Keywords Optimization · Decentralized · System · Subsystem · Differential evolution

1 Introduction

The optimization of a complex system with large scale and multi-variables is always a tough problem. In general, a complex system is composed of many subsystems, where each subsystem has a decision maker. Traditional centralized algorithms regard the system as a whole and deal with all the decision variables together. However, it will be more and more difficult when the scale of the problem becomes larger. Besides, in some cases, it is hard to find a center node or agent which can collect global information and make decisions. On the other hand, decentralized algorithms solve the problem on each subsystem. The subsystem only exchanges information with their neighbors to find the global solution and no center node is needed. Decentralized algorithms are less likely to have the curse of dimensionality and probably a powerful tool to solve the problems on complex systems.

G. Han · X. Chen (✉) · Q. Zhao
Department of Automation, Center for Intelligent and Networked Systems (CFINS),
Tsinghua University, Beijing, China
e-mail: bjchenxi@mail.tsinghua.edu.cn

Q. Zhao
e-mail: zhaoqc@tsinghua.edu.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_38

The research of decentralized algorithms has been carried out in many fields. Inalhan et al. proposed a decentralized optimization method to solve the multiple aircraft coordination problems [1]. In the field of architecture, a decentralized ordinal optimization for energy saving in an HVAC system was developed [2]. Wei Shi et al. proposed an exact first-order algorithm for decentralized consensus optimization [3]. There are more applications of the decentralized algorithm such as resource allocation in communication networks and collaborative estimation in wireless sensor networks [4].

Evolutionary algorithms are powerful tools to handle real-world optimization problems which are non-convex. Differential evolution (DE) is an efficient evolutionary computation method which performs well in continuous global search problems [5]. It was developed by Storn et al. in 1995 [6]. DE initializes a population of individuals randomly, and each individual represents a feasible solution in the optimization problem. In DE, new individuals are created by mutation and crossover operators. The next generation will be selected from the ones which show better adaptability. After limited times of evolution, the optimal individual will be achieved.

There are many cost minimization problems in engineering practice, where the total cost is the sum of the cost of all subsystems, such as the optimization for energy consumption of an HVAC system. These problems are hard to tackle with gradient-based optimization method in a centralized way, for each subsystem can include complex objective functions.

Focused on this kind of problems, a decentralized differential evolution algorithm (DDEA) was proposed. The structure of DDEA is designed for decentralized computation. It is implemented by introducing delegates and feedbacks to the original DE algorithm.

The remaining part of the paper is presented as follows. In Sect. 2, the decentralized optimization problem is formulated. In Sect. 3, the framework and detailed description of DDEA are given. In Sects. 4 and 5, simulations on three benchmark issues are carried out and comparison between centralized DE (denoted as DE) and DDEA is given. Finally, Sect. 6 summarizes this paper and shows the performance of DDEA.

2 Problem Formulation

In this paper, we focus on a multiple subsystem optimization problem, where each subsystem has a decision maker. The objective of i th subsystem is to minimize a local cost function $f_i: \mathbb{R}^{m_i} \rightarrow \mathbb{R}$, and the global cost function is the sum of all local cost functions. There are couplings in cost functions or constraints, so communication between neighboring subsystems is necessary.

2.1 Centralized Optimization Problem

First, consider an n -subsystem connected system, whose centralized optimization problem is defined as

$$\begin{aligned} & \underset{X \in \mathbb{R}^m}{\text{minimize}} && \sum_{i=1}^n f_i(X_i, X_{\{N_i\}}) \\ & \text{subject to} && g(X) \leq 0 \\ & && h(X) = 0 \end{aligned} \quad (1)$$

where $X_i \in \mathbb{R}^{m_i}$ is the local optimization variable from i th subsystem, and $X = [X_1^T, X_2^T, \dots, X_n^T]^T \in \mathbb{R}^m$ gathers all optimization variables of n subsystems. $m = \sum_{i=1}^n m_i$ is the dimension of X . N_i is neighborhood set of i th subsystem. $g(X) \leq 0$ and $h(X) = 0$ are the inequality and equality constraints, respectively. These constraints can be split into different subsystems.

2.2 Neighborhood of Subsystems

The neighborhood set N_i of i th. subsystem is defined as

$N_i = \{j: \text{the local cost function or constraints of subsystem } i \text{ involve subsystem } j\}$.

Then, the neighboring optimization variables can be defined as

$$X_{\{N_i\}} = \{X_k | k \in N_i\} \quad (2)$$

2.3 Decentralized Optimization Problem

The decentralized optimization is implemented in n subsystems. The optimization problem of i th subsystem is defined as

$$\begin{aligned} & \underset{X_i \in \mathbb{R}^{m_i}}{\text{minimize}} && f_i(X_i | X_{\{N_i\}}) \\ & \text{subject to} && g_j(X_i | X_{\{N_i\}}) \leq 0 \quad j = 1, 2, \dots, p_i \\ & && h_k(X_i | X_{\{N_i\}}) = 0 \quad k = 1, 2, \dots, q_i \end{aligned} \quad (3)$$

where functions $f(X_i | X_{\{N_i\}})$, $g(X_i | X_{\{N_i\}})$ and $h(X_i | X_{\{N_i\}})$ represents the functions of X_i when neighboring variables are given. In this paper, the i th subsystem can only decide X_i and its neighboring optimization variables are considered fixed.

2.4 Notations

n	the number of subsystems;
m	the dimension of the centralized problem;
m_i	the dimension of i th decentralized problem;
$\mathcal{N}_i$	the neighborhood set of subsystem i ;
N	the population size;
E	the number of delegates;
X_i^t	the i th individual of the t th generation;
U_i	the i th individual of children;
$I(i)$	the class number of i ;
$\phi(i)$	the feedback information of i th delegate;
F	the scale factor;
c_r	the crossover rate of DE algorithm.

3 Decentralized Differential Evolutionary Algorithm

Based on DE, we developed a practical algorithm to deal with the decentralized problems formulated by (3), and it is named decentralized differential evolutionary algorithm (DDEA). In each subsystem, a differential evolutionary algorithm is conducted. During each iteration, every subsystem should exchange necessary information with their neighbors. The flowchart of the algorithm in one subsystem is shown in Fig. 1. DDEA is a synchronous algorithm and all subsystem runs the same codes.

In this section, the local variable is denoted by X and neighboring variable is denoted by Y . Suppose the object of this subsystem is to minimize $f(X|Y): \mathbb{R}^m \rightarrow \mathbb{R}$. The algorithm is carried out as the following steps:

Step 1: Initialization.

The same as DE, the population of each subsystem is initialized randomly. The initial individuals can be generated by

$$X_i^{(0)} = X_l + R_i \circ (X^l - X^u), \quad i = 1, 2, \dots, N \quad (4)$$

where R_i is generated from a uniform distribution $U(0, 1)^m$ and $\circ$ means a dot product. X^l and X^u are lower bound and upper bound of X .

Step 2: Select delegates.

To evaluate an individual, we need to know its performance both in local and neighboring subsystems. That is, a local subsystem should exchange its population with neighbors. However, the communication cost will be large if the whole population is transmitted. In DDEA, delegates are selected from the whole population by k-means

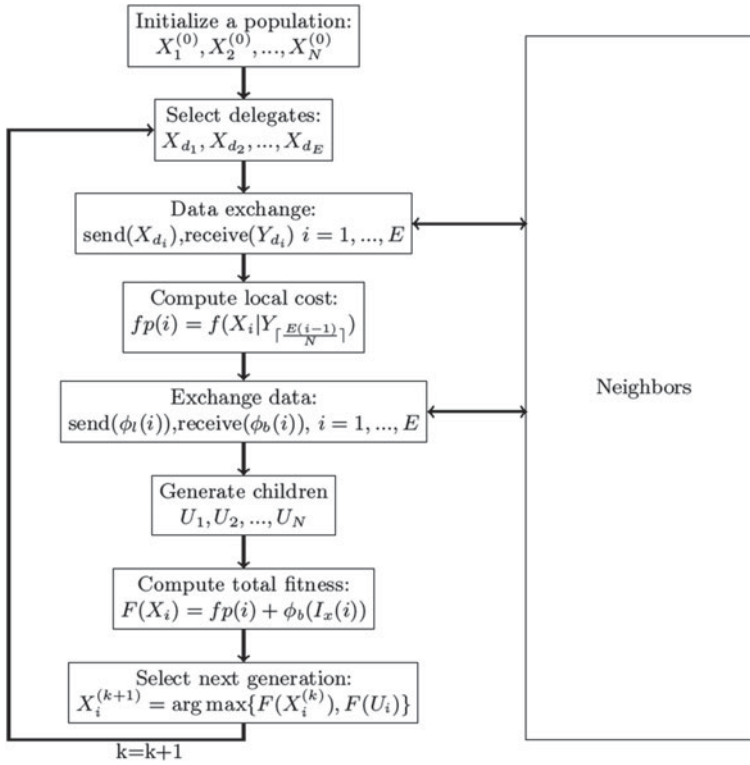


Fig. 1 The flowchart of DDEA in one subsystem

clustering algorithm. This selection method is based on the assumption that close individuals have similar fitness.

The k-means clustering will select E delegates $X_{d_1}, X_{d_2}, \dots, X_{d_E}$, which are cluster centers of all the individuals. Meanwhile, we get cluster indices $I_x(i)$, where $I_x(i) = j$ means $X_i \in \text{class } j$.

Step 3: Data exchange.

In this step, all subsystems will exchange delegates with their neighbors. As shown in Fig. 1, the communication is implemented by $\text{send}(X)$ and $\text{receive}(Y)$. Delegates received from neighbors are denoted by $Y_1, Y_2, \dots, Y_E$.

Step 4: Compute local cost.

When the neighboring delegates $Y_1, Y_2, \dots, Y_E$ are completely received, they will be allocated to N local individuals, and each delegate corresponds to $\frac{N}{E}$ individuals. Then, the local cost is computed by

$$fp(i) = f(X_i | Y_{\lceil \frac{E(i-1)}{N} \rceil}), \quad i = 1, 2, \dots, N \quad (5)$$

where $\lceil x \rceil$ is the nearest integer of x toward infinity.

The feedback of neighboring delegates Y_k is evaluated by

$$\phi(k) = \frac{E}{N} \sum_{i=1}^{\frac{N}{E}} fp\left(\frac{N}{E}(k-1) + i\right) \quad (6)$$

All neighbors will share the same feedback from this subsystem.

Step 5: Data exchange.

In this step, all subsystems will exchange the feedback of neighboring delegates $\phi(i)$. Suppose the feedback is $\phi_b(i)$.

Step 6: Generate children.

DDEA generates the next generation by mutation and crossover. First, N mutant vectors are generated by

$$V_i = X_{r_1} + F(X_{r_2} - X_{r_3}) \quad (7)$$

where r_1, r_2 , and r_3 are randomly chosen from $\{1, 2, \dots, N\}$, and $r_1 \neq i, r_1 \notin \{r_2, r_3\}, r_2 \neq r_3$.

After all the mutant vectors are created, they will cross with the parent individuals X_i to get children:

$$U_{i,j} = \begin{cases} V_{i,j} & \text{if } (r_c < c) \text{ or } (j = \tau_i) \\ X_{i,j} & \text{otherwise} \end{cases} \quad (8)$$

for $j = 1, 2, \dots, m$. Where m is the dimension of V_i and X_i ; c is the crossover rate $\in [0, 1]$; r_c is a random number taken from $U(0, 1)$; τ_i is randomly chosen from $\{1, 2, \dots, N\}$, which makes at least a component of V_i be obtained.

Step 7: Compute total cost.

The total cost is the sum of local cost and neighboring cost. In step 5, we have got the feedback of local delegates, which is an approximation of neighboring cost. For each individual in parent population, the total cost is computed by

$$F(X_i) = fp(i) + \phi_b[I_x(i)] \quad (9)$$

For the children population, no direct feedback information can be used. To estimate the neighboring cost of children individuals, U_i and its nearest parent individual will be classified as one class, and the indices idu is given by

$$I_u(i) = \underset{j}{\operatorname{argmin}} \|U_i - X_{d_j}\| \quad (10)$$

And, the total cost of U_i can be estimated by

$$F(U_i) = fu(i) + \phi_b[I_u(i)] \quad (11)$$

where $fu(i) = f(X_i|Y_{\lceil \frac{E(i-1)}{N} \rceil})$.

Step 8: Select next generation.

In DE, the next generation will be selected from the better individual in each (X_i, U_i) pairs, i.e.,

$$X_i = \begin{cases} U_i & \text{if } F(U_i) \leq F(X_i) \\ X_i & \text{otherwise} \end{cases} \quad (12)$$

When the selection is finished, go back to step 2 until the maximum generation count is reached.

4 Experimental Setup

Three benchmark functions are used to investigate the performance of DDEA. For comparison, we also test the centralized DE on these functions. All experiments are performed in MATLAB.

These problems are centralized optimization problems in the form of (Eq. 1), and they are transformed into decentralized optimization problems. These problems and their decentralized versions are listed in the appendix.

There are 4 hyper parameters in DDEA: the population size N , the number of delegates E , the maximum generations G_m and the scale factor F . These parameters are problem-dependent but in this experiment they are set to constants: $N = 100$, $F = 0.6$, $E = 10$, and G_m is set to 1000 or 200. For the DE algorithm, F is set to 0.6, either, and the crossover rate $cr = 0.2$. For each problem, 50 independent numerical experiments are performed.

5 Results and Discussions

Problem g01 is a 10-dimensional Rastrigin function taken from [7]. The decentralized problem contains 10 subsystems, where the subsystem i is adjacent to subsystem $i + 1$, and the subsystem 10 is adjacent to subsystem 1. The topology of this system is a ring.

Problem g02 is a 20-dimensional Rosenbrock function taken from [8]. This problem is divided into 19 subsystems, where the subsystem i is adjacent to $i + 1$, and the subsystem 19 contains x_{19} and x_{20} . So, the topology of this system is a chain.

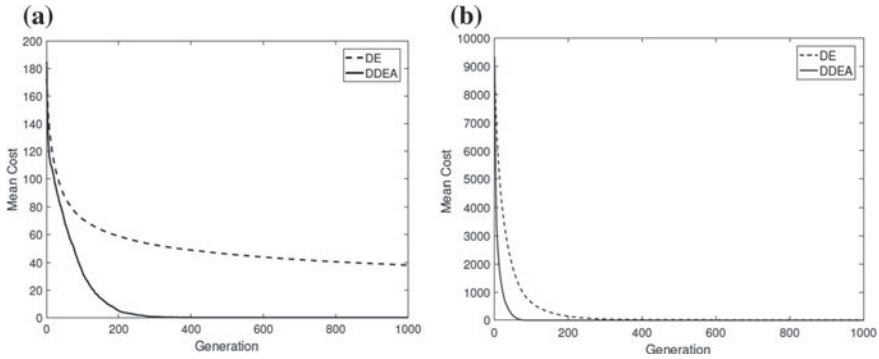


Fig. 2 The performance of DE and DDEA on g01 (a) and g02 (b)

Problem g03 is a 13-dimensional constrained optimization problem taken from. This problem is divided into 13 subsystems, and each subsystem has several neighbors. The problem has been transformed into an unconstrained optimization problem by introducing the penalty term [9].

DE and DDEA cost the same number of evaluations. For g01, the number of evaluations is $d \times N \times G_m = 10 \times 100 \times 1000 = 1000,000$ (an evaluation means calculating the cost function of a subsystem for one time). The numbers of evaluations on g01 and g02 are 2000,000 and 260,000 respectively.

For g01 and g02, the evolutionary process of DE and DDEA are shown in Fig. 2, where the traces show the mean cost of all the individuals at each generation, averaged over 50 simulations. It shows that the cost of DDEA decreases much faster than that of DE. For both g01 and g02, DDEA can find the optimal solution in about 400 generations, but the convergence of DE takes over 1000 generations. Table 2 shows that the performance of centralized DE becomes poorer with the increasing of dimension.

For g03, the evolutionary process is shown in Fig. 3. At first, the mean cost is lower than the optimum -15 , because the solutions are probably not feasible. After the 15th generation, the cost of DDEA decreases faster, and reaches optimum at the 60th generation, which is also faster than DE.

The statistical results of the two algorithms are listed in Table 1, where the mean result, best result and worst result of 50 simulations are included. For the three problems, the results of DDEA are much closer to the optima than the results of DE, and DE cannot reach the optima of g01 and g02 (Table 2).

6 Conclusions

In this paper, we implement the differential evolutionary algorithm in a decentralized way. Simulation results on three benchmark functions show that DDEA is able to find global optima when the problem is non convex, and it converges fast even on

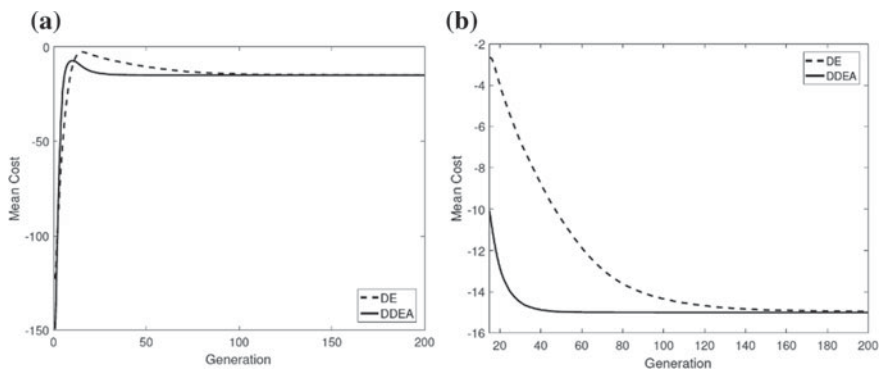


Fig. 3 The performance of DE and DDEA on g03. **a** The process from first generation. **b** The process from 15th generation, when the solutions are feasible

Table 1 Experimental results of DDEA and DE on three benchmark functions

Functions	Optimal	Methods	Mean	Best	Worst	Std
g01	0	DE	19.9606	11.5068	26.6292	3.2560
		DDEA	5.0192e-09	2.3908e-10	2.4350e-08	4.2443e-09
g02	0	DE	9.0136	7.7101	10.4818	0.6233
		DDEA	1.8063e-04	5.5635e-08	7.6266e-04	1.9804e-04
g03	-15	DE	-14.9479	-14.9945	-12.9937	0.2820
		DDEA	-15.0000	-15.0000	-15.0000	3.3212e-11

Table 2 The performance of DE on g02 of different dimensions

Dimension	5	10	15	20
Mean cost	5.9967e-15	0.0919	7.3648	16.3300

high-dimensional problems. For the three benchmark functions, DDEA outperforms than centralized DE. Hence, the proposed algorithm is an efficient and universal technique for decentralized optimization. It can be widely applied to various cost optimization problems of large-scale networked systems.

Acknowledgements This work was supported by National Key Research and Development Project of China (No. 2017YFC0704100 entitled New generation intelligent building platform techniques, and 2016YFB0901900), the National Natural Science Foundation of China (No. 61425027), the 111 International Collaboration Program of China under Grant B06002, and Special fund of Suzhou-Tsinghua Innovation Leading Action (Project Number: 2016SZ0202).

Appendix

In this appendix, the three benchmark problems are listed, followed by the corresponding decentralized problems on subsystems.

g01

$$\underset{\mathbf{x} \in \mathbb{R}^{10}}{\text{minimize}} \quad f(\mathbf{x}) = 10d + \sum_{i=1}^d x_i^2 - 10\cos(2\pi x_i) \quad (13)$$

where $x_i \in [-5.12, 5.12]$. $\mathbf{x}^* = 0$ and $f(\mathbf{x}^*) = 0$.

The problem of i th subsystem is

$$\underset{x_i \in \mathbb{R}}{\text{minimize}} \quad f(x_i | x_{i+1}) = 10 + x_i^2 - 10\cos(2\pi x_{i+1}) \quad (14)$$

where $i = 1, 2, \dots, 10$, and $x_{11} = x_1$.

g02

The problem of i th subsystem is

$$\underset{\mathbf{x} \in \mathbb{R}^{20}}{\text{minimize}} \quad f(\mathbf{x}) = \sum_{i=1}^{d-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \quad (15)$$

where $x_i \in [-2.048, 2.048]$, $x_i^* = 1, \forall i$ and $f(\mathbf{x}^*) = 0$.

The problem of i th subsystem is

$$\underset{x_i \in \mathbb{R}}{\text{minimize}} \quad f(x_i | x_{i+1}) = 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \quad (16)$$

where $i = 1, 2, \dots, 19$.

g03

$$\begin{aligned} \underset{\mathbf{x} \in \mathbb{R}^{13}}{\text{minimize}} \quad & f(\mathbf{x}) = 5 \sum_{i=1}^4 x_i - \sum_{i=1}^4 x_i^2 - \sum_{i=5}^{13} x_i \\ \text{subject to} \quad & 2x_1 + 2x_2 + x_{10} + x_{11} - 10 \leq 0 \\ & 2x_1 + 2x_3 + x_{10} + x_{12} - 10 \leq 0 \\ & 2x_2 + 2x_3 + x_{11} + x_{12} - 10 \leq 0 \\ & -8x_1 + x_{10} \leq 0 \\ & -8x_2 + x_{11} \leq 0 \\ & -8x_3 + x_{12} \leq 0 \\ & -2x_4 - x_5 + x_{10} \leq 0 \\ & -2x_6 - x_7 + x_{11} \leq 0 \\ & -2x_8 - x_9 + x_{12} \leq 0 \end{aligned} \quad (17)$$

where $0 \leq x_i \leq 1 (i = 1, \dots, 9)$, $0 \leq x_i \leq 100 (i = 10, 11, 12)$, and $0 \leq x_{13} \leq 1$. $\mathbf{x}^* = (1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 3, 3, 1)$ and $f(\mathbf{x}^*) = -15$.

For this problem, 13 subsystems have different forms of problems. For example, the problem of the first subsystem is

$$\begin{aligned} & \underset{x_1 \in \mathbb{R}}{\text{minimize}} && f(x_1) = 5x_1 - x_1^2 \\ & \text{subject to} && g_1(x_1|x_2, x_{10}, x_{11}) = 2x_1 + 2x_2 + x_{10} + x_{11} - 10 \leq 0 \\ & && g_2(x_1|x_3, x_{10}, x_{12}) = 2x_1 + 2x_3 + x_{10} + x_{12} - 10 \leq 0 \\ & && g_3(x_1|x_{10}) = -8x_1 + x_{10} \leq 0 \end{aligned} \quad (18)$$

References

1. Inalhan, G., Stipanovic, D.M., Tomlin, C.J.: Decentralized optimization, with application to multiple aircraft coordination. In: Decision and Control, 2002, Proceedings of the 41st IEEE Conference, vol. 1, pp. 1147–1155. IEEE (2002)
2. Zhuang, L., Chen, X., Guan, X.: A decentralized ordinal optimization for energy saving of an HVAC system. In: American Control Conference (ACC), 2016, pp. 611–616. IEEE (2016)
3. Shi, W., Ling, Q., Wu, G., Yin, W.: Extra: an exact first-order algorithm for decentralized consensus optimization. *SIAM J. Optim.* **25**(2), 944–966 (2015)
4. Johansson, B.: On distributed optimization in networked systems. Ph.D. Telecommunication (2008)
5. Das, S., Suganthan, P.N.: Differential evolution: a survey of the state-of-the-art. *IEEE Trans. Evol. Comput.* **15**(1), 4–31 (2011)
6. Price, K., Price, K.: Differential Evolution—A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Kluwer Academic Publishers (1997)
7. Yao, X., Liu, Y., Lin, G.: Evolutionary programming made faster. *IEEE Trans. Evol. Comput.* **3**(2), 82–102 (1999)
8. Rosenbrock, H.H.: An automatic method for finding the greatest or least value of a function. *Comput. J.* **3**(3), 175–184 (1960); Floudas, C.A., Pardalos, P.M.: A Collection of Test Problems for Constrained Global Optimization Algorithms. Springer Berlin Heidelberg (1990)
9. Fiacco, A.V., McCormick, G.P.: Nonlinear Programming: Sequential Unconstrained Minimization Techniques, vol. 4. Siam (1990)

Intelligent Building Fault Diagnosis Based on Wavelet Transform and Bayesian Network



Jundong Fu, Luming Huang, Li Chen and Yunxia Qiu

Abstract A novel fault diagnosis method is proposed in this paper in the distribution network based on wavelet transform and a Bayesian network. After the wavelet transform, decomposition, and reconstruction of various electrical basic quantities by amplitude, phase angle, and energy, the electrical fault feature quantity is combined according to various weights, and then the corresponding component switch fault characteristics are calculated by Bayesian. A simple Bayesian fusion of the electrical fault feature and component switch fault characteristics is used as the eigenvector of the Bayesian network, and then trained and predicted by Bayesian network. The experimental simulation results show that the fault diagnosis method for power distribution network based on wavelet transform and Bayesian network proposed in this paper has an obvious recognition degree according to the single fault feature. It is very accurate to identify the type and faulty components.

Keywords Intelligent building · Fault diagnosis · Pattern recognition · Bayesian network · Wavelet transform

1 Introduction

Pattern recognition [1, 2] is a relatively advanced classification method and there is no specific method for its specific classification method. At present, the main methods are decision-making theoretical method, syntax parsing method, and statistical pattern recognition. This paper is used to solve the fault diagnosis of the power supply and distribution network. It adopts decision-making theoretical method and statistical pattern recognition. Its specific classifier is a Bayesian classifier.

At present, the method for fault diagnosis of power supply and distribution network mainly include fault diagnosis method for power supply and distribution network

J. Fu · L. Huang · L. Chen (✉) · Y. Qiu
School of Electrical and Automation Engineering, East China Jiaotong University,
Nanchang, China
e-mail: 8755915@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_39

based on the expert system [3]. But this method is difficult to maintain the system; the knowledge base of various expert systems is complicated to build when a large-scale complex power supply and distribution network fails, poor learning ability, slow reasoning speed, and poor fault tolerance may occur. The network fault diagnosis method of power supply and distribution network based on fuzzy set theory [4] can deal with various uncertain and incomplete fault information in the fault diagnosis of power supply and distribution network, and it has strong fault tolerance, but when it is used, its membership function needs to be set manually, its subjectivity is strong, and the fault diagnosis results also have some errors. The network fault diagnosis method for power distribution based on Petri Network [5, 6] can be immediately used and has strong intuitiveness. But because of the protection action between the power distribution networks is a complex many-to-many relationship, it is necessary to establish multiple Petri net models to analyze these relationships. If this fault information is incomplete or indeterminate, when the fault occurs, the Petri net model will be difficult to build. At the same time, it will also cause its fault-tolerant ability to be poor and the diagnostic accuracy is not high. The network fault diagnosis method for power distribution based on the artificial neural network [7, 8] has a strong ability to learn functions, strong anti-interference capability, parallel processing mechanism, and fast troubleshooting. Of course, when using power supply network fault diagnosis based on the artificial neural network method to do fault diagnosis, a large number of specific and typical fault feature learning samples are required, which will lead to slow convergence of learning and training; if the power supply network fault diagnosis system topology changes, it will require relearning and training. Fault diagnosis method based on Bayesian Networks [9, 10] adopts probability and statistics theory, which is a kind of pattern recognition technology. It uses the prior probability and post-computation probability to make the reliability of the diagnostic results is high; if there is more uncertainty in the fault information of the power distribution network, the fault can be diagnosed effectively and timely, which fully highlights its fault-tolerant ability and high diagnostic capability. This article uses the Bayesian network for fault diagnosis.

2 Bayesian Reasoning

Bayesian reasoning is the process of finding the posterior probability of a node in a Bayesian network based on the prior probability of the root node and the conditional probability of each child node.

In the Bayesian-based power grid fault diagnosis model, the fault set E is given and is a subset of the fault variable information X . Therefore, the fault information given X_i is the probability $P(X_i = x_i | E = e)$ of specifying the fault x_i :

$$P(X_i = x_i | E = e) = P(X_i = 0, E = e) / P(E = e) \quad (1)$$

Equation (1) is a simple and unconstrained Bayesian probabilistic reasoning that can handle probabilistic inference of small random variable sets. In order to satisfy the generality, when its independent conditions are met, its variables are independent of each other and independent of other variables. It can be inferred:

$$P(X_1, X_2, \dots, X_k) = \prod_{i=1}^k P(X_i) = P(X_1)P(X_2) \dots P(X_k) \quad (2)$$

3 Supply and Distribution Network Fault Feature Extraction Method

3.1 Electrical Characteristics Extraction Method

The electrical quantity in this article refers to the information about the electrical fault when the supply and distribution network fails. It generally includes the voltage and current amplitude degree of fault, the phase angle degree of fault, and the energy intensity degree of fault. The three aspects of the fault degree information are subjected to a corresponding wavelet transform, and the weights are determined by the entropy method to obtain the final electrical fault characteristics. The specific calculation and extraction method is as follows (suppose take i components in a fault):

(1) Voltage and current amplitude degree of fault

Before taking the fault, after the wavelet transform of each voltage and current amplitude in the i components, take the comprehensive coefficient maximum value F_{ia} ; then after taking the fault, the voltage amplitudes of the i components are wavelet transformed, take the maximum coefficient Value F_{ib} ; The degree of change in amplitude before and after the failure is:

$$V_i = \max(F_{ia}, F_{ib}) / \min(F_{ia}, F_{ib}) \quad (3)$$

where F_{ia} and F_{ib} are the wavelet transforms of the fault features collected by the component, mainly using the Mallat algorithm. The basic idea of wavelet transform is to find a set of functions that are dense orthogonal bases in the function space of a finite energy signal by a simple translation and stretching or contraction of a function. The essence of wavelet transform is a filtering process on the original signal.

The corresponding degree of fault is

$$x_i = V_i^2 / V_1^2 + V_2^2 + \dots + V_n^2 \quad (4)$$

where n refers to the electrical components.

(2) Phase angle degree of fault

Before taking the fault wavelet transform, the phase angles of the i components, and form the coefficients into a matrix, perform singular decomposition, then take the diagonal matrix Λ_{ia} :

$$\Lambda_{ia} = \text{diag}(\lambda_{1a}, \lambda_{2a}, \dots, \lambda_{na}) \quad (5)$$

After taking the fault wavelet transform, the phase angles of the i components, forms the coefficients into a matrix, and performs singular decomposition, then takes its diagonal matrix Λ_{ib} :

$$\Lambda_{ib} = \text{diag}(\lambda_{1b}, \lambda_{2b}, \dots, \lambda_{nb}) \quad (6)$$

The phase angle degree of fault is:

$$S_i = \sum_{i=1}^n |\Lambda_{ia}(i) - \Lambda_{ib}(i)|/n \quad (7)$$

(3) Energy intensity degree of fault

Before taking the fault, the energy of the i components is subjected to the m -resolution wavelet transform coefficients $\{D_{ja}(k)\}$, $j = 1, 2, \dots, n; k = 1, 2, \dots, n$. After the wavelet transformation m times, the energy e_{ia} before the failure is:

$$W_{ia} = \sum_{j=1}^m \sum_k |D_{ja}(k)|^2 / m \quad (8)$$

$$e_{ia} = W_{ia}^2 / (W_{1a}^2 + W_{2a}^2 + \dots + W_{na}^2) \quad (9)$$

After taking the fault, the energy of the i components is subjected to the m -resolution wavelet transform coefficients $\{D_{jb}(k)\}$, $j = 1, 2, \dots, n; k = 1, 2, \dots, n$. After the wavelet transform m times, the energy e_{ib} before the failure is:

$$W_{ib} = \sum_{j=1}^m \sum_k |D_{jb}(k)|^2 / m \quad (10)$$

$$e_{ib} = W_{ib}^2 / (W_{1b}^2 + W_{2b}^2 + \dots + W_{nb}^2) \quad (11)$$

From this we can get the energy intensity degree of fault e_i :

$$e_i = (e_{ib} - e_{ia})^2 / (e_{ib} + e_{ia})^2 \quad (12)$$

After normalizing the three factors, the entropy method determines its weight, and determines its electrical fault characteristics $C = [C_1, C_2, \dots, C_n]$. The specific implementation process is as follows:

The entropy method determines the weight by calculating the “entropy”, which is to determine the weight of each indicator according to the degree of difference of each feature. When the value of a feature vector of the object to be determined differs greatly, the entropy value is small, indicating that the amount of information provided by the vector value is large and its weight should be large; conversely, if the difference between the values of a feature vector is small, the entropy value is large, indicating that the amount of information provided by the vector value is small and its weight should be small.

After the magnitude X_i , the phase angle amount S_i and the energy e_i collected from the distribution line are subjected to the corresponding wavelet transform decomposition and reconstruction, the following steps are performed to determine the weight.

Step 1: Standardize processing data.

$$X'_{ij} = X_{ij} - \min\{X_j\} / \max\{X_j\} - \min\{X_j\} \quad (13)$$

$$X'_{ij} = X_{ij} - \min\{X_{ij}\} / \max\{X_j\} - \min\{X_j\} \quad (14)$$

Step 2: Calculate the proportion of the feature vector value of the j th eigenvector of the i th component.

$$Y_{ij} = X'_{ij} / \sum_{i=1}^m X'_{ij} \quad (15)$$

Step 3: Calculate the index information entropy.

$$m_j = -k \sum_{i=1}^m (Y_{ij} \times \ln Y_{ij}) \quad (16)$$

Step 4: Calculate the information entropy redundancy.

$$d_j = 1 - e_j \quad (17)$$

Step 5: Calculate the indicator weight.

$$W_i = d_j / \sum_{j=1}^n d_j \quad (18)$$

where X_{ij} represents the value of the j eigenvector of the i electrical component (i.e., one of magnitude X_i , phase angle S_i and energy e_i), $\min\{X_j\}$ and $\max\{X_j\}$ are the minimum and maximum values of the j feature vector of all electrical component,

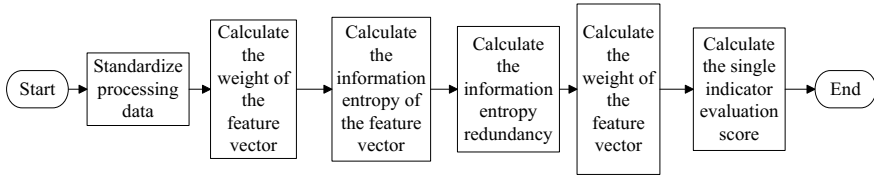


Fig. 1 Weight calculation flowchart

$k = 1/nm$, where m is the number of electrical components and n is the number of feature vectors, here is 3. The flow chart is shown in Fig. 1.

3.2 Component Switch Signal Feature Extraction Method

The switch signal of the electrical components of the power distribution network is generally Boolean, its degree of fault can be calculated by using Bayesian reasoning to determine the probability of each component under a certain fault condition P_i , so it can get the fault degree of its component switch $T = [T_1, T_2, \dots, T_n]$ for:

$$T_i = P_i^2 / (P_1^2 + P_2^2 + \dots + P_n^2) \quad (19)$$

3.3 Comprehensive Feature Extraction Method

In this paper, the comprehensive features of the input in the Bayesian network reasoning for learning and training are the voltage–current amplitude faults processed by the wavelet transform, the phase angle fault degree is processed through the wavelet singularity transformation, and the power fault degree of the wavelet resolution processed by m times of wavelet resolution, three characteristics through the entropy method so that the weight of electrical fault characteristics C_i obtained and the analysis and the Bayesian deduction calculated fault switch component of the degree of fault T_i , and then reconstruct the others to form a new comprehensive characteristics set $M = [M_1, M_2, \dots, M_n]$.

In the field of pattern recognition, information fusion applications based on Bayesian estimation methods are quite extensive. The general information fusion process is: assuming that two sensors start to observe the target M together, and extract the feature information x_1 and x_2 for the information of the two sensors, then the calculation probability of the target M is $P(M|x_1, x_2)$, and then according to the Bayesian estimation can get:

$$\begin{aligned}
P(M|x_1, x_2) &= P(M, x_1, x_2)/P(x_1, x_2) \\
&= P(x_2|M, x_1)p(M, x_1)/P(x_1, x_2) \\
&= P(x_2|M, x_1)P(x_1|M)P(M)/P(x_1, x_2)
\end{aligned} \tag{20}$$

If we assume that the two pieces of information are independent of each other (that is, the electrical fault characteristics and component fault switching quantity we are talking about here), then the above equation can be further written as:

$$P(M|x_1, x_2) = P(x_1, M)P(x_2, M)P(M)/P(x_1)p(x_2) \tag{21}$$

In the information fusion based on Bayesian estimation method, the importance of prior probability is found. When the prior probability is known or derived, the result of information fusion will be more accurate. Therefore, when performing information fusion based on Bayesian estimation, if there is no known prior probability, a reasonable theoretical method should be given for solving the prior probability.

Before the information in the grid fault is merged, the fault information quantity is subjected to the wavelet transform extraction feature, and the feature quantity containing the fault information is formed. Then, the Bayesian estimation based fusion is performed, and the true value representation of the fault state is estimated according to the Bayesian estimation criterion function.

4 Fault Diagnosis Process for Power Distribution Based on Bayesian

This paper proposes that when the power supply and distribution network fails, wavelet transform is performed on the amplitude, phase angle and energy, and the electrical fault characteristic quantities combined by various weights and the simple Bayesian reasoning are used to obtain the switch of each component. The quantity-phase fusion method is used as the characteristic information of the Bayesian network for learning and training, and can be used to complete the fault diagnosis of the power distribution network.

Figure 2 shows a flowchart of fault diagnosis based on Bayesian power supply and distribution:

In intelligent building supply and distribution network, assume that the electrical component is $U = [U_1, U_2, \dots, U_n]$, its main components are bus, circuit breakers, lines; action protection set to $M = [M_1, M_2, \dots, M_n]$, the main components are the probability of failure, the probability of protection from false actions, and the probability of protection against rejection; set electrical degrees of failure Feature $C = [C_1, C_2, \dots, C_n]$; the fault degree of component fault $T = [T_1, T_2, \dots, T_n]$; comprehensive characteristics $M = [M_1, M_2, \dots, M_n]$; fault

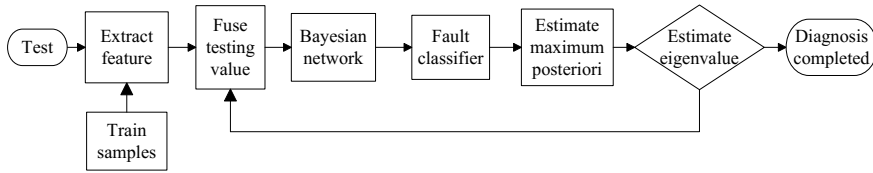


Fig. 2 Bayesian-based power supply and distribution fault diagnosis flowchart

mode $Y = [Y_1, Y_2, \dots, Y_n]$, the selected fault type of the power distribution network in this article is single-phase short circuit, two-phase short circuit, two-phase ground short circuit, and three-phase short circuit.

The main idea is to preprocess the data of the four types of faults and normal state through the abovementioned wavelet fusion and other methods and use it as the learning and training data of the Bayesian network, through the establishment of the Bayesian network and the diagnosis model of fault types and components. Finally, get the purpose and effect of its fault diagnosis. The specific steps are as follows:

Step 1: The voltage and current amplitude, phase angle, the strength of energy and other related basic electrical quantities can be obtained by the building monitoring center.

Step 2: After the wavelet transform is decomposed and reconstructed by the amplitude magnitude, phase angle magnitude, and energy, the electrical fault features are combined according to various weights, and the corresponding component switch fault feature is calculated by Bayesian recommendation at the same time.

Step 3: Combine the electrical fault feature quantity and component switch fault feature obtained in (2) with simple information fusion calculation to obtain the comprehensive fault feature information quantity.

Step 4: Construct a Bayesian network and learn and train the comprehensive fault feature information obtained in (3).

Its learning network is the maximum likelihood as the initial network. Its iterative correction is as follows:

$$P^{(k)}(x_1, \dots, x_n) = \prod_{i=1}^n P^{(k)}(x_i | \pi_i, \theta^{(k)}, G^{(k)}) \quad (22)$$

When there is no zero probability,

$$\overline{x_{mi}} = \begin{cases} x_i^l, 0 < \lambda \leq \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)}) \\ x_i^h, \sum_{j=1}^{h-l} \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)}) < \lambda < \sum_{j=1}^h \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)}) \\ x_i^{r_i}, \lambda > \sum_{j=1}^{r_i-1} \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)}) \end{cases} \quad (23)$$

When there is a zero probability,

$$\overline{x_{mi}} = \begin{cases} x_i^l, 0 < \lambda \leq \omega(1) \\ x_i^h, \sum_{j=1}^{h-l} \omega(j) < \lambda < \sum_{j=1}^h \omega(j) \\ x_i^{r_i}, \lambda > \sum_{j=1}^{r_i-1} \omega(j) \end{cases} \quad (24)$$

among them,

$$\omega(h) = \frac{\bar{P}^{(k)}(x_i^h | \pi_{mi}, D_{(i,m)}^{(k)})}{\sum_{j=1}^s \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)}) + \sum_{j=1}^{r_i-s} \bar{P}^{(k)}(x_i^j | \pi_{mi}, D_{(i,m)}^{(k)})} \quad (25)$$

If $x_{mi} \neq \overline{x_{mi}}$ is in the iterative process,

$$\begin{aligned} \bar{P}^{(k)}(x_{mi} | \pi_{mi}, D_{(i+1,m)}^{(k)}) &= \frac{\bar{P}^{(k)}(x_{mi} | \pi_{mi}, D_{(i+1,m)}^{(k)}) - \frac{1}{N}}{\bar{P}^{(k)}(x_{mi} | D_{(i+1,m)}^{(k)}) - \frac{1}{N}} \\ \bar{P}^{(k)}(x_{mi} | \pi_{mi}, D_{(i+1,m)}^{(k)}) &= \frac{\bar{P}^{(k)}(x_{mi} | \pi_{mi}, D_{(i+1,m)}^{(k)}) + \frac{1}{N}}{\bar{P}^{(k)}(x_{mi} | D_{(i+1,m)}^{(k)}) + \frac{1}{N}} \end{aligned} \quad (26)$$

(5) For a given fault feature, the Bayesian inference prediction using the integrated fault feature information obtained in (3), and fault diagnosis are implemented by checking and estimating the maximum posterior probability. The probability of the type of fault being diagnosed by the equation:

$$\begin{aligned} P(Y_i | M_1, M_2, \dots, M_n) &= P(M_1, M_2, \dots, M_n | Y_i) / P(M_1, M_2, \dots, M_n) \\ &= \alpha P(Y_i) P(M_1, M_2, \dots, M_n | Y_i) \end{aligned} \quad (27)$$

where $\alpha = 1/p(M_1, M_2, \dots, M_n)$, $p(Y_i)$ is the prior probability of class Y_i . It can also be expressed as:

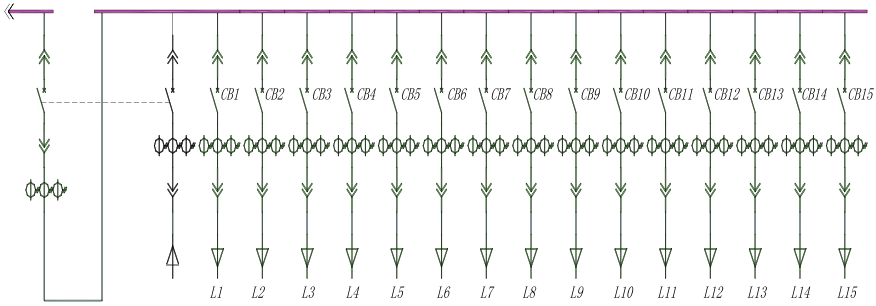


Fig. 3 Layer system diagram of the teaching building

$$P(Y_i|M_1, M_2, \dots, M_n) = \alpha \prod_{i=1}^n P(Y_i)P(M_i|M_1, M_2, \dots, M_{n-1}, Y_i) \quad (28)$$

Its prediction is based on the maximum probability. The process is:

$$\begin{aligned} P(c_{N+1}|D, G_B^h) &= \int P(c_{N+1}|\theta_B, D, G_B^h)P(\theta_B|D, G_B^h)d\theta_B \\ &= \int \prod_{i=1}^N \theta_{ijk} P(\theta_B|D, G_B^h)d\theta_B \\ &= E_{p(\theta_B|D, G_B^h)}(\prod_{i=1}^N \theta_{ijk}) \end{aligned} \quad (29)$$

5 Simulation Studies

5.1 Supply Network Fault Model and Simulation

The layered system of a teaching building is used as a research model. The system diagram is shown in Fig. 3.

To verify the simulation model, assume that the single-phase short-circuit lines are L_7 and L_8 , the circuit breaker is CB_{13} , circuit breaker is CB_{14} , circuit breaker is CB_{15} .

In the Bayesian network learning and training, a total of 60 sets of training data are taken, of which 10 sets of normal operating data and 10 sets of 5 types of fault type data, for a total of 50 sets; for the determination of faulty components, 30 sets of different types of data are selected for learning and training.

Table 1 Layered system of a teaching building

	Line	Bus	Transformer
Component fault	5.1×10^{-4}	4.7×10^{-6}	1.5×10^{-5}
Protection misoperation	1.4×10^{-6}	1.3×10^{-6}	3.3×10^{-6}
Protection refused action	3.7×10^{-4}	0.18	5.2×10^{-4}

5.2 Simulation Results and Analysis

As shown in Table 1, this table is a priori calculation probability calculated for searching various parameters of the layered system of a teaching building.

The procedure of simulation is as follows:

- (1) Processing of data. After decomposing and reconstructing the wavelet transforms of various amplitudes, phase angles, and energies on the layered system lines, the entropy method is used to obtain weights of 0.55, 0.3, and 0.15. The weighted combination of electrical fault features, while Bayesian recommended calculation of the corresponding component switch fault characteristics, extracting its fault characteristics and make information fusion. The fusion results of the values of various fault features are shown in Table 2.

Table 2 The value of each fault feature

Line	Electrical fault feature	Component switch fault feature	Comprehensive fault feature information volume
L ₁	0.1022	0.1001	0.0002
L ₂	0.2016	0.2000	0.0004
L ₃	0.7036	0.0005	0.5003
L ₄	0.0025	0.0003	0.0001
L ₅	0.0005	0.1000	0.0000
L ₆	0.0030	0.0000	0.2000
L ₇	0.9106	0.5431	0.6739
L ₈	0.2587	0.3594	0.3267
L ₉	0.0511	0.0007	0.0009
L ₁₀	0.0001	0.0000	0.0000
L ₁₁	0.0021	0.0002	0.0003
L ₁₂	0.0013	0.0005	0.0009
L ₁₃	0.0001	0.0008	0.0002
L ₁₄	0.0088	0.0003	0.0000
L ₁₅	0.0002	0.0000	0.0001

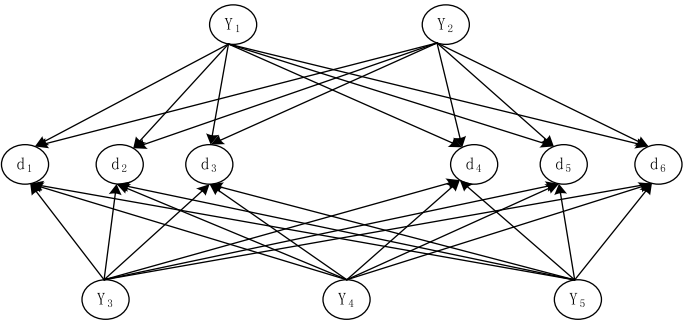


Fig. 4 Model based on Bayesian power distribution network

Table 3 Each feature enters the fault type diagnostic result

Fault type	Fusion maximum posterior probability	Single electrical fault characteristic maximum posterior probability	Single component fault switching maximum posterior probability
Single-phase breaking	0.134	0.162	0.226
Single-phase short circuit	0.856	0.642	0.465
Two-phase circuit	0.105	0.265	0.124
Two-phase short circuit	0.421	0.162	0.432
Two-phase short circuit	0.234	0.364	0.321

- (2) Establish a Bayesian network model. Its fault diagnosis model based on Bayesian power distribution network is shown in Fig. 4.
- (3) Fault diagnosis. For a given fault feature, the Bayesian inference is used to make use of the comprehensive fault feature information, and the fault diagnosis is implemented by checking and estimating the maximum probability.

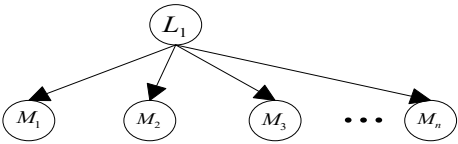
The algorithm of this paper uses a single electrical fault feature quantity as the feature input of Bayes, and compared with the single component fault switch quantity as the input. The simulation diagnosis results are shown in Table 3.

From Table 3 we can see that the maximum posterior probability is 0.856, which is a single-phase short circuit, it is consistent with the previous hypothesis. From the table, we can see that the combination of the two characteristics information used in this paper as a fault feature has a higher fault identification than a single electrical fault feature quantity as a fault feature and single component fault switch quantity as a fault feature, and has a clearer fault identification degree.

- (4) Component diagnostics.

Establish a circuit component Bayesian network diagnostic model. As shown in Fig. 5.

Fig. 5 Bayesian network diagnostic model for line elements



Similarly, a Bayesian network can be constructed to identify line faults, and the posterior probability of each line can be obtained through simulation.

The algorithm of this paper uses a single electrical fault feature quantity as the feature input of Bayes, and compared with the single component fault switch quantity as the input.

As shown in Table 4, it is the posterior probability of each line, that is, which probability of the line is large and which line is faulty.

From Table 4, it can be seen that the fault line are L_8 , L_9 , which are consistent with the previous assumptions. The blending of the two characteristic information amounts used in this paper as a fault feature has a higher degree of fault identification than a single electrical fault feature quantity as a fault feature and single component fault switch quantity as a fault feature, and has a clearer fault identification degree. Although a single fault feature can also be identified, the probability of the line closer to the fault is close to the probability of the fault, which cannot be well identified.

Table 4 Each feature input fault line diagnosis result

Fault line	Fusion maximum posterior probability	Single electrical fault characteristic maximum posterior probability	Single component fault switching maximum posterior probability
L_1	0.115	0.179	0.225
L_2	0.395	0.585	0.622
L_3	0.223	0.752	0.903
L_4	0.151	0.195	0.247
L_5	0.959	0.224	0.114
L_6	0.020	0.305	0.785
L_7	0.852	0.765	0.783
L_8	0.225	0.771	0.601
L_9	0.199	0.650	0.853
L_{10}	0.125	0.377	0.116
L_{11}	0.267	0.214	0.345
L_{12}	0.214	0.521	0.185
L_{13}	0.167	0.298	0.546
L_{14}	0.257	0.199	0.225
L_{15}	0.215	0.179	0.315

In this paper, the blending of two characteristic information quantities is used as the fault feature of the algorithm, it can clearly identify faulty lines.

6 Summary

A novel fault diagnosis method is proposed in this paper in the distribution network based on Bayes and wavelet transformation. After it is decomposed and reconstructed by amplitude, phase angle, and energy, it is combined with various weights. The amount of electrical fault characteristics, Bayes recommends calculating the corresponding component switch fault characteristics. A simple Bayesian fusion of its electrical fault characteristics and component switching fault features is used as a Bayesian network feature vector, which is then trained and predicted by Bayesian networks. Through experiments, the fault diagnosis method of the power supply and distribution network based on Bayesian and wavelet transform proposed in this paper has relatively obvious recognition degree corresponding to a single fault feature, which can accurately identify the type of fault and fault components.

References

1. Dash, P.K., Samantaray, S.R.: A novel distance protection scheme using time-frequency analysis and pattern recognition approach. *Int. J. Electr. Power Energy Syst.* **29**(2), 129–137 (2007)
2. Purkait, P., Chakravorti, S.: Wavelet transform-based impulse fault pattern recognition in distribution transformers. *IEEE Trans. Power Deliv.* **18**(4), 1588–1589 (2003)
3. Meng, X.P., Li, J.L., Zhang, Y.W.: Fault diagnosis of building automation system based on expert system. *Comput. Eng.* **37**(21), 273–275 + 278 (2011) (in Chinese)
4. Liu, X. R., Gao, Y. W., Wang, Z. L.: Method of Power distribution network fault diagnosis based on improved time fuzzy Petri Net. *J. Northeast. Univ. (Natural Science)* **37**(11), 1526–1529 (2016) (in Chinese)
5. Chen, X. Z., Chen, Q., Yu, Y. J., et al.: A fault diagnosis approach of power networks based on maximum likelihood decoding Petri Net models. *Trans. China Electrotech. Soc.* **30**(15), 46–52 (2015) (in Chinese)
6. Gao, Z. Z., Gong, Q. Y., Liu, L. J., et al.: Power system fault diagnosis based on rough set and Petri network optimized by BP algorithm. *Electr. Power* **49**(08), 12–16 + 30 (2016) (in Chinese)
7. Xiong, G. J., Shi, D. Y., Zhu, L., et al.: Fuzzy cellular fault diagnosis of power grids based on radial basis function neural network. *Autom. Electr. Power Syst.* **38**(05), 59–65 (2014). (in Chinese)
8. Lala, H., Karmakar, S.: Continuous wavelet transform and artificial neural network based fault diagnosis in 52 bus hybrid distributed generation system. In: *IEEE Students Conference on Engineering and Systems (SCES) 2015*, pp. 1–6 (2015)
9. Luo, X. H., Tong, X. Y.: Structure-variable bayesian network for power system fault diagnosis considering credibility. *Power Syst. Technol.* **39**(09), 2658–2664 (2015) (in Chinese)
10. Zhao, X., Yang, H. G.: Utility harmonic impedance estimation based on Bayes theorem. *Proc. CSEE* **36**(11), 2935–2943 (2016) (in Chinese)

Fault Location of Distribution Network for Wavelet Packet Energy Moment of Dragonfly Algorithm



Jundong Fu, Jinglin Yue, Li Chen and Tianhang Leng

Abstract The supply and distribution fault location method of wavelet packet energy gray-level moment wavelet neural network based on dragonfly algorithm (DA) is proposed. First, it makes convolutional wavelet energy gray-level moment for faulted electrical information, and then extracts the energy gray-level moment of fault features as the eigenvector of wavelet neural network for training, while the network parameters of wavelet neural network are globally optimized by DA. Finally, compared with the supply and distribution network fault location algorithm of support vector machine based on the DA and the pure wavelet neural network for supply and distribution network fault location algorithm, the experimental simulation shows that the proposed method has faster convergence speed. It also has the advantages of rapid descent speed, high precision, less iterations, rapid convergence, high locating accuracy for faults, and short locating time.

Keywords Fault location · Convolutional wavelet packet energy gray-level moment · Dragonfly algorithm · Wavelet neural network

1 Introduction

It is the fault location of the power distribution network that the fault position can be determined reliably, accurately, and quickly by the corresponding intelligent device and system when the fault occurs in the power distribution network.

At present, the main methods for fault location in a power distribution network are as follows: the fault location method for power distribution network [1] based on adjacent matrix can count the various switches and circuit breakers of the power distribution network in the form of nodes, and then use the matrix form describes the main fault feature information. It has fast operation speed and strong anti-interference

J. Fu · J. Yue · L. Chen (✉) · T. Leng
School of Electrical and Automation Engineering,
East China Jiaotong University, Nanchang, China
e-mail: 8755915@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_40

433

capability. Although the method has a clear locating target and strong practicability, the fault location time is long and the stability is poor due to the blindness of the search algorithm; the fault locating method based on the meta-heuristic intelligence algorithm mainly uses various intelligent optimization algorithms to optimize models. The method has strong search capability and high fault tolerance; the fault locating method based on artificial neural network is regarded as the most potential intelligent fault algorithm. This method has high error tolerance, fast fault locating, and high locating accuracy.

2 Wavelet Neural Network

2.1 Structure of Wavelet Neural Network

Wavelet neural network has a strong fault-tolerant capability and promising approximation ability of the neural network [2]. It also has good frequent domain resolution capability of wavelet subdivision and can accelerate its convergence speed.

There are m input nodes, N output nodes, and n hidden nodes as shown in Fig. 1.

Thus, we can describe the wavelet neural network model as

$$y_i = \sigma \left[\sum_{j=0}^n \omega_{i,j} \psi_{a,b} \left(\sum_{k=0}^m \omega_{jk} x_k(t) \right) \right] \quad (1)$$

where $i = 1, 2, \dots, N$; $\sigma(t) = \frac{1}{1+e^{-t}}$; $net_j = \sum_{k=0}^m \omega_{j,k} x_k$

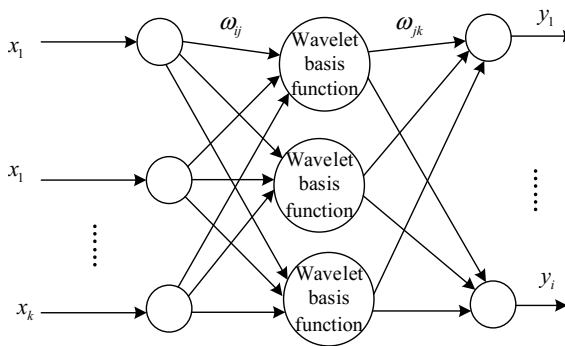


Fig. 1 Topological structure of the wavelet neural network

$$\psi_{a,b}(net_j) = \psi\left(\frac{net_j - b_j}{a_j}\right); y_i = \sigma\left(\sum_{i=0}^n \omega_{ij} \psi_{a,b}(net_j)\right) \quad (2)$$

The objective function is

$$E(\omega) = \frac{1}{2} \sum_{i=1}^N \|Y_i - d_i\|^2 \quad (3)$$

where ω is the weight vector formed by all the weights; d_i is the output vector of the wavelet neural network.

2.2 Training of Wavelet Neural Network

The training method of the wavelet neural network usually adopts the most rapid descent method which is the objective of minimizing the $E(\omega)$ by constantly correcting the weights ω [3].

Solving the partial derivative of the objective function:

$$\frac{\partial E}{\partial w_{ij}} = - \sum_{p=1}^Y (d_j^p - y_i^p) y_i^p (1 - y_i^p) \psi_{a,b}(net_j^p); \quad (4)$$

$$\frac{\partial E}{\partial w_{ij}} = - \sum_{p=1}^p \sum_{i=1}^N (d_j^p - y_i^p) y_i^p (1 - y_i^p) w_{ij} \psi'_{a,b}(net_j^p \delta) x_k^p / a_j \quad (5)$$

$$\frac{\partial E}{\partial a_j} = - \sum_{p=1}^p \sum_{i=1}^N (d_j^p - y_i^p) y_i^p (1 - y_i^p) w_{ij} \psi'_{a,b}(net_j^p \delta) \left(\frac{net_j^p - b_j}{a_j} \right) / a \quad (6)$$

$$\frac{\partial E}{\partial b_j} = - \sum_{p=1}^p \sum_{i=1}^N (d_j^p - y_i^p) y_i^p (1 - y_i^p) w_{ij} \psi'_{a,b}(net_j^p) / a_j \quad (7)$$

Adding additional momentum can reduce chance of shock in learning process and speed up convergence. The updated iterative equation for its weight is

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t); w_{jk}(t+1) = w_{jk}(t) - \eta \frac{\partial E}{\partial w_{jk}} + \alpha \Delta w_{jk}(t); \quad (8)$$

$$a_j(t+1) = a_j(t) - \eta \frac{\partial E}{\partial a_j} + \alpha \Delta a_j(t); b_j(t+1) = b_j(t) - \eta \frac{\partial E}{\partial b_j} + \alpha \Delta b_j(t); \quad (9)$$

3 Power Distribution Network Fault Feature Extraction

3.1 Wavelet Packet Energy Moment Principle

For extracting fault feature of the supply and distribution network, this article uses feature extraction method based on the convolutional wavelet packet energy gray-level moment [4] and the extraction process is

$$x_p^{n,j} = \int_R \frac{x(t)}{2'} \mu\left(\frac{p-t}{2'}\right) dt; 0 \leq j \leq s, 0 \leq n \leq 2^s \quad (10)$$

At this time, the information has become convolutional wavelet packet transform and its decomposition process is

$$x_p^{2n,j+1} = \frac{1}{\sqrt{2}} \sum_{k \in Z} h(K) x_{p-2^j k}^{n,j}; x_p^{2n+1,j+1} = \frac{1}{\sqrt{2}} \sum_{k \in Z} g(K) x_{p-2^j k}^{n,j} \quad (11)$$

After decomposing, the feature information of each frequency band in the last layer is extracted as signal feature, and then the extracted signal features are calculated for its energy gray scale moment [5]. Afterward, the required energy gray scale moment is constructed as feature vector. Finally, the preprocessing of related data is performed on the vectors, and is used as characteristic fault to perform fault location of the power distribution network. The reconstruction process is

$$x_p^{n,j} = \frac{1}{\sqrt{2}} \sum_{k \in Z} \bar{h}(K) x_{p-2^j k}^{2n,j+1} + \frac{1}{\sqrt{2}} \sum_{k \in Z} \bar{g}(K) x_{p-2^j k}^{2n+1,j+1} \quad (12)$$

3.2 Fault Feature Extraction Verification

The teaching building model selected as the simulation object is as follows (Fig. 2).

Take one of the branches as verification of fault feature extraction as shown in Fig. 3.

Assuming specific fault, through the electrical information of corresponding measurement equipment under the corresponding fault conditions, the energy gray-level moment extracted from electrical information is taken as fault feature.

The specific data of each node's convolutional wavelet packet energy gray-level moment is shown in Table 1.

It can be seen from the table that the lower transient resistance value the greater value of the wavelet packet energy gray scale moment. That is to say, its characteristic is obvious. What causes it is that we extract feature vector through the voltage-current flow. When its resistance value is smaller, we extract more information, so that the greater the feature vector, the more obvious the vector.

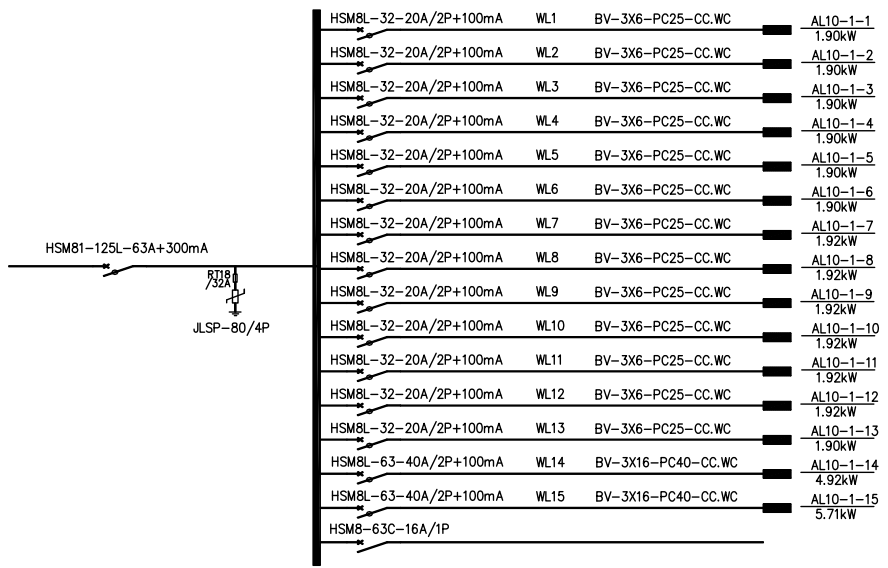


Fig. 2 Low-voltage floor distribution system diagram

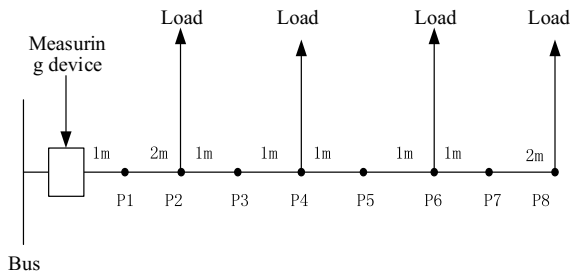


Fig. 3 Trunk for distribution lines

4 Fault Locating of Wavelet Neural Network Based on DA

4.1 Dragonfly Algorithm

Research shows that any populated behavior follows three basic guidelines: Separation, Alignment, and Cohesion. Separation aims to avoid static collisions between adjacent individuals; alignment refers to the speed at which individuals maintain consistency with adjacencies; cohesion refers to individuals tending to converge toward the center of adjacencies. In order to survive, all animals are instinctively close to food sources and away from natural enemies. Therefore, updating of individuals' location is related to these five behaviors.

Table 1 Node-level wavelet packet energy gray scale moment

Load point	Transition resis- tance/(Ω)	Energy gray-level moment node/($\mathbf{W} \cdot \mathbf{m}$)	Load point	Transition resis- tance/(Ω)	Energy gray-level moment of each node/($\mathbf{W} \cdot \mathbf{m}$)
P1	1	0.5717	P5	1	0.3759
P1	10	0.3014	P5	10	0.0891
P1	100	0.1146	P5	100	0.0789
P2	1	0.5231	P6	1	0.3276
P2	10	0.2658	P6	10	0.0885
P2	100	0.1102	P6	100	0.0658
P3	1	0.4725	P7	1	0.2752
P3	10	0.2046	P7	10	0.0880
P3	100	0.1014	P7	100	0.0452
P4	1	0.4256	P8	1	0.1572
P4	10	0.1236	P8	10	0.0875
P4	100	0.0945	P8	100	0.0214

The mathematical method can be expressed as follows:

(1) Degree of separation

$$S_i = \sum_{j=1}^N X - X_j \quad (13)$$

(2) Degree of alignment

$$A_i = \frac{\sum_{j=1}^N V_j}{N} \quad (14)$$

(3) Degree of cohesion

$$C_i = \frac{\sum_{j=1}^N X_j}{N} - X \quad (15)$$

(4) Food attraction

$$F_i = X^+ - X \quad (16)$$

(5) Natural enemy repulsion

$$E_i = X^- - X \quad (17)$$

where X is the current position of the individual; X_j is the position of the j th neighbors; N is the number of individual xx adjacent to the i th; V_j is the speed of the j th neighbors; X^+ is the location of food source; and X^- is the location of natural enemy.

The next step is to calculate the position and update step size.

$$\Delta X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta X_t; X_{t+1} = X_t + \Delta X_{t+1} \quad (18)$$

where t is the current iteration; i is the i th individual; X_t is the individual location of T -generation population; ΔX_{t+1} is the increment of the next generation; X_{t+1} is the individual location of the next population; s is the separation weight; S_i is the separation of the i th individual; a is the alignment weight; A_i is the alignment of the i th individual; c is cohesion weights; C_i is the cohesion of the i th individual; f is the food factor; F_i is the attraction of the food position to the i th individual; e is the natural enemy factor; E_i is the repulsive force of the natural enemy position on the i th individual; and w is the inertia weight.

What conditions are satisfied by individuals? To solve this problem, Seyedali Mirjalili assumed that each individual is at the center of a circle with radius r . If the Euclidean distance between individual i and individual j is smaller than r , then individual i is considered to be adjacent to individual φ . The radius r increases with the number of iterations in order to speed up the convergence. When none is adjacent to the individual i , the Lévy flight random walk method is introduced to improve the random features of the searching. The walking step satisfies a heavy-tailed Lévy distribution and is calculated as

$$L\acute{e}vy(x) = 0.01 \times \frac{r_1 \times \sigma}{|r_2|^{\frac{1}{\beta}}} \quad (19)$$

where r_1 and r_2 are random numbers between zero and one, β is a constant (here is $3/2$), σ can be calculated as follows:

$$\sigma = \left(\left(\Gamma(1 + \beta) \times \sin\left(\frac{\pi\beta}{2}\right) \right) / \left(\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)} \right) \right)^{\frac{1}{\beta}} \quad (20)$$

when no individual is adjacent, the population location is updated as follows:

$$X_{t+1} = X_t + L\acute{e}vy(d) \times X_t \quad (21)$$

4.2 Wavelet Neural Network Fault Location Process Based on Dragonfly Algorithm

The wavelet neural network fully embodies the time–frequency localizing features and the self-learning ability of the neural network of wavelet analytical theory. There is a disadvantage that the network will easily fall into local minimum if the initial weights and thresholds are randomly assigned which lead to the instability of network training. So, it adopts DA to optimize the parameters [6]. The basic idea of the wavelet neural optimization based on the DA is that the weights and thresholds connecting each other are used as individual position vectors in the DA. The optimal values of the weights and thresholds are sought by the individual position algorithm. Taking into account, all possible factors of group behavior makes DA quickly converge to near the optimal value and have good global search ability and stability. Based on the above points, it is applied to the parameter optimization of the wavelet neural network to find the most suitable training network parameter while ensuring the minimum training error.

Its adaptation function is as follows:

$$f = \frac{1}{1 + E(N)} \quad (22)$$

The specific locating process is as follows:

- (1) Initialize the parameter settings. Set the number of dragonfly and the maximum number of iterations.
- (2) Set the learning task of the wavelet neural network.
- (3) Initialize dragonfly position of X and position change step of ΔX .
- (4) Calculate each value of weight. According to the initial value in(1), initialize adjacent radius r , inertia weight w , separation weight s , alignment weight a , cohesion weight c , food factor f , and enemy factor e .
- (5) Calculate the fitness value. The function constructed based on the error function is taken as the current fitness function of each individual. As soon as dragonfly conducts behavioral activities, the maximum adaptation of the current individual is calculated. Then, if it is greater than the saved fitness value, replace the saved optimal fitness value with the current value and use it as the current optimal value. Simultaneously, keep the dragonfly parameters corresponding to the current optimal value, otherwise, the original fitness value and its corresponding parameter are still saved.
- (6) Use the Euclidean distance to determine if there are neighbors. If it exists, the position update is calculated by Eq. (18), otherwise use Eq. (21).
- (7) Calculate the degree of behavior. According to equation from (13) to (17), the degree of separation, alignment, degree of cohesion, food attraction, and repulsion of natural enemies were calculated in order.
- (8) Location update. Use (18) to update the position and step size.

- (9) Determine whether the maximum number of iteration is satisfied. If reached, output maximum fitness value of dragonfly and parameter corresponding to the optimal value, otherwise, increase the number of iteration and return to execute the fourth step.

5 Simulation Studies

5.1 Model and Simulation

The article uses the DA to optimize the training process of wavelet neural network in order to get the optimal network parameters.

Suppose the single-phase ground fault has been diagnosed and the line is 24 m, randomly select five groups of transition resistance from 0 to 100 Ω from the database as training samples and specifies fault phase angle of 30°. Continuously, randomly select five groups of distribution network dataset from 0 to 100 Ω with the fault phase angle of 45° as the test sample (1,10, and 100 are selected here). First, the fault feature electrical information is extracted by the three-layer wavelet packet on the condition of the frequency of 50 Hz. Then the convolutional wavelet packet energy gray scale moment is calculated, and it takes the fault feature information formed as the input of the wavelet neural network. Finally, it uses the algorithm procedure to precisely locate the fault.

5.2 Results and Analysis

In order to clearly understand the fault location algorithm of the wavelet neural network based on the DA, in this section, it compares with the fault location of the support vector based on the DA. The training error curve is shown in Fig. 5.

The locating results of each fault location algorithm are shown in Tables 2, 5, and 8. The relative error of algorithm are shown in Tables 3, 6 and 9, where *DA-WNN* is the fault location algorithm of the wavelet neural network based on the DA; *DA-SVM* is the fault location algorithm of the support vector machine based on the DA; and *WNN* is the pure wavelet neural network fault location algorithm (Tables 4, 7, 10).

Table 2 Locating result of 1 Ω

Fault distance/m	DA-WNN/m	DA-SVM/m	WNN/m
0.8	0.81	0.82	0.86
4.9	5.00	5.01	5.28
12.3	12.52	12.44	12.87
15.0	15.36	15.47	15.93
18.6	18.92	19.18	19.78
21.0	21.21	21.22	21.97
23.1	23.23	23.34	23.60

Table 3 Relative error of 1 Ω

Fault distance/m	DA-WNN/100%	PSO-WNN/100%	WNN/100%
0.8	1.35	2.92	6.18
4.9	2.06	2.24	5.65
12.3	1.76	1.12	2.79
15.0	2.43	3.16	3.67
18.6	1.73	3.10	4.55
21.0	0.99	1.07	3.58
23.1	0.55	1.02	1.62
error	1.55	2.09	4.01

Table 4 Locating time of 1 Ω

Fault distance/m	0.8	4.9	12.3	15.0	18.6	21.0	23.1	Average time
DA-WNN/s	29.42	28.94	29.19	29.59	29.35	29.42	29.46	29.34
PSO-WNN/s	44.56	43.97	44.44	45.31	44.16	44.56	45.91	44.70
WNN/s	76.54	76.46	77.41	75.66	77.00	75.66	76.21	76.42

Table 5 Locating result of 10 Ω

Fault distance/m	DA-WNN/m	DA-SVM/m	WNN/m
0.8	0.81	0.83	0.86
4.9	5.01	5.02	5.21
12.3	12.54	12.45	12.68
15.0	15.40	15.53	15.61
18.6	18.95	19.24	19.55
21.0	21.22	21.24	21.84
23.1	23.22	23.35	23.50

Table 6 Relative error of 10 Ω

Fault distance/m	DA-WNN/100%	PSO-WNN/100%	WNN/100%
0.8	1.46	3.23	6.95
4.9	2.26	2.46	6.34
12.3	1.92	1.19	3.09
15.0	2.68	3.51	4.09
18.6	1.88	3.44	5.09
21.0	1.04	1.13	3.99
23.1	0.55	1.08	1.76
error	1.68	2.62	4.47

Table 7 Locating time of 10 Ω

Fault distance/m	0.8	4.9	12.3	15.0	18.6	21.0	23.1	Average time
DA-WNN/s	30.75	30.25	30.52	30.94	30.69	30.75	30.80	30.67
PSO-WNN/s	46.59	45.97	46.46	47.37	46.16	46.59	48.00	46.73
WNN/s	80.02	79.93	80.93	79.10	80.50	79.10	79.67	79.89

Table 8 Locating result of 100 Ω

Fault distance/m	DA-WNN/m	DA-SVM/m	WNN/m
0.8	0.81	0.83	0.85
4.9	5.01	5.02	5.21
12.3	12.54	12.46	12.68
15.0	15.41	15.54	15.62
18.6	18.97	19.25	19.56
21.0	21.23	21.25	21.85
23.1	23.24	23.37	23.52

Table 9 Relative error of 100 Ω

Fault distance/m	DA-WNN/100%	PSO-WNN/100%	WNN/100%
0.8	1.53	3.31	7.02
4.9	2.33	2.54	6.41
12.3	2.00	1.26	3.16
15.0	2.76	3.58	4.17
18.6	1.96	3.51	5.17
21.0	1.12	1.21	4.06
23.1	0.63	1.16	1.83
error	1.53	2.32	5.11

Table 10 Locating time of 100 Ω

Fault distance/m	0.8	4.9	12.3	15.0	18.6	21.0	23.1	Average time
DA-WNN/s	33.42	32.88	33.17	33.62	33.35	33.42	33.47	33.67
PSO-WNN/s	32.88	49.97	50.49	51.48	50.17	50.64	52.17	51.12
WNN/s	86.97	86.88	87.97	85.97	87.50	85.97	86.59	86.68

Transient resistance values are determined in advance, of which 100 Ω is the most common. The wavelet energy gray scale moment feature vector value of 1 and 10 Ω are large. Comparing the three kinds of transient resistance values of 1, 10, and 100 Ω , for the same distance, the wavelet energy gray scale moment feature vector value of 1 and 10 Ω are larger than 100 Ω . That is to say, its characteristic is obvious. So the greater extracted feature vector, the higher locating accuracy and less locating time.

We can see from the data analysis in Fig. 4 that the network training error of the power distribution fault locating algorithm based on the wavelet neural network of the DA is from the beginning of 0.3 close to 0.05, and the training error of its algorithm finally stabled at 0.05 after 80 iterations. However, the training error of the support vector machine based on DA is close to 0.15, and the network training error finally stabled at 0.15 after 150 iterations. The network training error of the pure wavelet neural network is close to 0.15 after 200 iterations, but at this time, the network error does not converge steadily. It can be seen from the data analysis in the table three and four that the average relative error of the locating results of each fault point of the wavelet neural network fault location algorithm based on the DA is 1.5304%, and the average time required for fault point location is 33.6772s. The fault location algorithm based on the support vector machine based on DA has an average relative error of 2.3251% for each fault location, and the average time for fault point location is 51.1241s. The fault location algorithm based on pure wavelet neural network has an average relative error of 5.1101% for each fault location, and the average time required for fault point location is 86.6681s. In summary, the fault location algorithm based on the wavelet neural network proposed by this chapter has the advantages of high accuracy, less iterations, and rapid convergence. It also has high locating accuracy and short locating time (Fig. 5).

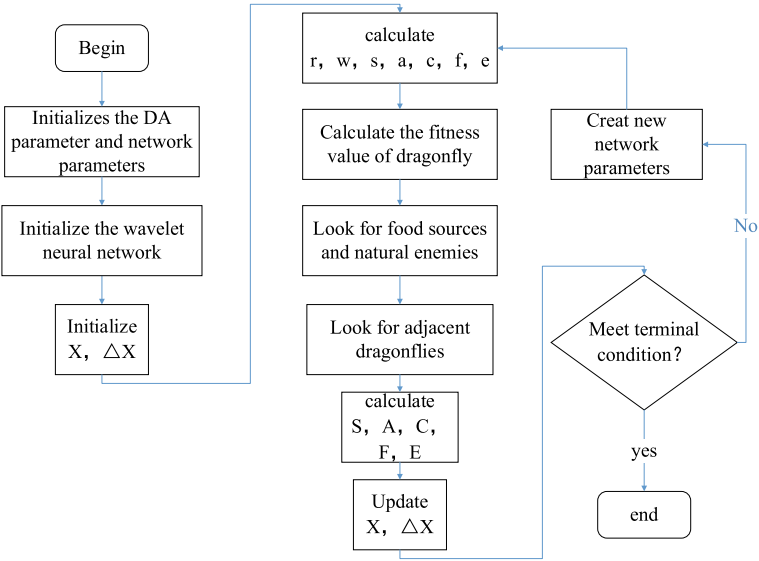


Fig. 4 DA algorithm flowchart

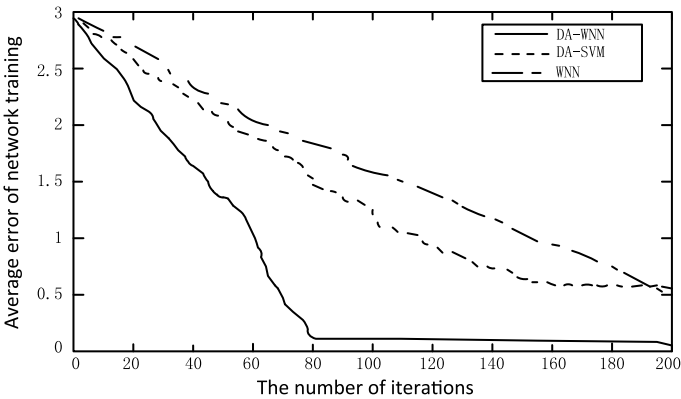


Fig. 5 Training error curve

6 Conclusions

This article proposed a fault location algorithm for power distribution based on the wavelet neural network based on DA [7]. First, it performed convolutional wavelet energy gray-level moments on the faulted electrical information, and then extracted the energy gray scale moments of its fault features as eigenvectors of wavelet neural network to train on the network. The network parameters of the wavelet neural network are globally optimized using the DA. Experimental simulations showed that

comparing with the supply and distribution network fault location algorithm of support vector machine based on the DA and the pure wavelet neural network for supply and distribution network fault location algorithm, the supply and distribution fault location algorithm of wavelet neural network based on the DA has faster convergence speed. The training error of the wavelet neural network optimized by DA has advantages of falling fast, high precision, less iterations, rapid convergence, high locating accuracy for faults, and short locating times. Compared with its algorithm, fault location algorithm of pure wavelet neural network and fault location algorithm of support vector machine based on DA do not have the advantages of fast convergence, high precision and short locating times.

References

1. Mirjalili, S.: Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **27**(4), 1053–1073 (2016)
2. Daely, P.T., Shin, S.Y.: Range based wireless node localization using Dragonfly Algorithm. In: Eighth International Conference on Ubiquitous and Future Networks. IEEE (2016)
3. He, Y., Tan, Y., Sun, Y.: Wavelet neural network approach for fault diagnosis of analogue circuits. *IEE Proc. Circ. Dev. Syst.* **151**(4), 379–384 (2004)
4. Peng, B.: Fault diagnosis of rotating machinery based on improved wavelet neural network. *J. Vib. Meas. Diagn.* **27**(1), 32–35 (2007)
5. Lin, F.J., Shieh, H.J., Huang, P.K.: Adaptive wavelet neural network control with hysteresis estimation for piezo-positioning mechanism. *IEEE Trans. Neural Netw.* **17**(2), 432–444 (2006)
6. Dong, L., Xiao, D., Liang, Y., et al.: Rough set and fuzzy wavelet neural network integrated with least square weighted fusion algorithm based fault diagnosis research for power transformers. *Electr. Power Syst. Res.* **78**(1), 129–136 (2008)
7. Isermann, R.: Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance, vol. 28, no. 2, pp. 195–197 (2006)

Graphical Programming Language Design for Decentralized Building Intelligent System



Shuo Zhao, Qiliang Yang, Jianchun Xing and Guangtong Xue

Abstract The decentralized building intelligent system is a new type of building intelligent control system, which forms a decentralized network through intelligent nodes. It solves the difficulties in the configuration, modification, and information sharing of traditional centralized building control systems. However, because of the characteristics of no-center and parallelism, the decentralized system still lacks effective programming language support. In this regard, this paper designs a dedicated graphical programming language for the decentralized building intelligent system, then describes and explains from the static definition and dynamic interaction of graphics element two aspects. Finally, combining a case of control, this paper draws the connection diagram with the graphical programming language and makes an explanation, which verifies the feasibility of the graphical programming language for the decentralized system.

Keywords Decentralized · Graphical programming language · Graphic element · Building members · Building community

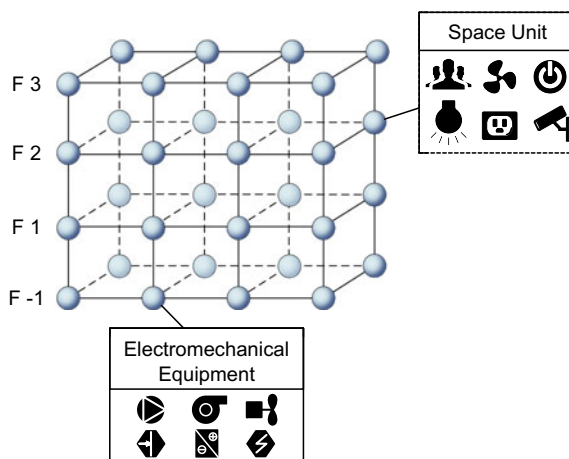
1 Introduction

The decentralized building intelligent system is a new type of intelligent building control system. It regards the entire building and its electromechanical system as a decentralized network connected by building space units and large electromechanical equipment. Figure 1 shows the structure of the decentralized building intelligent system. Each building space unit and electromechanical equipment corresponds to an intelligent node called computing process node (CPN), which contains standard data sets to integrate and manage various types of building information, and can bring standardized production and massive copying. The CPNs are connected on a plug-

S. Zhao · Q. Yang (✉) · J. Xing · G. Xue
College of Defense Engineering, The Army Engineering University of PLA,
Nanjing, China
e-mail: yql@893.com.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_41

Fig. 1 The structure of the decentralized building intelligent system



and-play basis according to the topological relationship of the building space to form a network and form a decentralized distributed computing platform. It effectively solves the difficulties faced by traditional centralized building control systems [1, 2] and satisfies the new needs of modern intelligence buildings, bringing great potential for development and application in building automation.

Automatic control of traditional buildings is mainly implemented by programmable logic controllers (PLC). There are three kinds of PLC graphical programming languages being defined in the IEC61131-3 standard [3, 4], which are widely accepted and applied. However, these programming languages are only oriented to electromechanical components and not intuitive and friendly in terms of programming logic and methods and it requires some programming foundation and learning cost when used.

Therefore, the graphical programming language of conventional building automation and the existing graphical programming methods are not able to achieve a good development and programming of the decentralized system, and the decentralized building intelligent system still lacks efficient, practical, convenient, and friendly programming language support.

2 Requirements of Decentralized System Graphical Programming Language

Based on the decentralized building intelligent system and its own capabilities, its graphical programming language should own the following characteristics.

Friendliness: Engineers involved in the field of intelligent buildings are mainly noncomputer professionals. To reduce the barriers between intelligent buildings and computer programming, an operating-friendly programming language is prerequi-

site. Friendly development means the human-centered interaction process [5], which fully considers the user's background and provides vision, thinking, and operation methods in simple, reasonable, easy way.

Openness: Based on plug-and-play features, it requires that the system development environment is open, not only open to new devices, but also structural engineer so that the system can support extensions and additions efficiently and easily.

Event Driven: The thinking way of non-programming professionals is based on control flow, that is, discovering events and handling events [6]. Intelligent buildings have a large number of sensors and controlled objects, which requires the system to use these devices to implement programming data flow with event-driven, making the logic and control strategy simple and clear. Programming in an event-driven data flow mode makes the logic simple, and the control strategy clear at a glance.

Network parallelism: In decentralized network, each CPN is an intelligent node with computing capability, and the computing power of the entire system is distributed throughout the building with no central processing unit. So, the decentralized network has powerful parallel computing capabilities, which requires the development of the system programming can effectively use the capabilities, and represent and implement task processing based on the parallel network.

3 Graphical Programming Language Design

3.1 Graphic Elements Design

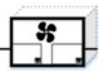
The basis of the graphical programming language is the graphic elements, corresponding to the smallest graphical visible entity in the programming interface [7]. This section combines the structure and composition of the architecture-free system, and graphically describes and formalizes the graphic elements.

Classification of Graphic Elements

Based on the composition of the physics world of intelligent building, the elements of reality can be mapped to the elements of the programming language [8]. Sensors and actuators, CPNs, functional subsystem are controlled objects of three different levels in decentralized buildings, from which making classification of the graphic elements [9]. Classification and examples of graphic elements are shown in Table 1.

- (1) **Member Block (MB):** To correspond to the architectural building space unit and large-scale electromechanical equipment in the decentralized system, and also to the basic division of the CPN. It mainly includes six categories and all kinds of member blocks can be further refined into more specific blocks. For example, the air-conditioning unit includes air-conditioning box, cold/heat source, water pump, fan, heat exchanger, controller, etc.
- (2) **Community Block (CB):** To correspond to the sub-nets that implement relatively independent and complete function in decentralized building sys-

Table 1 Classification and examples of graphic elements

Element categories	Modules	Basic graphics	Examples
Member block	Six categories: space units, air-conditioning, water system, power distribution, fire fighting and security, which can be further refined based on devices		 Air conditioning member block
Community block	With two modes: read/write mode and para-region mode		 Read/write mode  Para-region mode
Variable block	Analog input variable block, analog output variable block, switch value input variable block, switch value output variable block		 Analog input variable block  Switch value output variable block
Operate block	Mathematical operations, logic operations, and advanced network algorithms		 Sum
Process block	Delay, loop, interrupt, select, jump, time count, number count, synch-send and synch-receive		 Select

tem. The community can be manually configured by setting its members, or configured automatically by identifying and interacting with parameters. The community block contains two modes, corresponding to different usage conditions: Read/Write mode, used for uniformly reading and writing the data of members; Para-region mode, community block unfolds to form a para-region and the programming operations in the para-region will be concurrently executed in all members of the community for network parallel operations. The form of its expression is shown in Fig. 2.

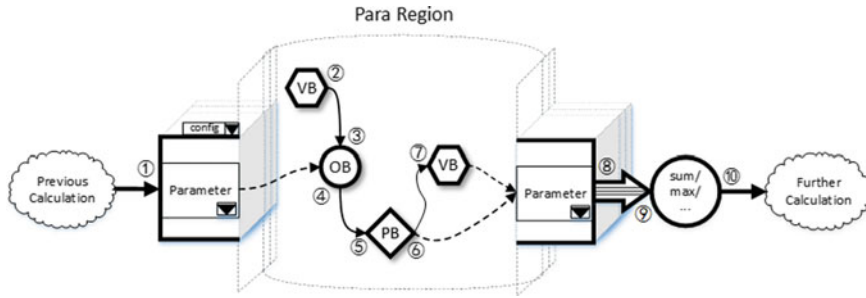


Fig. 2 The concept connection diagram of a program

- (3) **Variable Block (VB)**: To correspond to the physical quantities in intelligent buildings, representing different sensors and actuators, which can be configured according to attributes in standard data sets of member block.
- (4) **Operate Block (OB)**: To complete the most basic calculations such as mathematical operations, logical operations, and advanced network algorithms.
- (5) **Process Block (PB)**: To intuitively support system programming and display the process flow of data progression for explicitly presenting the control flow.

Synch-send and synch-receive are designed for parallel network computing and used to exchange information between members in the community. When synchronously sending, each member sends the data to all neighbors at the same time; when synchronously receiving, each member simultaneously receives all data from neighbors to the local.

Definitions of Graphic Elements

Each kind of graphic elements can be represented by a multiple array:

$$MB = \langle \text{Basic}, \text{SDS}, \text{IF} \rangle; CB = \langle \text{Basic}, \text{SDS}^*, \text{IF} \rangle$$

$$VB = \langle \text{Basic}, \text{SDS}, \text{IF} \rangle; PB = \langle \text{Basic}, \text{Pro}, \text{IF} \rangle; OB = \langle \text{Basic}, \text{Ope}, \text{IF} \rangle$$

SDS is the mapping of standard data set in CPN corresponding to member block; SDS* is the mapping of the common part of the standard data set of the members within the community; IF is the interface group that transfers the data between the element; Pro and Ope are the specific operation code of process control and calculation; Basic is the fundamental elements group, and it can be represented by a multiple array:

$$\text{Basic} = \langle N, A, I, Lo, L, O, C, F \rangle$$

N—Name; A—Attribute; I—Image; Lo—Location; L—Link; O—Operation, identifying the changed amount, to update information; C—Code, corresponding

to the code calling method of graphic element's—Flag, store status flags such as selection, connection, and conversion.

3.2 *Interface of Graphic Elements*

The interface refers to the input and output of the graphic element, and is the basis of the graphic element connection and the channel for data exchange. Depending on the function of different graphic elements, the interfaces can be divided into two categories according to different data types.

Single-line Interface

Existing in the graphic elements with unidirectional single-line connection, and the connection line is represented by a single directed solid line. The input and output of most graphic elements except community block and a part of operator block generally belong to single-line interface.

Multi-line Interface

Existing in the graphic elements associated with parallel operations, mainly referring to community block and a part of operator block. The connecting line is represented by an arrow with several parallel lines and the input or output data is a specific vector, in which an element represents related data of a member in the community.

Figure 2 shows the concept connection diagram of a program. The community block passes the received data to all members in the community through the single-line input interface (①); Each member synchronously executes the programs in the parallel region, in which the graphic elements are connected to each other with single-line interfaces (②–⑦); After the parallel operation, all members summarize the output results to the multi-line output interface (⑧) of the community block with a output vector, one element in the vector represents a member's operation result; The output vector enters the multi-line input interface (⑨) of some specific operator block to realize the processing of summation, maximizing and so on, and the output data through the single-line type output interface (⑩) of the operator block continues to pass backward.

4 Graphical Programming Case

Case Description

The case is constant-pressure control of decentralized building water supply system, taken from the literature [2]. The water supply system is one of the most basic systems in buildings. In order to ensure the quality of water supply for large buildings and avoid the problem of insufficient water pressure caused by large-scale horizontal and vertical pipe networks, the water supply system is equipped with a pressure pump. The pump controls its revolving speed according to the return value of the water

pipe pressure at each place and completes the increase or decrease of the pressure to ensure the water supply is safe and reliable. The ends of the water pipe network are distributed within the building space unit. Although it is not possible to measure the pressure values at every position of the water pipe, the measurement values of each end can be obtained through the standard data set of the space unit. The water pump can get the values transmitted through the network to monitor and adjust the pressure of the water pipe network.

Control Strategy

If the water pressure at the end is lower than the set value and continues for a period of time, the pump revolving speed will be increased. When the amount of the low-pressure ends increases, the revolving speed should be increased more; analogously, the pressure reduction control has a contrary progress. The pump revolving speed is proportional to the sum of the deviation between water pressure and the set value.

Because of the large scale of entire water supply network, it is very difficult to program for the whole network containing the pump and all the water supply ends, and the distribution of ends in the network is very extensive, so it is difficult to configure and program one by one. However, the quantity of the pressure pumps is very small, and the pumps connect to the water pipe network. So, it is very easy to program the pump and control the system with the community and para-region.

Programming Process Description

Figure 3 is the graphical programming diagram of the pressure reduction control of the decentralized building water supply system. The procedure for running the program is as follows.

- (1) Set the timer to make the pump periodically initiate water pressure testing of the water pipe network. When the timing triggers, the pump initiates the parallel interaction, constructs the water pipe network community through the configuration and the standard data set parameters, and distributes the set value of water pressure to the members.
- (2) All members simultaneously execute the program in the para-region: compare the water pressure set value and the measured value P , and output 1 to AND block when $P > P_{set}$, otherwise output 0; After the timer times out for 3 min, the comparison between P_{set} and P is again triggered, and the result value is output to the AND block for judging whether the water pipe end is continuously in an overpressure state.
- (3) After AND block obtaining the two values, the AND result is output to the variable S . If the pipe end sustains overpressure state, $S = 1$; otherwise, $S = 0$.
- (4) The end community block aggregates the value of the variable S of all members into a summation block, to judge the total overpressure degree in the community.
- (5) After the subsequent calculation, the total overpressure degree of the water pipe network is converted to the correction value of pump revolving speed, and the pressure reduction control of the water supply system is realized. The pressure raise control of water supply system is similar.

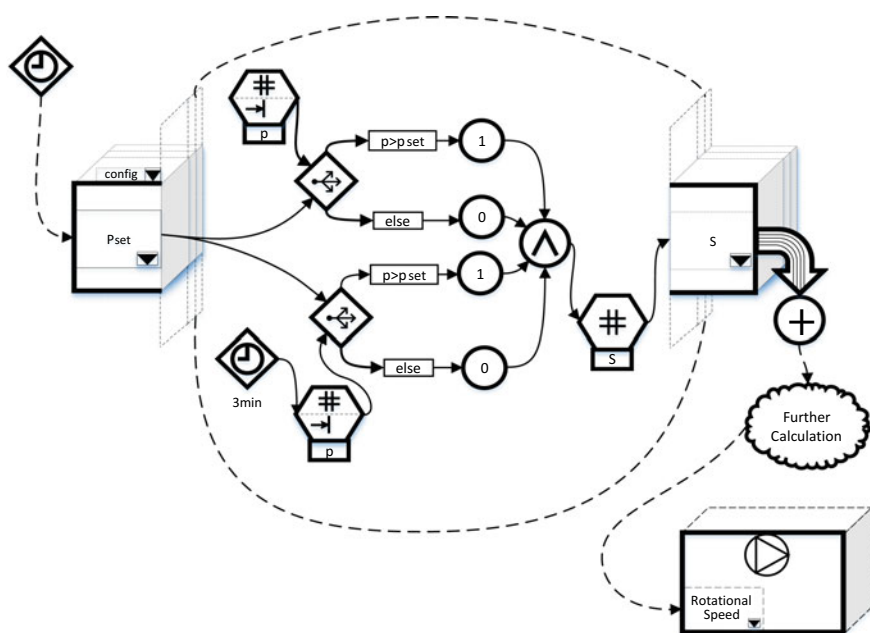


Fig. 3 The graphical programming diagram of the pressure reduction control of the decentralized building water supply system

The control strategy can be uniquely represented using this graphical programming language, and can meet the changes in the actual requirements by changing the use and connection of the Graphic Elements, indicating the feasibility of this language.

5 Conclusions and Outlook

It is designed, in this paper, a kind of dedicated graphical programming language based on the decentralized building intelligent system combining with the characteristics of the system and related engineers, which supports friendly human-computer interaction and parallel computing, and can effectively reduce the difficulty and the cost of learning of development and programming. However, this is only a tentative theoretical design and there are also many deficiencies in the completeness of the theory. Much work centering on the elements need to be further developed, such as designing a complete elements library, adapting to more network algorithms and many more. Another important task is the transformation between graphical and textual languages, and developing a complete development environment to achieve a real sense of graphical programming and development.

References

1. Dai, Y.C., Jiang, Z.Y., Shen, Q., Chen, P.Z., Wang, S.Q., Jiang, Y.: A decentralized algorithm for optimal distribution in HVAC systems. *Build. Environ.* **95**, 21–31 (2016)
2. Shen, Q.: Studies on architecture of decentralized system in intelligent building. Ph.D. Dissertation, Tsinghua University (2015) (in Chinese)
3. International Electrotechnical Commission. IEC61131–3 Programmable Controllers—Part 3: Programming Language, 2nd edn. (2003)
4. Thayer, T.: Speaking in tongues: understanding the IEC61131-3 programming languages. *Control Eng.* **56**(1), 47–47 (2009)
5. Chang, C.K., Yang, Y.F., Tsai, Y.T.: Exploring the engagement effects of visual programming language for data structure courses. *Educ. Inf.* **33**(3), 187–200 (2017)
6. Christoph, R., Herbert, K.: A process-oriented modeling approach for graphical development of mobile business apps. *Comput. Lang. Syst. Struct.* **53**, 43–58 (2018)
7. Kleppe, A.: *Software Language Engineering: Creating Domain-Specific Languages Using Meta-models*. Pearson Schweiz Ag. (2009)
8. Hsiao, W.S., Lee, C.H., Yang, M.H., Chen, R.Q.: User interface based on natural interaction design for seniors. *Comput. Hum. Behav.* **75**, 147–159 (2017)
9. Spinellis, D.: Notable design patterns for domain-specific languages. *J. Syst. Softw.* **56**(1), 91–99 (2001)

Insect Intelligent Building (I²B): A New Architecture of Building Control Systems Based on Internet of Things (IoT)



Qianchuan Zhao and Ziyang Jiang

Abstract This paper introduces Insect Intelligent Building (I²B), a new architecture of building control systems based on Internet of Things (IoT). It has the following key features that are missing in the traditional solutions: (1) it has built-in agent models for space units and control devices; (2) it has built-in link models for immediate connections between space units and control devices; (3) it assigns a smart controller (called insect) for each space unit or each control device with six data ports (called legs) and these controllers communicate with each other only through their data ports according to the link models; (4) it decompose various building operation/management commands into computing tasks that run on smart controllers and depends on the collaborations among neighboring controllers to achieve the desired effects of the commands. Different from the traditional centralized solutions, this new architecture allows the building control systems to be developed in before the buildings are really built and simplify the installation and configuration of the building control systems.

Keywords Building control · Peer-to-Peer (P2P) · Insect Intelligent Building (I²B) · Internet of Things (IoT) · Distributed optimization

1 Introduction

In this research work, the building control systems as shown in Fig. 1 have become an indispensable part of smart buildings. These systems work together to provide

Q. Zhao (✉)

Tsinghua National Laboratory for Information Science and Technology (TNList),
Department of Automation,
Center for Intelligent and Networked Systems (CFINS),
Tsinghua University, Beijing 100084, China
e-mail: zhaoqc@tsinghua.edu.cn

Z. Jiang

Building Energy Research Center, Tsinghua University, Beijing 100084, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_42

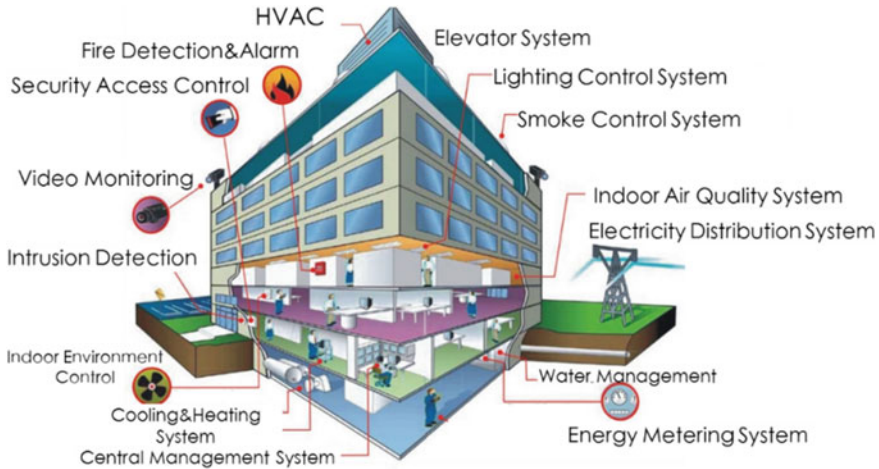


Fig. 1 An illustration of related systems in buildings (Adapted from a source picture on the Internet)

a comfortable, energy saving, and safe indoor environment for occupants. With the advances of the latest information technology, integration of individual control systems could further improve the efficiency of overall building performances. Network technology plays a key role in integrating various control systems. Among the existing solutions, BACnet is an ISO-approved international standard data communication protocol for building automation and control systems. A strength of BACnet is to allow the sensors and the control devices to be connected to the Internet and is accessible from remote [1]. This partially addresses the interpretability problem of devices from different vendors and allows cloud technology to collect big data for analysis [2].

In the existing integration solutions to the building control systems, which is leveled and centralized [3], however does not address the problem of intensive labor cost in constructing and commissioning the control codes for various devices (usually called field devices since they are connected through field buses to their controllers) such as pumps, chillers, and lighting control. One key issue of the current development of building control codes lies in the fact of control logic written in the controllers (typically in PLC code) depends on the specific building configuration of field devices. As a result, control codes developed in one building, usually cannot be reused in a different building. When the field device is upgraded or adding/removal of devices happens, human intervention is needed to make adjustments in the control codes accordingly. From this standpoint, the current control systems for field devices are far from the convenient “plug and play” mode of the Internet or computer devices.

To address the limitations of the current field bus based building control systems, we introduce Insect Intelligent Building (I²B), a new architecture of building control systems based on Internet of Things (IoT), whose basic idea has been outlined in our previous paper [4]. It is a distributed peer-to-peer system consisting of interconnected

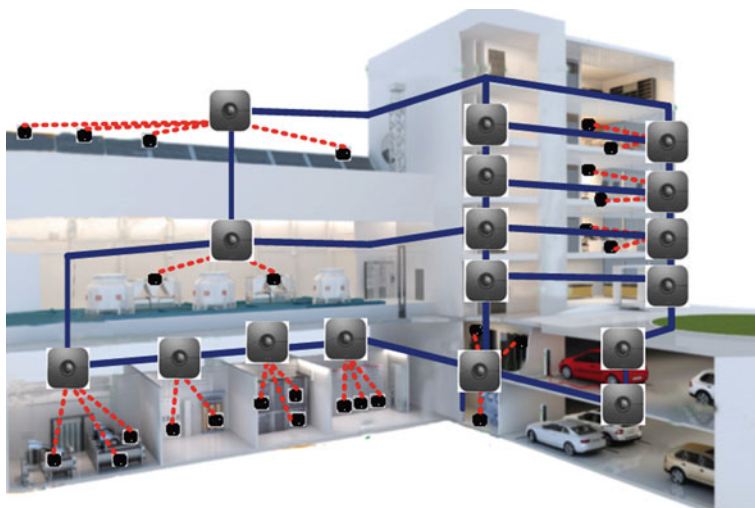


Fig. 2 An application layout of Insect Intelligent Building (I²B)

smart nodes embedded in a building as shown in Fig. 2. It has the following key features: (1) it has built-in agent models for space units and device units; (2) it has built-in link models for immediate connections between space units and device units; (3) it assigns a smart controller (called insect) for each space unit or each device unit with six data ports (called legs) and these controllers communicate with each other only through their data ports according to the link models¹; (4) it decomposes various building operation/management commands into computing tasks that run on the smart controllers and depends on the collaborations among neighboring controllers to achieve the desired effects of the commands. Different from the traditional centralized solutions, this new architecture allows the building control systems to be developed in before the buildings are really built and simplify the installation and configuration of the building control systems.

2 The New Architecture and Key Functional Features

The proposed new building control systems architecture is based on Internet of Things (IoT) and is shown in Fig. 3. The new architecture consists of four parts:

- Tags to recognize building basic modules such as space units (called zones) or control devices and tubes connecting basic modules
- A decentralized building control platform

¹It is the feature that the smart controllers with six legs (refers to the data ports) interconnect to achieve the overall control/management function of the buildings makes us call buildings controlled under our new control architecture “Insect Intelligent Buildings (I²B)”.

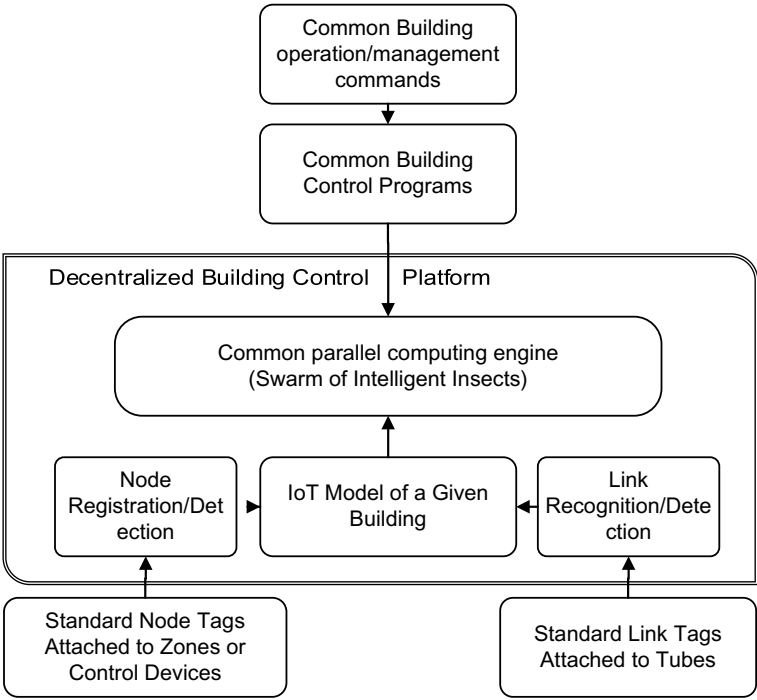


Fig. 3 Building control systems in the proposed new architecture

- Computing process nodes for each space unit or device unit
- General-purpose building control programs.

The functions of the three parts are as follows. Tags are introduced to describe the properties of basic modules which consist a building and its physical fields such as temperature field, air flow field, and cooling/heating water fields. Here, we regard basic modules as a piece of space or device in a building that does not need to be divided further from the viewpoint of building operation or management.

For example, a room is a natural space unit which can be regarded as a basic module. A cooling machine is also a natural basic module. The decentralized building control platform is introduced to allow the execution of general-purpose building control programs at arbitrary buildings. To do so, the platform constructs an IoT model for a given building by recognizing the tags that are deployed in the building. The IoT model is a network consisting of a node set and link set with nodes corresponding to the basic modules and links corresponding to the neighboring relation between modules. Physically, the IoT model is stored separately in a set of the interconnected computing nodes called the parallel computing engine (or the swarm of intelligent insects). The structure of the parallel computing engine is shown in Fig. 4

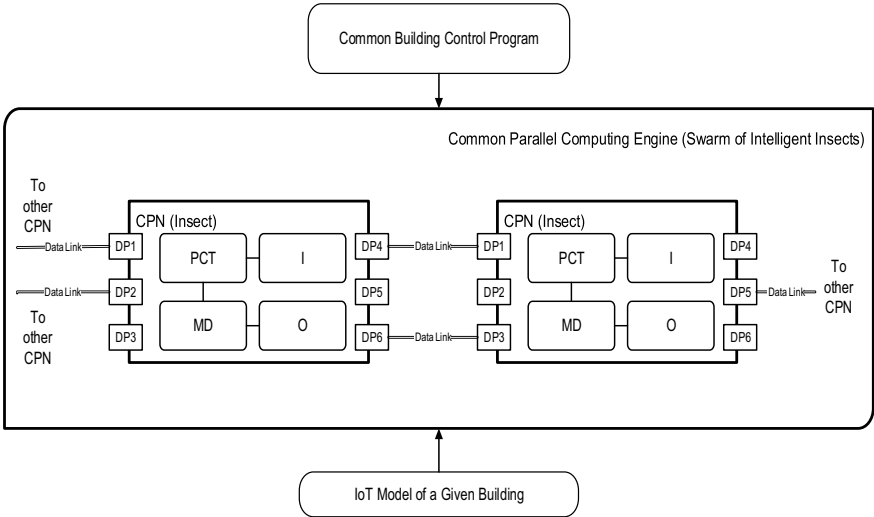


Fig. 4 The structure of the parallel computing engine (swarm of intelligent insects)

with details of a node shown in Fig. 5. Each node of the engine² hosts a complete model data of one basic module and the complete model data of all links connecting to the module. Two nodes of the engine are connected through the data links if the basic modules in the building are connected through tubes with tags. Each data link can represent one tube. The nodes of the parallel computing engine also host a set of building control programs that are written as a sequence of parallel computing tasks. Each computing task is an elementary instruction that can be executed in parallel by the node computer with the help of neighboring nodes through data exchange over data links. It should be emphasized that all computing tasks are written local in the sense that tasks only refer to node physics property data and communication demand over data links but not refer to global data such as unique names or addresses of specific modules or tubes. In this way, the building control programs can be developed separately from the construction of the IoT model of specific buildings and it is the parallel computing engine that binds them to generate the building specific sensing data and control the device commands.

The general-purpose building control programs is a formal description of common building operation/management requirements that are not dedicated to a specific building. They are taken by the parallel computing engine within the decentralized building control platform and can be interpreted as sequences of parallel computing tasks that can be executed on the node computers of the engine. The new architecture has the following key features:

- It enables the automatic modeling of the building control systems, saving time and labors on installation and configuration

²A node is known as a CPN (Computing Process Node) in [4].

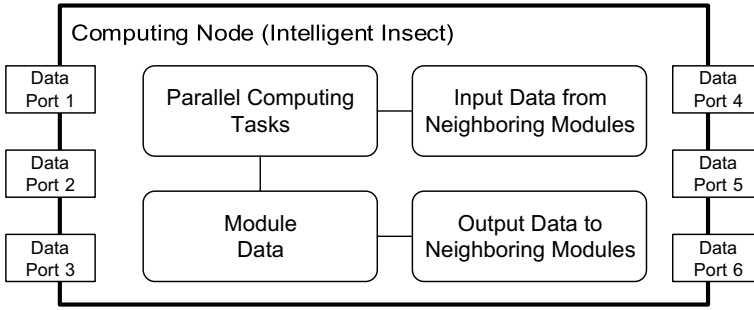


Fig. 5 Details of a computing node (intelligent insect)

- It enables the integration of device control
- It enables the maintainability and extensibility of the building control system
- It enables large quantities of building control programs and diagnosis programs, such as sensor faults diagnosis.

3 Theoretical Foundations

One of the key foundation to model the building control systems is an automatic modeling ability provided by the network topology recognition. This can be done by using RFID tags which have been widely used [5] or by topology matching that was introduced recently [6].

Based on the IoT model of a given building, we can leverage existing results from distributed computing [7], especially the recent advances of the peer-to-peer algorithms which have been demonstrated feasible for large scale real-time video data services over the Internet [8]. Theory of the networked control systems or multi-agent system models may also provide useful tools to design distributed control algorithms for various building control and management problems. In fact, one can cast many general building control and management problems into constrained optimization problems, which can be solved by general distributed algorithms. As an example, let us consider a typical constrained optimization problem.

$$f^* = \min_x \sum_{i \in D} f_i(x_i) \quad (1)$$

Subject to:

$$\sum_{i \in D} g_i(x_i) = C \quad (2)$$

or

$$\sum_{i \in D} g_i(x_i) \leq C \quad (2')$$

where C is a nonnegative constant and domain $D = \{i | P(x_i) \leq 0\}$ is defined by a subset of nodes satisfying the condition P given by inequality of local variables x_i .

We assume that the constraint functions g_i are nonnegative and there is a spanning tree T with the root node (labeled as node 0) belongs to the interested domain D .

Under these assumptions, propose a simple randomized insect intelligent algorithm framework as shown in Fig. 6. Below is an explanation of the algorithm framework. Each node (insect) in domain D continuously does two things, one is to generate random samples $x_i(u_j)$, $j = 1, 2, \dots, M$ of its local decision variable x_i and evaluate the corresponding performance functions $f_i(x_i)$ and $g_i(x_i)$. The performance evaluation process could be done through either a simulation model or measurement of certain variables attached to a given node i . The second thing is to calculate its subtree T_i , sum of $f_i(T_{ij})$ and $g_i(T_{ij})$, and drops infeasible samples $x_i(u_j)$ that violate the constraint (2) or (2') in the sense that $g_i(T_{ij}) > C$. Note the first thing can be done in parallel and locally for all nodes except the root. The root node utilizes its unique position to decide samples $x_0(u_j)$ by fitting the gap left by all other nodes in D for the requirement (2) or (2'). The second thing can be done through communication with immediate neighbors along the spanning tree.

The advantage of this algorithm is that it is quite simple and general and avoids the trouble of violating the constraint (2) if all nodes generate sample purely independently. An implementation on a line topology parallel pump control algorithm can be found in [9].

More sophistic optimization tools are available such as the one reported in [10].

4 Case Studies

With people's demands of increasing for indoor environment, HVAC system energy consumption has occupied 65% of the total building energy consumption, and has been growing. Almost half of these energy consumptions are consumed in the water system, and pumps are the considerable power consuming elements in the water system. Some distributed optimization algorithms have been developed that compatible with our new architecture [9, 11, 12].

Occupant distribution information can be used to predict the heating and cooling load of each zone and therefore to coordinate the heating, ventilation, and air conditioning (HVAC) system and energy-storage devices to improve the energy efficiency of the building under normal conditions. Under emergent conditions, this information can not only provide an initial condition to generate an evacuation plan, but also adjust the guidance in real time to avoid congestion.

Therefore, zonal occupant-level estimation is of great practical interest. It has been shown the zonal occupant-level estimation problem could be solved through distributed computing [13] based on the conservation of occupants. In other words,

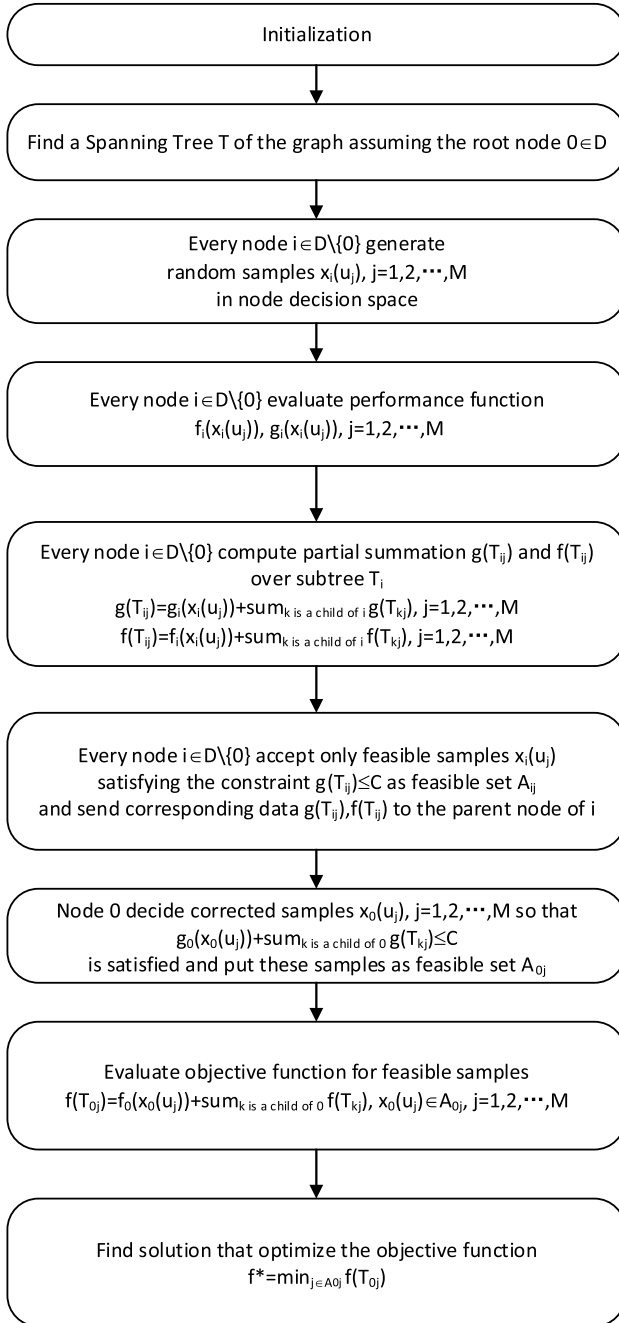


Fig. 6 An insect intelligent algorithm framework to solve constrained optimization problems

when occupants travel among the zones, the total number of occupants in these zones should keep constant. The collaboration among smart nodes in charge of neighboring zones can dramatically reduce the estimation error.

For distributed solutions to cases such as fire source or strength estimation problems, please refer to [4].

5 Conclusions

In this paper, we introduce a new peer-to-peer architecture of building control systems based on Internet of Things (IoT) aiming at reducing labor cost for building control systems construction and commissioning. It is consisting of interconnected smart nodes embedded in a building. Building control and management tasks are executed on the parallel computing platform formed as a multi-agent system of all smart nodes. The new architecture allows the control devices to be connected through smart nodes in a “plug and play” way based on standard information model of devices. It also allows control codes to be developed as downloadable applications that are independent of specific network topologies and thus independent of specific buildings empowered by on the idea of peer-to-peer programming. There are a lot of research challenges and opportunities under the new architecture. For example, the development of a standard description of the space unit and device unit, systematic steps to convert general building control programs into parallel computing tasks running on smart nodes and their communication protocols.

Acknowledgements This work was supported by the National Key Research and Development Project of China (No. 2017YFC0704100 entitled New generation intelligent building platform techniques, and 2016YFB0901900), the National Natural Science Foundation of China (No. 61425027), the 111 International Collaboration Program of China under Grant B06002, BP2018006 and Special fund of Suzhou-Tsinghua Innovation Leading Action (Project Number: 2016SZ0202).

We appreciate the advice of Prof. Yi Jiang from Tsinghua University, Beijing, China, during the procedure in developing the proposed architecture and a lot of discussions with Prof. Qiansheng Fang from Anhui Jianzhu University, Hefei, China, Jilin Zhang from Dalian University of Technology, Dalian, China and Jianchun Xing from The Army Engineering University of PLA, Nanjing, China and insightful comments from Prof. Zhenya Zhang from Anhui Jianzhu University, Hefei, China, Xi Cheng from Tsinghua University, Beijing, China, Li Xia from Tsinghua University, Beijing, China, Tianyi Zhao from Dalian University of Technology, Dalian, China, Qiliang Yang from The Army Engineering University of PLA, Nanjing, China and Dr. Yunchuang Dai from Tsinghua University, Beijing, China.

We also appreciate constructive comments from the anonymous reviewers.

References

1. Bushby, S.T.: BACnet today: Significant new features and future enhancements. *ASHRAE J.* **44**, 10–18 (2002)
2. Fernbach, A., Granzer, W., Kastner, W.: Interoperability at the management level of building automation systems: a case study for BACnet and OPC UA. In: *Proceedings of the IEEE 16th Conference on Emerging Technologies Factory Automation (ETFA)*, pp. 1–8. IEEE Press, New York (2011)
3. Domingues, P., Carreira, P., Vieira, R., et al.: Building automation systems: concepts and technology review. *Comput. Stand. Interfaces* **45**, 1–12 (2016)
4. Jiang, Z., Dai, Y.: A decentralized, flat-structured building automation system. *Energy Procedia* **122**, 68–73 (2017)
5. Weinstein, R.: RFID: A technical overview and its application to the enterprise. *IT Prof.* **7**(3), 27–33 (2005)
6. Wang, Y., Zhao, Q.: A distributed algorithm for building space topology matching. In: Fang, Q., et al. (eds.) *Advancements in Smart City and Intelligent Building, Advances in Intelligent Systems and Computing* 890, Chap. 37 (2019)
7. Lynch, N.A.: *Distributed Algorithms*. Morgan Kaufmann, San Francisco (1996)
8. Buford, J., Yu, H., Lua, E.K.: *P2P Networking and Applications*. Morgan Kaufmann, San Francisco (2008)
9. Zhao, Q., Wang, X., Wang, Y., et al.: A P2P algorithm for energy saving of a parallel-connected pumps system. In: Fang, Q., et al. (eds.) *Advancements in Smart City and Intelligent Building, Advances in Intelligent Systems and Computing* 890, Chap. 36 (2019)
10. Han, G., Chen, X., Zhao, Q.: Decentralized differential evolutionary algorithm for large-scale networked systems. In: Fang, Q., et al. (eds.) *Advancements in Smart City and Intelligent Building, Advances in Intelligent Systems and Computing* 890, Chap. 38 (2019)
11. Dai, Y., Jiang, Z., Shen, Q., et al.: A decentralized algorithm for optimal distribution in HVAC systems. *Build. Environ.* **95**, 21–31 (2016)
12. Yu, H., Zhao, T., Zhang, J., Li, X.: A decentralized algorithm to optimize multi-chiller systems in the HVAC system. In: *2017 Chinese Automation Congress (CAC)* (2017)
13. Jia, Q.S., Wang, H., Lei, Y., et al.: A decentralized stay-time based occupant distribution estimation method for buildings. *IEEE Trans. Autom. Sci. Eng.* **12**(4), 1482–1491 (2015)

Part VII

Smart Home and Smart Utility

The Fault Diagnosis Model Established Based on RVM



Yahui Wang, An Yun, Qinghong Ye and Yunfeng Zhao

Abstract This paper introduces the application of Relevance Vector Machine (RVM) in fault diagnosis. First, the theoretical contents of RVM including the algorithm characteristics, the derivation of mathematical models, the characteristics of kernel functions and multimode classification are introduced. Then, the multi-classification fault diagnosis model of building electrical system is established by using RVM. Finally, the experimental results show that the RVM model has good classification effect on small sample data.

Keywords Fault diagnosis · Relevance vector machine · Small sample data

1 Introduction

As one of the key technologies in whole building systems, building electricity is a subsystem of fault diagnosis that can be not ignored. Traditional building electrical fault diagnosis algorithms, such as naive Bayes, need a large number of sample data to improve the accuracy of the diagnosis, which is not suitable for the classification vector machine (RVM) to solve the problem of building electrical fault diagnosis. In this paper, a fault classification model for small sample data is discussed.

At present, there are many types of research on intelligent diagnosis methods with high classification accuracy for small sample data [1–3]. RVM has good classification performance for small sample data, and its output in the form of probability can reveal the uncertainty of the classification result. The application of RVM in the early diagnosis of mechanical faults, bearing fault detection, power load forecasting,

Y. Wang · A. Yun (✉) · Y. Zhao
School of Electrical and Information Engineering, Beijing University of
Civil Engineering & Architecture, Beijing, China
e-mail: 2269146452@qq.com

Q. Ye
Shanghai Aerospace Energy LTD, Shanghai, China

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_43

power system transient stability assessment, bearing fault detection, and other fields have achieved very good results.

2 The Model of Relevance Vector Machine

In the sample training set, the input vector can be expressed as $\{x_i\}_{i=1}^N$. The target vector can be represented as $t = [t_1, t_2, \dots, t_N]^T$ and it can also be expressed as

$$t_i = y(x; w) + \varepsilon_i \quad (1)$$

The model vector output function of SVM can be expressed as

$$y(x; w) = \sum_{i=1}^N \omega_i K(x, x_i) + \omega_0 \quad (2)$$

where $w = [w_0, w_1, w_2, \dots, w_N]^T$ represents a weight function and $K(x; x_i)$ represents a weight function.

In a sparse Bayesian framework, the probability distribution of the objective function can be expressed as

$$p(t_i) = N(t_i | y(x_i; w), \sigma^2) \quad (3)$$

The likelihood estimation probability of a sample set can be expressed as

$$p(t | w, \sigma^2) = \prod_{i=0}^N N(t_i | y(x_i; w), \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{\|t - \Phi w\|^2}{2\sigma^2}\right) \quad (4)$$

$$\Phi = [\phi(x_1), \phi(x_2), \dots, \phi(x_N)]^T \quad (5)$$

$$\phi(x_i) = [1, K(x_i, x_1), K(x_i, x_2), \dots, K(x_i, x_N)]^T \quad (6)$$

Using the maximum likelihood estimation method to solve w and σ^2 directly under the Bayesian framework can lead to over-learning problems. Therefore, RVM assigns a zero mean Gaussian prior distribution to the weight vector w as

$$p(w | \alpha) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}) \quad (7)$$

Using Bayes' theorem, solve the posterior probability as

$$p(w, \alpha, \sigma^2 | t) = \frac{p(t|w, \alpha, \sigma^2)p(w, \alpha, \sigma^2)}{p(t)} \quad (8)$$

For the test data sample y , the prediction distribution of the solution objective function y can be expressed as

$$p(t_y | t) = \int p(t_y | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2 | t) dw d\alpha d\sigma^2 \quad (9)$$

$$p(w, \alpha, \sigma^2 | t) = p(w | t, \alpha, \sigma^2) p(\alpha, \sigma^2 | t) \quad (10)$$

$$p(t | \alpha, \sigma^2) = \int p(t | w, \sigma^2) p(w | \alpha) dw = (2\pi)^{-\frac{N}{2}} |\Omega|^{-\frac{1}{2}} \exp \left\{ -\frac{t^T \Omega^{-1} t}{2} \right\} \quad (11)$$

$$\Omega = \sigma^2 I + \Phi A^{-1} \Phi^T, A = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$$

$$\Sigma = (\sigma^{-2} \Phi^T \Phi + A)^{-1} \quad (12)$$

$$\mu = \sigma^{-2} \Sigma \Phi^T t \quad (13)$$

A posterior probability distribution about weights can be expressed as

$$p(w | t, \alpha, \sigma^2) = \frac{p(t | w, \sigma^2) p(w | \alpha)}{p(t | \alpha, \sigma^2)} = (2\pi)^{-\frac{N+1}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{(w - \mu)^T \Sigma^{-1} (w - \mu)}{2} \right\} \quad (14)$$

$$\begin{cases} p(t_y | t) \approx \int p(t_y | w, \alpha_{MP}, \sigma_{MP}^2) p(w | t, \alpha_{MP}, \sigma_{MP}^2) dw \\ (\alpha_{MP}, \sigma_{MP}^2) = \arg \max_{\alpha, \sigma^2} p(\alpha, \sigma^2 | t) \end{cases} \quad (15)$$

After solving the α and σ^2 partial differentials by the Eq. (11), we can get the iterative update equation of α and σ^2 as

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \quad (16)$$

$$(\sigma^2)^{new} = \frac{\|t - \Phi \mu\|^2}{N - \sum_i \gamma_i} \quad (17)$$

with $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ where Σ_{ii} is the i th diagonal element of Σ . The prediction model of the correlation vector machine RVM can be expressed as

$$p(t_y | t) = N(t_y | y, \sigma_y^2) \quad (18)$$

In essence, RVM and SVM belong to the same classifier, so RVM can use the Logistic sigmoid function to solve the likelihood estimation of the sample set as

$$p(t | w) = \prod_{n=1}^N \sigma \{y(x_n; w)\}^{t_n} [1 - \sigma \{y(x_n; w)\}]^{1-t_n} \quad (19)$$

2.1 Selection of Kernel Functions

The number of jobs to be scheduled: The notable feature of the kernel function is that the inner product operation in the high-dimensional feature space is implicitly implemented by the inner product kernel function defined in the original space [4]. The choice of the kernel function has a strong correlation with the performance of the correlation vector machine. Different kernel functions are selected to correspond to different nonlinear transformations and feature spaces, and an RVM classifier corresponding to the kernel function type is generated. The commonly used kernel functions in the following four RVM algorithms are briefly introduced [5]:

Linear Kernel: $K(x, z) = x \cdot z$

Polynomial Kernel: $K(x, z) = [(x \cdot z) + 1]^d$

Radial Basis Function: $K(x, z) = e^{-\|x-z\|^2/2\sigma^2}$, $\sigma > 0$

Sigmoid Kernel: $K(x, z) = \tanh(a \langle x, z \rangle + r)$

Cauchy Kernel: $K(x, y) = 1/(1 + \|x - y\|^2/\sigma)$

The appropriate kernel function is selected according to the characteristics of the sample data to be diagnosed, and the corresponding RVM fault diagnosis model is established. In order to construct an RVM classifier with good classification performance, we need not only to consider what type of kernel function to select, but also to determine the parameters of the type kernel function. In this paper, the RBF is selected according to the sample data characteristics of building electrical system. It is mainly based on the good performance of RBF for small sample data classification, and its less related parameters.

2.2 Multimode Classification of RVM

To set the problem of k classification, the One Against Restart (OAR) requires k RVMs to construct multiple classification models, and each classifier needs to distinguish this class from the rest of the classes as a binary classification problem, so the training of each classifier needs all samples to participate in. Figure 1 shows an example of using the OAR method to construct RVM multiple classifiers. By comparing the output of four RVM classifiers, the category with the highest probability is the fault category to which the test sample belongs.

For k -classification problem, One Against One (OAO) needs to construct multi-classification model through $k(k-1)/2$ RVMs, divide the training sample set into $k(k-1)/2$ subsets, and each classifier only needs to participate in the training of the corresponding subsets of samples. Figure 2 shows an example of using the OAO to construct RVM multiple classifiers. The multi-classification model in the figure needs to classify four types of failures, using six RVM classifiers, two more than the OAR method. The classifier compares the output of the six RVM classifiers and selects the category with the most votes as the diagnosis result.

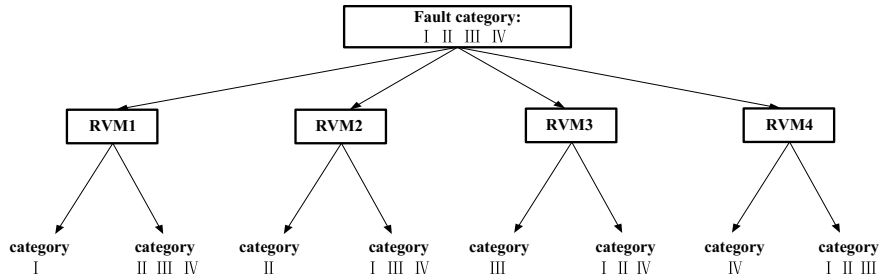


Fig. 1 Multiple classifiers of one against rest

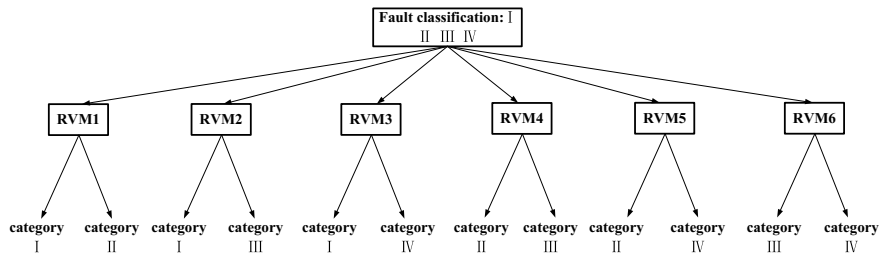


Fig. 2 Multiple classifiers of one against one

The fault diagnosis of building electrical system is essentially a multi-classification problem, and the basic RVM is a binary classifier. According to the requirements of diagnostic accuracy and fault classification of the multi-classifier, the OAO model is chosen, which is mainly due to the high accuracy of OAO classification, the small rejection classification area and the ease of sample training. For the classification of 5 fault states, 10 RVM binary classifiers need to be constructed, and the number of classifiers required is small.

3 Case Verification: Establishment of Building Electrical Fault Diagnosis Model Based on RVM

Failure data diagnosed by RVM algorithm is from Fault Simulation Platform (FSP). There are 10 different positions on the FSP that can be used as the resistance value of the feature vector.

3.1 Establishment of RVM Fault Diagnosis Model

The method of voting is used to classify the test sample data. Each test sample needs to be judged by 10 RVM classifiers, and the classification results of 10 RVM classifiers are accumulated. The class with the most votes is used as the classification result. Figure 3 shows the RVM fault diagnosis model for building electrical systems.

3.2 The Choice of RVM Kernel Function

The RVM binary classifier fault diagnosis model is applied in MATLAB to verify the fault diagnosis effect of PKF, RBF, and Cauchy kernel function. The diagnosis results are shown in Table 1.

In this paper, the RBF is used to establish the RVM-based fault diagnosis model for building electrics. The RBF kernel function is as follows:

$$K(x,z)=e^{-\frac{\|x-z\|^2}{2\sigma^2}}$$
 (20)

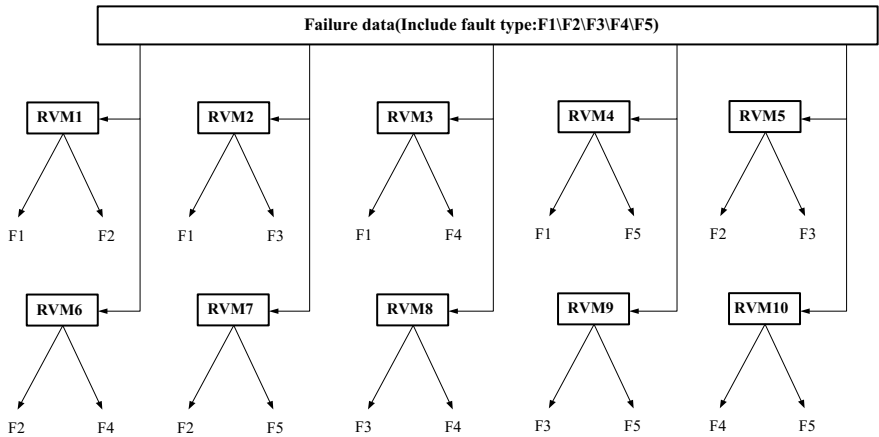


Fig. 3 Diagnostic model of RVM for building electrical system

Table 1 Fault diagnosis result of different kernel functions

Kernel function	Training time/s	Fault diagnosis accuracy/%
Polynomial	0.6787	88.25
Cauchy	0.7525	88.25
RBF	0.6855	100

In the Eq. (20), σ is expressed as kernel width coefficient. RVM does not need to set a penalty factor, and in SVM, penalty parameter is a constant to balance the empirical risk and confidence interval, the improper setting will lead to problems such as over-learning [6]. But in RVM, the parameters are automatically assigned without the use of K-cross-validation method to determine the optimal parameters [2]. This article sets σ to 0.5.

3.3 Simulation Experiment of Electrical Fault Diagnosis Based on RVM

FSP generates 60 sets of fault data and uploads 45 sets as training samples and 15 sets as test samples through the data acquisition system. The sample distribution is shown in Table 2.

After establishing the RVM framework for fault diagnosis of building electrical systems, 45 training samples were input to establish a classification model, and then the test samples were input for fault diagnosis. First, select five groups of test samples. The normal state of the building electrical system F1, line impedance fault F2, grounding resistance fault F3, insulation resistance fault F4, and continuity fault are grouped in each category. The statistical RVM fault diagnosis algorithm voting results are shown in Table 3 shows.

The total number of votes for the five fault categories is 10, and the maximum number of votes for each fault category is 4 votes. In the actual classification diagnosis, the class with the number of votes in 3–4 is the fault class diagnosed by RVM.

Table 2 Sample distribution of simulation

Fault status	Number of training samples	Test sample number
F1	9	2
F2	9	2
F3	9	4
F4	9	2
F5	9	5

Table 3 Diagnosis voting results of RVM

Actual fault category	Algorithm for diagnosing votes				
	F1	F2	F3	F4	F5
F1	3.75	1.39	2.11	0.94	1.81
F2	1.31	3.84	0.98	1.95	1.92
F3	1.92	1.02	3.68	1.55	1.83
F4	1.86	1.28	0.89	3.6	2.37
F5	1.18	1.44	1.92	1.57	3.89

As shown in Table 3, the remaining votes in the OAO algorithm will be randomly assigned to the other four categories when the output result of the algorithm is that the fault belongs to F1 and the number of votes obtained is 3.75, corresponding to the sample data of the actual fault F1 test. Other F2 got: 1.39, F3 got 2.11, F4 got 0.94, F5 got 1.81. In general, the algorithm to determine the category of votes to be significantly higher than other categories, generally higher than 3 votes or more, the distinction is obvious.

4 Conclusions

In this work, a multi-class fault diagnosis model for building electrical systems based on RVM algorithm is established, which not only achieves a good classification of small sample data, but also reveals the uncertainty of fault classification results through the output of its probability form. We use the one-versus-one multi-classification algorithm to establish the simulation model on MATLAB. The simulation results show that the RVM-based electrical fault diagnosis model has a good effect on the classification of small sample data.

Acknowledgements The project was partially supported by “The Fundamental Research Funds for Beijing University of Civil Engineering and Architecture”, Beijing, China with the grant No. X18191.

References

1. An, Y., Wang, Y.H., Zhang, F.C.: Research on gas pressure regulator fault diagnosis based on deep confidence network (DBN) theory. CAC (2017). <https://doi.org/10.1109/cac.2017.824344>
2. Shen, Y., Ding, S.X., et al.: A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. J. Process Control **22**(22), 1567–1581 (2012)
3. Tran, V.T., Althobiani, F., Ball, A.: An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. Expert Syst. Appl. **41**(9), 4113–4122 (2014)
4. Ishigaki, T., Higuchi, T., Watanabe, K.: Fault detection of a vibration mechanism by spectrum classification with a divergence-based kernel. IET Signal Process. **4**(5), 518–529 (2010)
5. Ando, S., Kato, M.: Criticism of the Desmarais method for kernel function computation. Trans. Jpn. Soc. Aeronaut. Space Sci. **31**(93), 161–170 (1988)
6. Li, C., Sanchez, R.V., Zurita, G., et al.: Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. Neurocomputing **168**(C), 119–127 (2015)

PLC-Based Intelligent Home Control System



Liming Wei, Yangyun Wu and Xiujuan Guo

Abstract This paper presents an intelligent home control system based on PLC. The system combines the BP neural network algorithm with the traditional PID algorithm. The proposed algorithm has better dynamic performance, faster adjustment speed, and smaller overshoot. It can track to control and adjust the complex and nonlinear data with a higher standard. At the same time, the overall design plan of the system is described in details. The system designed in this paper provides a good reference for the practical application of intelligent home.

Keywords Intelligent home control system · PLC · Improved BP neural network

1 Introduction

With the progress of society and the enhancement of China's national strength, millions of families have already a well-off life. Therefore, the safety, comfort, convenience, and reliability of family housing are increasing day by day, and then intelligent home comes into being. However, due to the high cost, cumbersome operation and no national uniform standards and many other reasons, the intelligent home has not entered a completely ordinary family. Therefore, it is particularly important to study the meaning of intelligent home control system with convenient operation and low price. This paper mainly studies the current situation of intelligent home domestic and abroad and analyzes the shortcomings of the existing intelligent home control system. Based on the principle of simple operation and convenient use, the overall framework, hardware selection and software configuration of the system are determined.

L. Wei (✉) · Y. Wu · X. Guo

School of Electrical and Computer Engineering, Jilin Jianzhu University, Changchun 130118, China

e-mail: 982396524@qq.com

L. Wei · Y. Wu · X. Guo

Art Science Architecture Creative Center, Jilin Jianzhu University, Changchun 130118, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_44

2 Overall Design of the System

The object of the research work in this paper takes a model of a living house with two bedrooms, one hall, one kitchen, and one bathroom. Through real-time monitoring and control of various parameters in the home, the user can enjoy a convenient and quick life. According to the structure of home and the actual use requirement of home in the daily life, the intelligent home control system designed in this paper mainly realizes the following functions: (1) intelligent lighting control: according to the acquisition, analysis, control of the light intensity in the room, the requirements of the indoor lighting by daily household and other factors, the indoor lighting is realized to be automatically controlled so as to meet the overall requirements of the use of lights in the house. (2) Intelligent temperature and humidity control: through the analysis of data collected from the temperature and humidity sensors in each room in the house and compare it with the desired ideal value, after the arithmetic operation and comparison, it is returned to the air conditioner and other executing agencies so as to adjust the temperature and humidity of the room. (3) Intelligent curtains control: This article realizes through the analysis of individual needs and the collection of outdoor light intensity to achieve the intelligent control of the curtain to achieve the protection of personal privacy. (4) The collection and alarm of natural gas concentration in the house: This paper collects the natural gas concentration in the kitchen in real time, and once the natural gas leak is found, the system will automatically open the kitchen window and alarm the user to ensure the safety of the personal property in the house. (5) Hardware design: the programmable controller of this paper adopts SIEMENS PLC S7-300 series. It has the advantages of an anti-interference device, low price, convenient operation and so on. RS-485 bus is used to realize communication between the upper and lower monitors. (6) Software design: the editor of the S7-300 of the lower monitor uses its Step 7 programming software to edit the corresponding system. The upper computer monitors the data of the whole intelligent home control system by Kingview 6.55 and can simultaneously control the expectation value setting and target value of the whole system at the king view. In this paper, MATLAB is used to simulate and debug the algorithm adopted by the system. (7) Touch screen design: This article uses the SIEMENS touch screen TP700 to complete the design of the interface of the whole system and realize the real-time monitoring of all the data of the whole home network.

According to the analysis of the hardware selection and software design of the whole system, the SIEMENS PLC S7-300 is used as the controller of the whole system. Kingview is used as the control software of the system. And RS485 bus is used to real-time communication between the hardware and software of the upper and lower computer.

The central idea used in the design is “human-oriented”, considering the problem from the angle of the user, so that the intelligent home control system is more humanized and convenient. The two operation modes (ie. manual and automatic) are employed in this design to control the entire intelligent home to achieve a better user experience. It can allow users to switch freely and users can adjust the actual

situation of the home according to the personal needs. It is mainly divided into two parts, named automatic control and manual control. Automatic control is to measure all kinds of data indoor by the sensor of the system, and then transfer the data to the control device (PLC) which is calculated and analyzed by PLC and compared with the ideal value set by the user. When there is a deviation from the ideal value, PLC will command the relevant executor to action. If the indoor temperature is lower than the ideal value, the system will first calculate the size of the error, judging the temperature that the air conditioner needs to open by this error value, and then transfer the temperature to the air conditioner to make the room's temperature rise, so as to maintain the comfort of the indoor temperature. In manual control, when the user's entire system is transferred to the manual control option, the automatic control device of the whole system will be closed. At the time, the collector is still working to carry out real-time data acquisition for the whole room. The user can detect the data in the room through the display screen, and manually adjust the relevant data to achieve the purpose of manual control.

3 PID Control Algorithm Based on BP Neural Network

3.1 Algorithm Principle

The algorithm proposed in this paper is mainly to combine the BP neural network control algorithm with the traditional PID control algorithm, to make up for the shortcomings of the traditional PID control algorithm in function and make the results more accurate and faster. The principle of the algorithm is to modify the parameters K_p , K_i , and K_d by BP neural network algorithm to achieve the best value of the experiment, thus improve the shortcomings of the slow calculation and the inaccurate results of the traditional PID algorithm. First, the indoor temperature is collected through the indoor sensor, then the gap between the temperature and the ideal value is compared. Then the gap between the indoor temperature and the preset value is reduced by the weighting, mapping, and feedback of the BP neural network algorithm, making sure that the indoor temperature is kept at the optimum temperature.

The mathematical expressions of the classical PID control algorithm are as follows:

$$u(k) = u(k-1) + K_p[e(k) - e(k-1)] + K_i e(k) + K_d[e(k) - 2e(k-1) + e(k-2)] \quad (1)$$

where $u(k)$ is the system function, k_p is the proportional factor, k_i is the integral factor, k_d is the differential factor.

If the three parameters are regarded as the adjustable coefficients of the system, the expression can be written as

$$u(k) = f[u(k-1), K_p, K_I, K_D, e(k), e(k-1), e(k-2)] \quad (2)$$

where $f()$ is a nonlinear function related to equal data. Through the continuous learning of BP neural network, the best rule of this function can be obtained and can be applied. Because there is no BP neural network function model in MATLAB/Simulink, this paper uses the Sigmoid function to edit the BP neural network system, and then directly introduces the S function in Simulink to achieve the simulation of the PID control algorithm based on the BP neural network (Fig. 1).

This paper assumes that the BP neural network has three-layer network with the number of nodes of the input layer is M , the number of hidden layer nodes is Q and the number of the output layer nodes is 3. The output nodes are the three parameters named K_p , K_i , and K_d that need to be controlled in the PID algorithm. The structure of the BP neural network is as follows Fig. 2:

From Fig. 2, the conclusion can be drawn that the input layer of BP neural network is

$$\left. \begin{array}{l} O_j^{(1)} = x_j, j = 0, 1, 2, \dots, M-1 \\ O_M^{(1)} = 1 \end{array} \right\} \quad (3)$$

where O is the input layer function, the selection of the number of nodes in the input layer depends on the complexity of the controlled system. The input and output equations of the hidden layer in the BP neural network structure diagram are

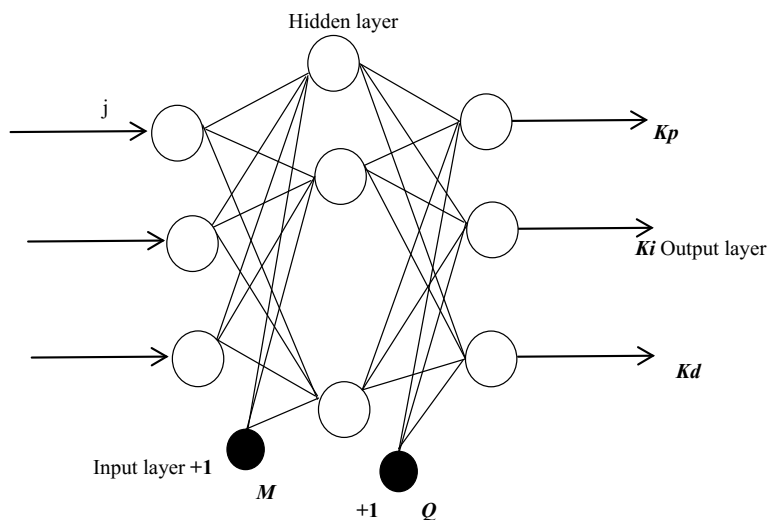


Fig. 1 BP neural network structure diagram

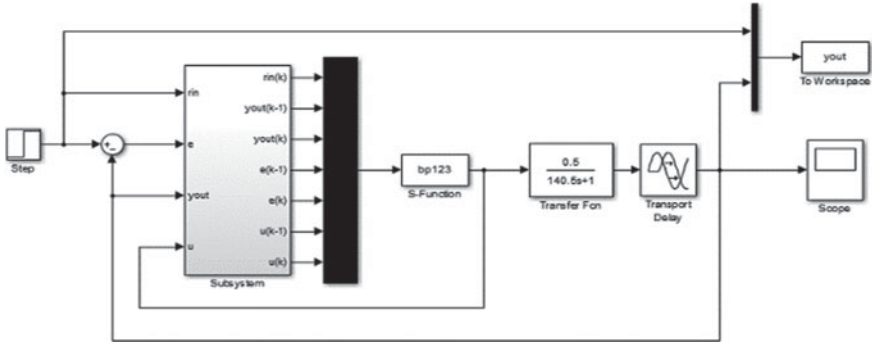


Fig. 2 PID controller model based on BP neural network

$$\left. \begin{aligned} net_i^{(2)}(k) &= \sum_{j=0}^M \omega_{ij}^{(2)} O_j^{(1)}(k) \\ O_i^{(2)}(k) &= f[net_i^{(2)}(k)] \quad (i = 0, 1, 2, \dots, Q-1) \\ O_Q^{(2)}(k) &= 1 \end{aligned} \right\} \quad (4)$$

where $\omega_{ij}^{(2)}$ is the weighted coefficient. $f(\bullet)$ is the activation function of the system ($f(\bullet) = \tanh(x)$)

The input and output equations of the output layer in the BP neural network structure diagram are

$$\left. \begin{aligned} net_i^{(3)}(k) &= \sum_{l=0}^Q \omega_{il}^{(3)} O_l^{(2)}(k) \\ O_l^{(3)}(k) &= g[net_l^{(3)}(k)], \quad (l = 0, 1, 2) \\ O_0^{(3)} &= KP \\ O_1^{(3)} &= KI \\ O_2^{(3)} &= KD \end{aligned} \right\} \quad (5)$$

where $\omega_{il}^{(3)}$ is weighting coefficient of the output layer, the number of the upper corner is represented by the specific position of the equation. $g[\bullet]$ is the activation function of this layer, because the output value K_p , K_i , and K_d is in the PID system, so the results of $g[\bullet]$ needs to be positive, so the equation is

$$g(x) = \frac{1}{2}[1 + \tanh(x)] \quad (6)$$

In the BP neural network algorithm, the system mainly uses the gradient descent method to speed up the update speed of the weight of the system, and the specific direction of the update is the fastest direction of the gradient descent. The selection of specific performance indicators is based on the equation:

$$J = \frac{1}{2}[r(k+1) - y(k+1)]^2 = \frac{1}{2}e^2(k+1) \quad (7)$$

$$\Delta\omega_{li}^{(3)}(k+1) = -\eta \frac{\partial J}{\partial \omega_{li}^{(3)}} + \alpha \Delta\omega_{li}^{(3)}(k) \quad (8)$$

where η is the rate of learning, α is an inertial factor. The output layer weighting coefficient equation of BP neural network is shown in (9):

$$\left. \begin{aligned} \Delta\omega_{ij}^{(3)}(k+1) &= \eta\delta_l^{(3)}O_i^{(2)}(k) + \alpha\Delta\omega_{li}^{(3)}(k) & l = 0, 1, 2 \\ \delta_l^{(3)} &= e(k+1)\text{sgn}\left(\frac{\partial y(k+1)}{\partial u(k)}\right) \times \frac{\partial u(k)}{\partial O_l^{(3)}(k)}g[net_l^{(3)}(k)] & l = 0, 1, 2 \end{aligned} \right\} \quad (9)$$

BP neural network adopts the following equation for weighing the hidden layer coefficient.

$$\left. \begin{aligned} \Delta\omega_{ij}^{(2)}(k+1) &= \eta\delta_i^{(2)}O_j^{(1)}(k) + \alpha\Delta\omega_{ij}^{(2)}(k) \\ \delta_i^{(2)} &= f[net_i^{(2)}(k)] \sum_{l=0}^2 \delta_l^{(3)}\omega_{li}^{(3)}(k) & i = 0, 1, \dots, Q-1 \end{aligned} \right\} \quad (10)$$

In the equation:

$$\begin{aligned} g'[\bullet] &= g(x)[1 - g(x)] \\ f'[\bullet] &= [1 - f^2(x)]/2 \end{aligned}$$

The procedure of the algorithm is as follows:

- (1) The network structure of BP neural network is selected, that is to determine the number of nodes of each layer, and initialize the weight of the layer, set the learning rate and the value of the coefficient $k = 1_\eta$.
- (2) The sample is obtained as $r(k)$ and $y(k)$ and calculate $e(k) = z(k) = r(k) - y(k)$.
- (3) Normalize the data, which means to take normalization of $r(i)$, $y(i)$, $e(i)$ ($i = k, k-1, \dots, k-p$) after processing it is used as input to BP neural network system.
- (4) Bring the normalized data into the equation before, and calculate the value of the adjustable parameters K_P , K_I , K_D in the PID controller after the calculation is completed.
- (5) Calculate the results of the control output of the PID controller $u(k)$.
- (6) Calculate again $\omega_{li}^{(3)}(k)$ and $\omega_{ij}^{(2)}(k)$.
- (7) Make $k = k + 1$, return to step (2).

3.2 Simulation Studies

Because of the shortcomings of the conventional PID controller, such as poor anti-interference ability and lack of stability, this paper adopts the way of adding BP neural network to improve the stability and accuracy of the whole algorithm. And

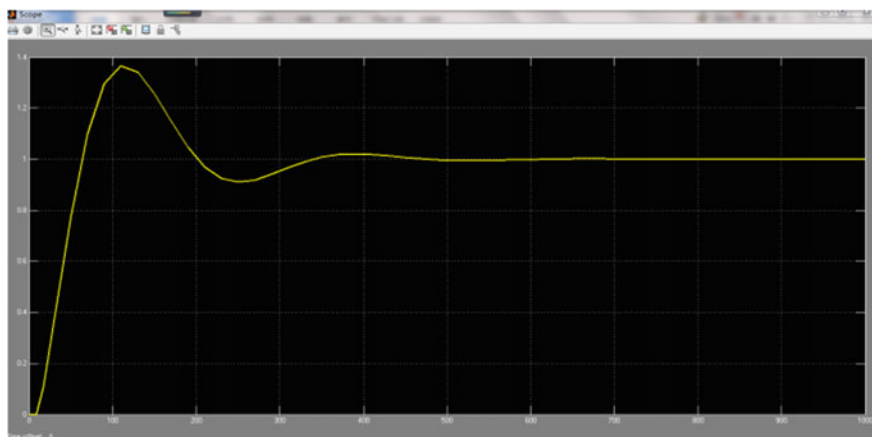


Fig. 3 Simulation result of the traditional PID controller

because there are no modules for BP neural network in Matlab/Simulink, this paper completes the simulation of the PID controller based on the BP neural network. The specific simulation model is shown in Fig. 2. In this paper, Matlab is used to simulate the traditional PID controller. The experiment is to set the simulation time to 1000 s, and the traditional PID simulation curve can be obtained as shown in Fig. 3. After the algorithm is improved, the simulation diagram is shown in Fig. 4.

Through the analysis of the simulation results of the two algorithms, it can be found that the overshoot of the traditional PID controller is about 48%, and the overshoot of the simulation results is greatly reduced when the BP neural network is added. This is because the BP neural network can control 3 parameters K_P , K_I , K_D in real time. After adding the BP neural network, the control time of the whole system also decreases, from the original 700 s to less than 360 s, which is almost twice the time, and the steady-state error is almost zero. It can be seen that the whole algorithm has better dynamic performance, faster adjustment speed and smaller overshoot after adding BP neural network algorithm. It can carry on higher tracking control and regulation for complex and nonlinear data.

4 Conclusions

In this paper, an intelligent home control system based on improved BP neural network PID algorithm is designed. The algorithm combines the PID and BP neural network algorithm to effectively solve the problem of unstable dynamic performance and low adjustment speed. The effectiveness of the algorithm is proved by using

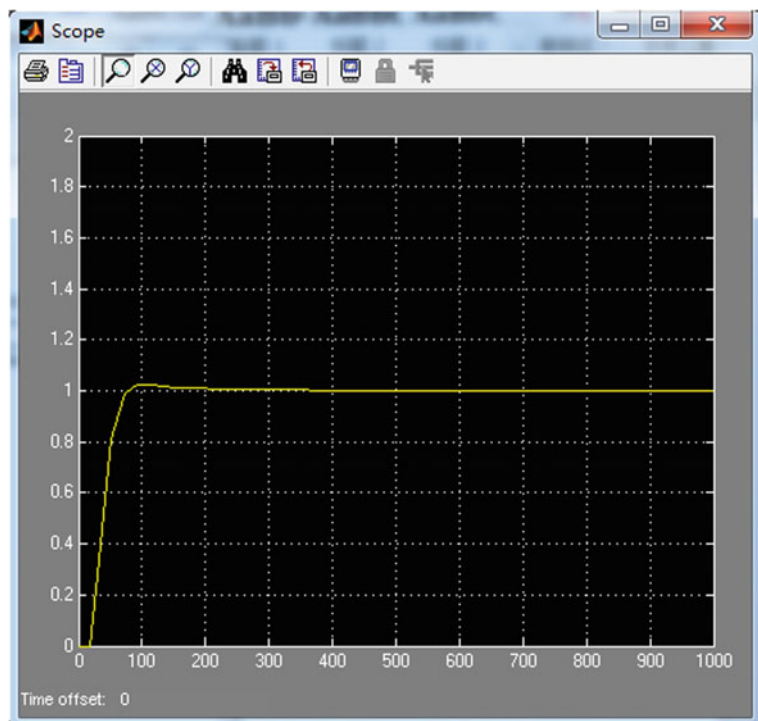


Fig. 4 Simulation result of PID control algorithm based on BP neural system

MATLAB simulation. At the same time, the overall design scheme of the system is presented, and the corresponding flowcharts and structural diagrams are given to show the effectiveness and practicability of the designed system.

Acknowledgements The work was supported by the “13th Five-Year” Science and Technology Research Project of the Department of Education, Jilin Province, China with grant no. [2016] 143.

A Classification-Based Occupant Detection Method for Smart Home Using Multiple-WiFi Sniffers



Ping Wang, Huaqian Cao, Si Chen, Jiake Li, Chang Tu and Zhenya Zhang

Abstract Knowing the number of occupants and where they are located proves crucial in many smart home applications such as automated home control, anomaly detection and activity recognition. In this paper, we propose a novel classification-based occupant counting method that makes use of existing and prevalent WiFi probe requests that are originally designed for WiFi devices to scan WiFi APs at certain channels. First, we employ a binary-location-classification model to determine each detected occupant inside or outside a targeted area; then the neural network is introduced to act as the classifier. Moreover, multiple WiFi sniffers for each given target area are deployed to generate multiple features for the neural network to perform classification and it proves mathematically to be more accurate than one WiFi sniffer only used. Finally, we validate our proposed method through real experiments. Results show that our classification-based occupant detection method using multiple WiFi sniffers outperforms the 1-WiFi-sniffer-based method, and its accuracy makes it suffice to be a viable approach to occupant estimation for smart home.

Keywords Classification · Occupant detection · WiFi probe request · Smart home · Multiple-WiFi sniffers

1 Introduction

Smart home is becoming more and more important to provide a safe, sound and secure living environment. Specifically, being able to accurately count the number of occupants in a target area can be extremely useful in saving significant energy,

P. Wang · H. Cao · S. Chen · J. Li · C. Tu · Z. Zhang (✉)
Anhui Province Key Laboratory of Intelligent Building & Building Energy Saving,
Anhui Jianzhu University, Hefei, Anhui, China
e-mail: zzychm@ustc.edu.cn

P. Wang · H. Cao · S. Chen · J. Li · C. Tu · Z. Zhang
School of Electronic and Information Engineering, Anhui Jianzhu University,
Hefei, Anhui, China

elderly care and anomaly detection in smart home. A significant amount of work has been done in the past two decades to enable accurate and robust occupant counting. Occupant counting in smart buildings can be implemented via video processing and camera systems or deployment of occupancy sensors, including CO_2 concentration sensor, RFID sensor, motion sensor, light sensor, temperature sensor, acoustic sensor, humidity sensor, passive infrared sensor and so on, throughout the building [1]. These sensors are supposed to collect information that is relevant to occupant, rendering occupant counting in buildings feasible. As been stated and summarized in [2, 3], these existing approaches have their own drawbacks in terms of implementation cost, user privacy and intrusiveness. Consequently, it is worthy to investigate other potential methods to overcome these shortcomings. An alternative option is to utilize the pervasive wireless signals that uniquely match to building occupants nowadays, namely the WiFi probe requests [4–6]. In this paper, we will use the WiFi sniffers to capture the WiFi probe requests for occupant counting in smart home.

Authors in [7] are the first to use WiFi probe request to coarsely estimate the location of users using the range-based maximum likelihood estimation and the weighted k-nearest neighbour algorithm. And authors in [8] focus on dealing with the intermittent nature of probe request using various number of WiFi sniffers to estimate the locations of users. Time series models including Auto-Regression Integrating Moving Average models and Long Short-Term Memory are used to forecast the number of occupants at a given location [9]. And in [10], a detection method called One-Class SVM (OC-SVM) is employed and a mathematical approximation of accuracy is derived, which utilises the characteristics of wireless signals as well as OC-SVM. Specifically, the received signal strength of probe requests at the APs is elaborately processed to improve the accuracy.

Instead of range-based localization of users in buildings based on the WiFi probe requests required to predict the occupant counting as in [7–10], in our work, to estimate the number of occupants in a target area, a simple binary-location-classifier is first proposed to determine whether a detected object is inside or outside of the target area without any range-based localization technologies required. Then, the sum of the individual target smart terminal in a zone gives the occupant count of that zone. Moreover, another contribution is that, in order to improve the accuracy of the classifier, multiple WiFi sniffers are utilized in our work which proves effective both theoretically and experimentally.

The remainder of this paper is organized as follows: Sect. 2 presents the basic assumption and the system framework. In Sect. 3, we discuss the multiple WiFi sniffers-based classification model for occupant counting. Section 4 presents our experimental results, and finally, Sect. 5 concludes the paper.

2 Basic Assumptions and the System Framework

Since smartphones have become an important part of our daily lives due to their capabilities of accessing the web using WiFi and mobile data networks, we assume that the occupants in buildings are always carrying their smartphones with themselves,

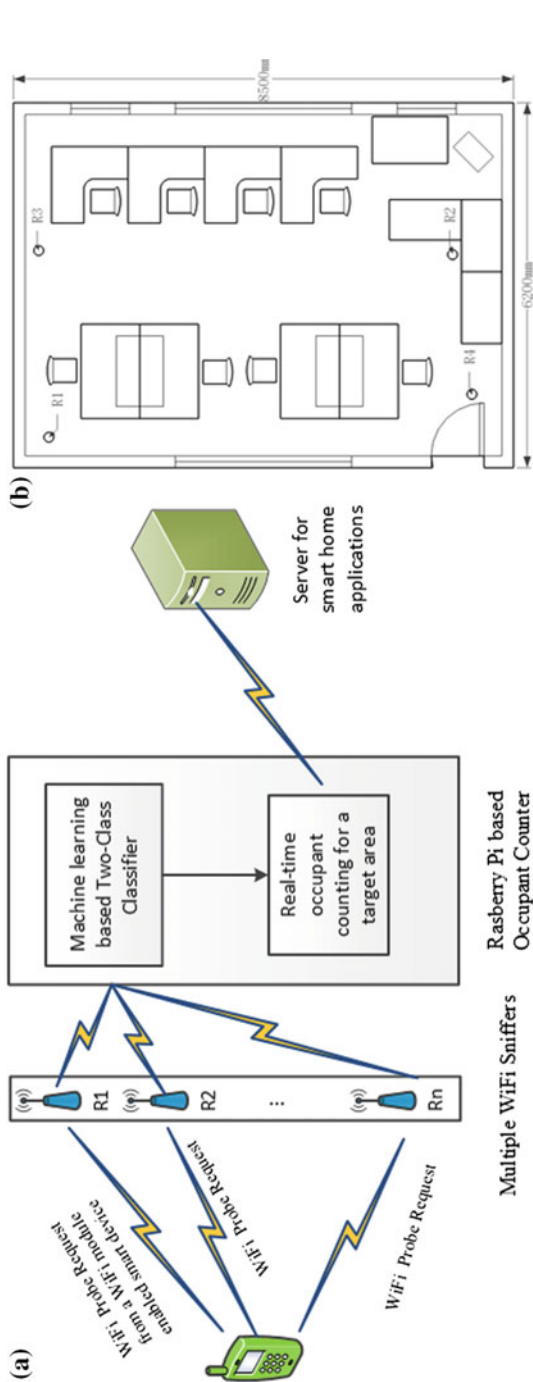


Fig. 1 Classification-based real-time occupant counting using probe requests that are captured at multiple WiFi sniffers deployed in each targeted area/zone. **a** System framework; **b** multiple WiFi sniffers deployed in one given targeted area/zone

rendering that once a smartphone is detected inside our targeted area, the number of occupants in the targeted area will be increased by one. However, in our method it is impossible to detect occupants, who do not carry smartphones or without their WiFi modules enabled; providing some other methods (e.g. using surveillance video or images) that can supplementarily correct this error is part of our future work and not considered to be within the scope of this study.

As shown in Fig. 1a, the system framework of our proposed classification based occupant counting method is simple and easy to implement. First, multiple wireless sniffers $R = (R_1, R_2, \dots, R_n)$, are deployed in the given target area as shown in Fig. 1b, to track the probe requests sent by those WiFi equipment. The data captured at WiFi sniffers includes time stamps, MAC address of the WiFi-enabled device, and the signal strength of the probe request. Then, a classification model based on the received signal strength vector by multiple WiFi sniffers is introduced to determine whether a detected smart terminal is inside or outside of the given target area. The classifier is achieved by machine learning techniques. Specifically, the RSSI vector is the input, and the binary result denoting whether or not the WiFi terminal is in the designated area is the output. Finally, the sum of the individual detected targets in a zone gives the occupant counting of that zone.

3 Multiple WiFi Sniffers-Based Classification Model for Occupant Counting

Instead of accurately positioning of an occupant, whether an occupant is inside or outside of a specific area is required to solve the occupant counting problem. Therefore, a simple binary classifier will solve our problem. The goal is to learn a function (hypothesis) $h: X \mapsto Y$, where X is the set of features that we want to classify with $Y = \{0, 1\}$. The training of the classifier is made with a sample set S of n labeled instances:

$$S = \{(x^{(i)}, y^{(i)} \triangleq h(x^{(i)}))\}_{i=1}^n \quad (1)$$

where $x \in X$, $y \in Y$ and S is assumed to be drawn from a fixed distribution $D_{X \times Y}$. Notice the use of the superscript to index the different training instances, and the notation $h(x^{(i)})$ is used as the target value of $x^{(i)}$. Specifically, $x^{(i)} = \{RSSI_1^{(i)}, RSSI_2^{(i)}, \dots, RSSI_n^{(i)}\}$, which is obtained by n WiFi sniffers.

With the above defined binary classifier, in this paper we focus on the accuracy performance of the classifier. The following will give the definition of accuracy of the binary classifier.

3.1 Accuracy Analysis for Multiple-WiFi Sniffers-Based Binary Location Classifier

We refer $RSSI_{ij}$ as the representative RSS value for terminal i at WiFi sniffer j , which can be written as

$$RSSI_{ij} = P_0 - 10n\log_{10} \frac{d_{ij}}{d_0} + \chi \quad (2)$$

where P_0 is the signal strength at the reference distance d_0 , n is the path loss exponent, d_{ij} is the Euclidian distance between terminal i and WiFi sniffer j , and χ is the additive white Gaussian noise with a variance of σ^2 and mean of zero.

Once the RSS of terminal i at WiFi sniffer j is obtained, assuming that the event Y that terminal i inside of the target area happens, $Y = 1$, otherwise, $Y = 0$, the accuracy of 1-WiFi-sniffer-based binary classifier (take WiFi sniffer j as an example) can be written as:

$$\begin{aligned} p_j &= p(1)p(1|1) + p(0)p(0|0) \\ &= p(1) \int_{R_{th_1}}^{+\infty} f_1(x)dx + p(0) \int_{-\infty}^{R_{th_0}} f_0(x)dx \end{aligned} \quad (3)$$

where $p(1)$ is the probability of one smart terminal being inside of the target area, $p(1|1)$ is the condition probability of one terminal being determined inside the targeted area when it is inside the targeted area, $p(0)$ is the probability of one smart terminal being outside the target area, $p(0|0)$ is the condition probability of one terminal being determined outside the targeted area when it is outside the targeted area, $f_1(x)$ is the probability density function (pdf) of the RSS for one terminal being inside the target area, R_{th_1} is the RSS threshold for a terminal inside the targeted area to be determined inside the targeted area, $f_0(x)$ is the probability density function (pdf) of the RSS for one terminal being outside the targeted area, R_{th_0} is the RSS threshold for a terminal outside the targeted area to be determined outside the targeted area.

Let $P(m)$ be the accuracy incurred m -WiFi sniffers-based classifier, where m is a positive integer. We will show that the accuracy incurred by multiple- WiFi-sniffers-based classifier is greater than or equal to that of only 1-WiFi sniffer-based classifier.

Theorem For $\forall m, n, m \in N^*, n \in N^*,$ if $m \geq n$, then $P(m) \geq P(n)$.

Proof First, without loss of generality, $P(m)$ can be written as

$$P(m) = 1 - \prod_{k=1}^m (1 - p_k),$$

Assume $m \geq n$, then

$$\begin{aligned}
 P(m) - P(n) &= \left(1 - \prod_{k=1}^m (1 - p_k)\right) - P(n) \\
 &= \left(1 - \prod_{k=1}^n (1 - p_k) \prod_{k=n+1}^m (1 - p_k)\right) - P(n) \\
 &= \left(1 - (1 - P(n)) \prod_{k=n+1}^m (1 - p_k)\right) - P(n) \\
 &= (1 - P(n)) \left(1 - \prod_{k=n+1}^m (1 - p_k)\right) \\
 &= (1 - P(n)) P(m - n) \\
 &\geq 0.
 \end{aligned}$$

Therefore, if $m \geq n$, then $P(m) \geq P(n)$. ■

With the Theorem, let $n = 1$, it is easily obtained that the accuracy incurred by multiple-WiFi sniffers-based classifier is greater than or equal to that of only 1-WiFi-sniffer-based classifier.

3.2 Neural Networks Based Binary Classifier Design

As described in Eq. (1) and shown in Fig. 2, the Neural Networks (NNs) considered here correspond to maps of the form:

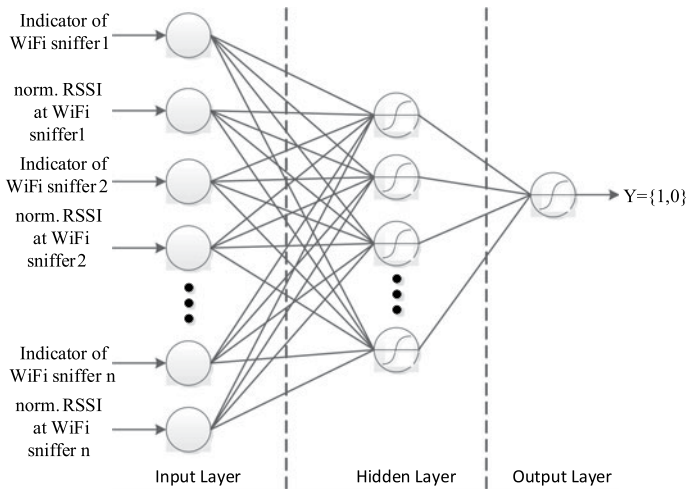


Fig. 2 Implementation of binary location classifier using neural networks

$$y_k = \psi'' \left(\sum_i \omega_i'' h_i(x_k) + \theta'' \right) \quad (4)$$

where the structure of the function ψ'' , and three parameters ω_i'' , h_i and θ'' are required to be determined. Once the structure of the function ψ has been chosen, training the network corresponds to search that particular set of parameters ω_i'' , h_i and θ'' for which the corresponding NNs best fits the training items. It is noted that an indicator variable for the specified WiFi sniffer is introduced in Fig. 2 to demonstrate whether the WiFi sniffer is used. Specifically, the indicator variable is a boolean variable. When it is true, it means that the RSSI monitored by the specified WiFi sniffer is used in the Neural Network, otherwise it is neglected. The output of the Neural Network is also a boolean variable, where true stands for the monitored occupant is inside the target zone and false is outside the target zone.

In our tests, in order to get good performance, different settings for the number of hidden layers and the number of neurons for the NNs will be implemented.

3.3 Occupant Counting

Our final goal is real-time occupant counting for a given zone which can be used for smart home applications, by counting the number of occupants within coarsely defined zones. The results by BP Neural Networks-based binary classifier lead to mapping each smart terminal to a zone. The sum of the individual smart terminals in the zone is the occupant count result for that zone. In this case, it might be a problem of how to choose a proper time interval to estimate the number of occupants in the target zone. If the time interval is too small, not enough WiFi probe requests can be tracked by multiple-WiFi sniffers because of the intermittent nature of WiFi probe requests, while if the time interval is too long, the timeliness of these results will be not good for smart home applications. Therefore, based on some results from questionnaire survey, usually fifteen minutes to half an hour will be reasonable as the recommended time interval. In our experiments, half an hour is used to verify the effectiveness of our proposed approach.

4 Experimental Results

4.1 Experimental Setup

As shown in Fig. 1b, 4 WiFi sniffers are deployed in Room 401 of Teaching Building A at Anhui Jianzhu University to capture available WiFi probe requests for occupant counting. It is noted that both the number and the locations of the WiFi sniffers are

not elaborately optimized which will be our future work. We use 30 min as the time interval for occupant counting. The data collection lasts two days from December 11, 2017 to December 12, 2017, and the time duration for each day is from 9:00 to 18:00. Specifically, the data collected on Dec. 11, 2017 is used to train the Neural Network which is executed offline on a computer, and data for the other day is used to predict the real-time number of occupants in Room 401 that runs on a Raspberry Pi-based node in real time and distributively. And in the test, the real occupant number of Room 401 at time k , is got by real counting by a person.

4.2 Performance Evaluations and Discussion

First, we evaluate the mean accuracy of the binary-location-classifier using N WiFi sniffers, where $N \in \{1, 2, 3, 4\}$, labeled with 1-WiFi sniffers-based, 2-WiFi sniffers-based, 3-WiFi sniffers-based, and 4-WiFi sniffers-based, respectively, as shown in Fig. 3, and the variance is about 1% which is not shown in the figure. In this paper, the neural network is chosen to perform as the classifier. As in Fig. 3, in order to evaluate impacts of different structures of the neural networks on the accuracy, we test different number of hidden layers with various neurons, for instance, [3], [5], [7], [20], [4, 2], [5, 2], [5, 3], [6, 3], shown as the x axis in Fig. 3. The length of the row vector for the x axis is the number of hidden layers, and the j th element of the row vector is the number of neurons included in the j th hidden layer. It is shown that, the accuracy of multiple-WiFi-sniffers-based classifier varies little with the configurations of the neural networks. However, the number of WiFi sniffers used in the classifier plays an important role in the accuracy. The mean accuracy of 4-WiFi sniffers-based Classifier is higher than that of n -WiFi sniffers-based (i.e., $n = 1, 2, \text{ or } 3$) Classifier, which coincides with Theorem proved in Sect. 3.1.

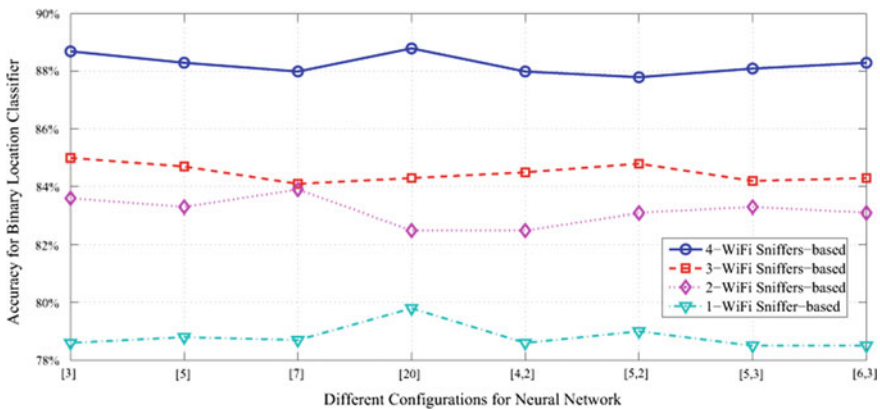


Fig. 3 Accuracy of the binary-location-classifier with different number of WiFi sniffers

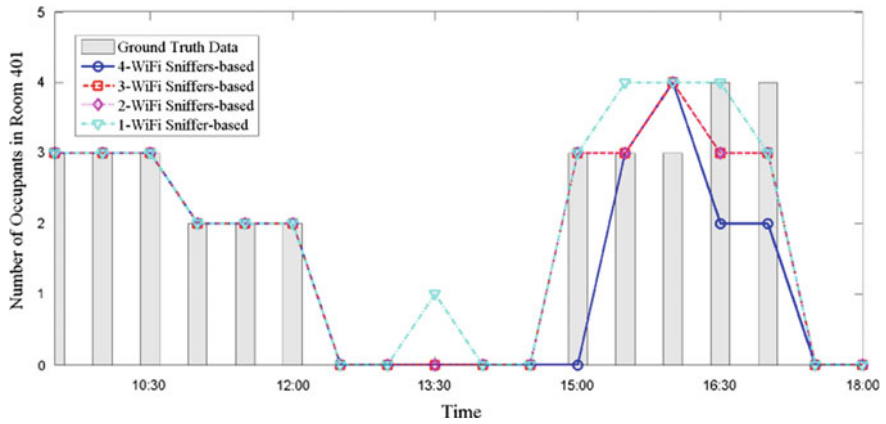


Fig. 4 Number of total people in Room 401 per half an hour on Dec. 12, 2017

Secondly, as shown in Fig. 4, the total number of occupants in Room 401 is given with respect to time. Specifically, we record the number of occupants at the discrete time slices predicted by our 4 methods, namely 1-WiFi sniffers-based, 2-WiFi sniffers-based, 3-WiFi sniffers-based, and 4-WiFi sniffers-based, respectively, and then sum them up from 9:00 to 18:00 for each method to get the overall number of occupant for the whole day. The results are 34, 31, 31, and 26, respectively, for 1-WiFi sniffer-based, 2-WiFi sniffers-based, 3-WiFi sniffers-based, and 4-WiFi sniffers-based. And the real number of occupants obtained by real counting is 32. Therefore, the occupant counting accuracy of the above four mentioned methods is, 93.75%, 96.88%, 96.88%, and 81.25%, respectively. That is because, for 1-WiFi sniffers-based method, since the accuracy of the classifier is lower than that of 2-WiFi sniffers-based method and 3-WiFi sniffers-based method, the occupant counting accuracy is reasonably lower. Moreover, because of the unreliable wireless links, the probability of probe requests captured by 1 WiFi sniffer is higher than that of multiple-WiFi sniffers, the occupant number predicted by 1-WiFi sniffers-based method is larger than the real number, and the occupant number predicted by 4-WiFi sniffers-based method is always less than the real number. Therefore, although the accuracy of the 4-WiFi sniffers-based classifier is high, when the probe request is not captured simultaneously by 4 WiFi sniffers, our occupant counting method based on 4-WiFi sniffers-based method will not do any classification for counting, resulting in a lower counting and decrease the accuracy of occupant counting. That is why 4-WiFi sniffers-based occupant counting method has relatively low accuracy as seen in Fig. 4. Meanwhile, for 2-WiFi sniffers-based and 3-WiFi sniffers-based classifiers, since both of them have satisfactory accuracy and WiFi probe requests can be captured by them with a high probability, they have the best prediction ability under our test environments.

5 Conclusions and Future Work

In this paper, occupant detection for smart home is studied by passively capturing WiFi probe requests from smart terminals using multiple-WiFi sniffers. We propose a binary-location-classifier to determine the detected object whether it is inside or outside the target area, which is implemented by neural networks. Our results show that, for each target area, the accuracy of multiple-WiFi sniffers-based binary classifier is higher than that of 1-WiFi-sniffer-based classifier. This proposed classification-based method based on multiple-WiFi sniffers can be a viable solution for occupancy monitoring in future smart buildings, which can have application such as building energy management and surveillance. One of the challenges is that those who are not carrying a device will not be counted. This could be another research project that estimate the number of occupants that are missed and make it part of the overall accuracy calculation. Besides, our future work will also include the optimization of the number and the deployment of the WiFi sniffers to further improve the zone-level occupant counting performance.

Acknowledgements This work was supported by National Key Research and Development Project of China, No. 2017YFC0704100 (entitled New generation intelligent building platform techniques), National Experimental Teaching Demonstration Center (entitled Building Control and Energy Saving Optimization Experiment Center, Anhui Jianzhu University), National Natural Science Foundation of China (Grant No. 11471304), and Ph.D Research Startup Foundation of Anhui Jianzhu University (Grant No. 2017QD07).

References

1. Akkaya, K., Guvenc, I., Aygun, R., Pala, N., Kadri, A.: IoT-based occupancy monitoring techniques for energy-efficient smart buildings. In: *Proceedings of IEEE Wireless Communications Networking Conference Workshops*, pp. 58–63 (2015)
2. Huang, Q., Cox, R., Shaurette, M., Wang, J.: Intelligent building hazard detection using wireless sensor network and machine learning techniques. In: *International Conference on Computing in Civil Engineering*, pp. 485–492 (2012)
3. Labeodan, T., Zeiler, W., Boxem, G., Zhao, Y.: Occupancy measurement in commercial office buildings for demand driven control applications-a survey and detection system evaluation. *Energy Build.* **93**, 303–314 (2015)
4. Cisco: Location analytics. https://meraki.cisco.com/lib/pdf/meraki_whitepapeccmxcmx.pdf
5. Musa, A.B.M., Eriksson, J.: Tracking unmodified smartphones using Wi-Fi monitors. In: *Proceedings of ACM Conference on Embedded Network Sensor Systems*, ser. SenSys'12, New York, NY, USA, pp. 281–294 (2012)
6. Kropet, T.: Don't trust open hotspots: Wi-Fi hacker detection and privacy protection via smartphone, BS Thesis (2015)
7. Vattapparamban, E., Ciftler, B.S., Guvenc, I.G., Akkaya, K., et al.: Indoor occupancy tracking in smart buildings using passive sniffing of probe requests. In: *IEEE International Conference on Communications Workshops*, pp. 38–44 (2016)
8. Ciftler, B.S., Dikmese, S., Guvenc, I.G., et al.: Occupancy counting with burst and intermittent signals in smart buildings. *IEEE Internet Things J.* 1–11 (2017)

9. Qolomany, B., Al-Fuqaha, A., Benhaddou, D., Gupta, A.: Role of deep LSTM neural networks and Wi-Fi networks in support of occupancy prediction in smart buildings. In: The 15th IEEE International Conference on Smart City, Bangkok, Thailand, 18–20 Dec 2017
10. Nguyen, C.L., Khan, A.: WiLAD: wireless localisation through anomaly detection (2018). https://www.researchgate.net/publication/319416168_WiLAD_Wireless_Localisation_through_Anomaly_Detection

A p -Persistent Frequent Itemsets with 1-RHS Based Correction Algorithm for Improving the Performance of WiFi-Based Occupant Detection Method



Ping Wang, Huaqian Cao, Si Chen, Jiake Li, Chang Tu and Zhenya Zhang

Abstract Considering that existing device-based occupant detection methods cannot count those who do not carry a device, in this paper, for buildings where the behaviour of the occupants tends to be regular, taking the WiFi-based occupant detection method as a basis, we propose a p -persistent frequent itemsets with 1-right-hand-side (RHS)-based occupant detection algorithm to improve the occupant detection performance in terms of accuracy. Association analysis using apriori algorithm is utilized to predict the occupancy of buildings through mining the relationships among occupants. We mathematically prove the reasonability of frequent itemsets with 1-RHS chosen in our algorithm and show the experimental results of applying this approach with different p . The results show that our proposed method can improve the accuracy performance in that it can see the occupant in buildings that the WiFi-based occupant detection method cannot see.

Keywords Occupant detection · Association analysis · Apriori algorithm · Frequent itemsets · p -persistent

1 Introduction

Getting the real-time estimation of occupancy in buildings (knowing the number of occupants and where they are located in buildings or in a zone) proves crucial in many applications, such as automated HVAC control, anomaly detection and activity recognition. A significant amount of work has been done in the past two decades to enable accurate and robust occupants counting. Occupancy counting in

P. Wang · H. Cao · S. Chen · J. Li · C. Tu · Z. Zhang (✉)
Anhui Province Key Laboratory of Intelligent Building & Building Energy Saving,
Anhui Jianzhu University, Hefei, Anhui, China
e-mail: zzychm@ustc.edu.cn

P. Wang · H. Cao · S. Chen · J. Li · C. Tu · Z. Zhang
School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei,
Anhui, China

smart buildings can be implemented via video processing and camera systems or deployment of occupancy sensors, including CO₂ concentration sensor, RFID sensor, motion sensor, light sensor, temperature sensor, acoustic sensor, humidity sensor, passive infrared sensor and so on, throughout the building [1]. It is noted that for those device-based methods, once people do not carry a device with themselves or the device cannot work properly, the accuracy will decrease since not all the required inputs are available for the model to work [2, 3]. Consequently, it is worthy to investigate other potential methods to overcome these shortcomings of existing systems. An alternative option is to utilize the association rule mining for learning the relationship among the occupant behaviour patterns using the Apriori algorithm [4]. The association analysis is first applied to the problem of predicting occupant locations in [5], which allows prediction based on the time associations between the historical occupant movements and any available real-time information. In this paper, we will take the WiFi-based occupant counting method as an example, use the association analysis to find all rules that correlate the presence of one occupant with others, which can count the occupant that WiFi sniffers cannot see.

WiFi probe requests have been extensively used for location analytics and tracking, as well as privacy and security issues, and have recently spawn much interest in research [6, 7]. The authors in [8] are the first to use WiFi probe request to report efficiency of building occupancy tracking using WiFi probe requests and utilize the range-based maximum likelihood estimation and the weighted k -nearest neighbour algorithm to coarsely estimation the location of users. Furthermore, the authors in [9] focus on dealing with the intermittent nature of probe request using various number of WiFi sniffers to estimate the locations of users. Time series models including Auto-Regression Integrating Moving Average models and Long Short-Term Memory are used to forecast the number of occupants at a given location [10]. And in [11], a detection method called One-Class SVM (OC-SVM) is employed and a mathematical approximation of accuracy is derived, which utilizes the characteristics of wireless signals as well as OC-SVM. For the WiFi, probe requests based occupant counting approach as in [8–11], one of the challenges is that those who are not carrying a device will not be counted. In order to solve this problem, in this paper, we propose a p -persistent frequent itemsets with 1-right-hand-side (RHS)-based correction algorithm for occupant detection.

The remainder of this paper is organized as follows: Sect. 2 presents the basic assumptions and system model. In Sect. 3, we discuss the characteristics of frequent itemsets with k -RHS and develop a p -persistent frequent itemsets with 1-right-hand-side (RHS)-based occupant detection algorithm. Section 4 presents our experimental results, and finally, Sect. 5 concludes the paper.

2 Basic Assumptions and System Model

It is also assumed that the occupant detection for buildings, where the behaviour of the occupants tends to be regular. The regularity may be based on time of a day, or particular days of a week or times of a year, rendering that we can find

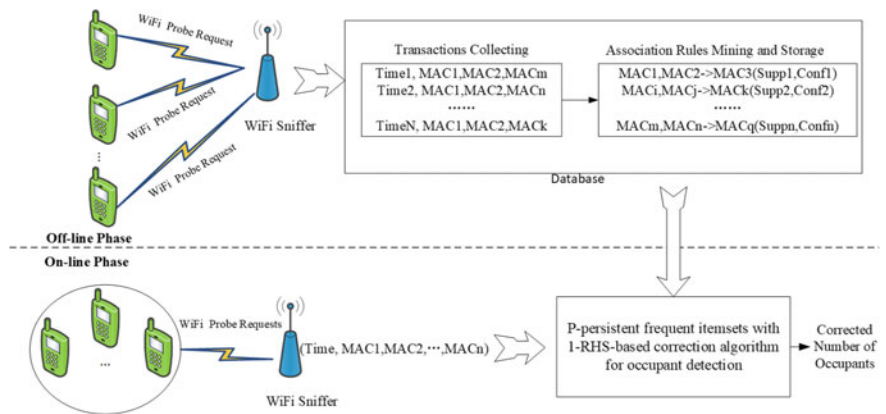


Fig. 1 Framework of an association analysis method using apriori algorithm for improving the performance of WiFi-based occupant detection approach

association rules among the occupants. Moreover, all the MAC addresses of the potential occupants' devices with wireless connection ability have been registered in advance.

All the registered MAC addresses $M = \{MAC_1, MAC_2, \dots, MAC_m\}$ can be modelled as a set of m attributes called *items*. Let $D = \{t_1, t_2, \dots, t_n\}$ be a set of transactions called the *database*. Each transaction in D is a set of MAC addresses which are detected by WiFi sniffers simultaneously and located in the same area or zone, has a unique transaction ID and contains a subset of the items in M . A rule is defined as an implication of the form $X \rightarrow Y$ where $X, Y \subset M$, and $X \cap Y = \emptyset$. The sets of items (for short itemsets) X and Y are called antecedent (left-hand-side or LHS) and consequent (right-hand side or RHS) of the rule, respectively.

With the above assumptions and definitions, our proposed p -persistent frequent itemsets with 1-right-hand side (RHS) correction algorithm for improving the performance of WiFi-based occupant detection approach is easy to implement. Generally, as shown in Fig. 1, it generally has two phases: the offline phase and the online phase. In the offline phase, association rule mining is executed based on the long-time collected data for those registered WiFi terminal by multiple WiFi sniffers, and the obtained frequent itemsets are all stored in an efficient way. In the online phase, for each time window, for each (pair or group of) detected MAC address frequently cooccurring in the same zone, the k -extension of high frequent itemsets is searched in the saved frequent itemsets. If the RHS of the detected MAC address is not included in the current detected set of MAC addresses, the number of the RHS will be counted with probability p ; otherwise, the current number of occupants keeps invariable. One of the challenges is the determination of k , which plays an important role in the performance of our method. Also, p will be discussed in the following.

3 Design of k -Extension of High Frequent Itemsets Based Correction Algorithm for Occupant Detection

In this paper, we focus on improving the accuracy of WiFi-based occupant detection algorithms using association analysis, which will solve the problem in the traditional WiFi-based occupant detection methods that those who are not carrying a device will not be counted. Specifically, the association rules, a powerful predictive tool, is utilized to discover groups of occupants that appear frequently together in the same zone in buildings. In this section, we will discuss the characteristics of the k -extension of high frequent itemsets, and propose a p -persistent frequent itemsets with 1-right-hand side (RHS)-based correction algorithm for occupant detection.

3.1 Discussion on the Characteristics of the k -Extension of High-Frequent Itemsets

To select interesting rules from the set of all possible rules, the minimum thresholds on support and confidence should be taken into consideration. In our work, the set of MAC addresses that are collected by WiFi sniffers simultaneously is assumed one transaction. The support $\text{supp}(\mathbf{X})$ of an itemset X is defined as the proportion of transactions in the data set which contains the itemset, and the *confidence* of a rule is defined:

$$\text{conf}(X \rightarrow Y) = \text{supp}(X \cup Y) / \text{supp}(X) \quad (1)$$

With the above definition, confidence can be interpreted as an estimate the probability $P(Y|X)$, the probability of finding the RHS of the rule in transactions under the condition that these transactions also contain the LHS. Therefore, in our association analysis-based correction algorithm for occupant detection, we prefer to utilize the association rules with high confidence. In the following, we will give a theorem to explain why we propose a 1-right-hand-side (RHS)-based p -persistent correction algorithm for occupant detection.

Theorem For $\forall$ high-frequent itemset X , define its k -extension as Y_k , and its $k+1$ -extension as Y_{k+1} , then $\text{conf}(X \rightarrow Y_k) \geq \text{conf}(X \rightarrow Y_{k+1})$.

Proof First, for association rules $X \rightarrow Y_k$ and $X \rightarrow Y_{k+1}$, their confidence can be written as

$$\text{conf}(X \rightarrow Y_k) = \text{supp}(X \cup Y_k) / \text{supp}(X) \quad (2)$$

$$\text{conf}(X \rightarrow Y_{k+1}) = \text{supp}(X \cup Y_{k+1}) / \text{supp}(X) \quad (3)$$

Obviously, it is known that $Y_k \subseteq Y_{k+1}$, then $(X \cup Y_k) \subseteq (X \cup Y_{k+1})$ holds.

Moreover, according to the anti-monotone property of $\text{supp}(\mathbf{X})$, $\text{supp}(X \cup Y_k) \geq \text{supp}(X \cup Y_{k+1})$. Therefore, $\text{conf}(X \rightarrow Y_k) \geq \text{conf}(X \rightarrow Y_{k+1})$. ■

Based on the above **Theorem**, it shows us that for two association rules, when their LHSs are the same, the confidence of k -extension is larger than or equal to that of $(k + 1)$ -extension. Therefore, to ensure the accuracy of our association analysis-based correction algorithm for occupant detection, we use the 1-extension of high-frequent itemsets to correct the number of occupants in buildings.

3.2 Design for p -Persistent Frequent Itemsets with 1-Right-Hand-Side (RHS)-Based Correction Algorithm for Occupant Detection

Algorithm 1: High Frequent Itemsets with 1-RHS Mining Algorithm	Algorithm 2: 1-RHS-based p -persistent Correction Algorithm
Input: DataCollectedByWiFiSniffers DataCol, Support Threshold Supp _{th} , Confidence Threshold Conf _{th} Output: High frequent itemsets with 1-RHS HFI – 1RHS	Input: MacAddCollectedByWiFiSniffers at TimeWindow t MAC _{t} , HighFrequentItemsets with 1-RHS HFI – 1RHS, Probability p Output: CorrectedNumber at TimeWindow t ΔN_t
<pre> 1 begin 2 if DataCol is NULL then 3 return; 4 end 5 // Initialization; 6 HFI – 1RHS = $\emptyset$; 7 TimeDiffRef = 1e10; 8 TransSet = $\emptyset$; 9 while DataCol is not NULL do 10 if TimeDiff(DataCol_{i}, DataCol_{j}) <= TimeDiffRef then 11 // Transform the data collected into transaction set 12 Store DataCol_{i}, DataCol_{j} in TransSet; 13 Remove DataCol_{i}, DataCol_{j} from DataCol; 14 end 15 while TransSet is not NULL do 16 if Supp_{i} >= Supp_{th} && Conf_{i} >= Conf_{th} then // Apriori 17 algorithm is used to generate association rules 18 if HighFrequentTrans_{i} 19 if Number of HighFrequent's RHS == 1 then 20 Store HighFrequent in HFI – 1RHS; 21 end 22 end 23 Remove Trans_{i} from TransSet; 24 end </pre>	<pre> 1 begin 2 if MAC_{t} is NULL then 3 return; 4 end 5 // Initialization; 6 ΔN_t = 0; 7 NewMACSetAdded = $\emptyset$; 8 TempMAC = $\emptyset$; 9 for MAC_{i} $\in$ MAC_{t} do 10 Use MAC_{i} to construct the LHS set; 11 Based on the LHS, match the association rules in HFI – 1RHS to get 12 the corresponding 1-RHS RHSofMAC_{i}; 13 TempMAC = RHSofMAC_{i}; 14 if TempMAC $\notin$ (MAC_{t} $\cup$ NewMACSetAdded) then // Include the 15 1-RHS in the MACSet if not in the current detected MAC set 16 TempVar = rand; 17 if TempVar $\leq p$ then // Include the 1-RHS with probability p 18 ΔN_t = ΔN_t + 1; 19 Store TempMAC in NewMACSetAdded; 20 end 21 end </pre>

According to the above analysis, in order to count people that do not carry a smart phone or without their phone's WiFi module enabled, we design a p -persistent frequent itemsets with 1-right-hand-side (RHS)-based correction algorithm for occupant detection.

Based on the framework as shown in Fig. 1, two core procedures need to be elaborately designed, namely, high frequent itemsets with 1-RHS mining and p -persistent frequent itemsets with 1-RHS-based correction algorithm.

For high-frequent itemsets with 1-RHS mining, apriori algorithm is utilized to generate rules as in Algorithms 1. First, the MAC addresses detected simultaneously by WiFi sniffers are grouped into one transaction. Then, based on the support requirement, high-frequent itemsets can be generated. Third, according to the confidence threshold, associate rules can be obtained. Finally, delete the association rules with more than 1-RHS and store the association rules with 1-RHS in an effective way.

For p -persistent frequent itemset with 1-RHS-based correction algorithm, for each time window, based on the detected MAC address set, the LHS is first constructed. Then, through searching the stored association rules, all the corresponding 1-RHS can be found out efficiently. If the 1-RHS is not included in the current detected MAC address set, it will be counted with probability p , where $0 < p \leq 1$. Here, we

count the matched MAC with probability p in that it is a probabilistic event not an absolute event. Specifically, in our work, we test two values for p , one is 1, and the other is equal to the confidence. Details can be found in Algorithm 2 and results will be given in Sect. 4.

4 Experimental Results

To test our approach, we conducted experiments in Room 401 of Teaching Building A at Anhui Jianzhu University. As shown in Fig. 2a, 4 WiFi sniffers are deployed in Room 401 to capture the available WiFi probe requests for occupant counting. We use 30 min as the time interval for occupant counting. The data collection lasts 32 days from December 23, 2017 to January 23, 2018, and the time duration for each day is from 9:00 to 18:00. Specifically, the data collected from December 23, 2017 to January 22, 2018 is used to do association rules mining which is executed offline on a computer, and data for the other day is used to predict the real-time number of occupants in Room 401 that run on a Raspberry Pi-based node in real time and distributively. And in the test, the real occupant number of Room 401 at time k , is got by real counting by person. Specifically, to verify the effectiveness of our proposed algorithm, on January 23, 2018, one occupant is required to turn off his smartphone’s WiFi module so that WiFi sniffers cannot detect him.

Based on the framework of our proposed method as in Fig. 1, transactions are first obtained and part of them are shown in Fig. 2b. Second, 1-RHS of high-frequent itemsets are generated by Algorithm 1 with support threshold of 50% and confidence

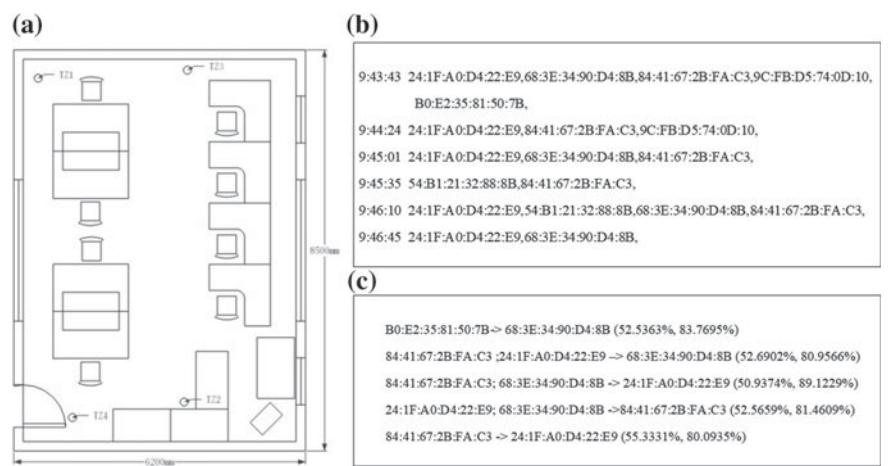


Fig. 2 Layout of the experiment and steps of association rules mining. **a** 4 WiFi sniffers deployed in Room 401 for occupant counting. **b** Group the collected MAC addresses into transactions. **c** Part of high frequent itemsets with 1-RHS in our test

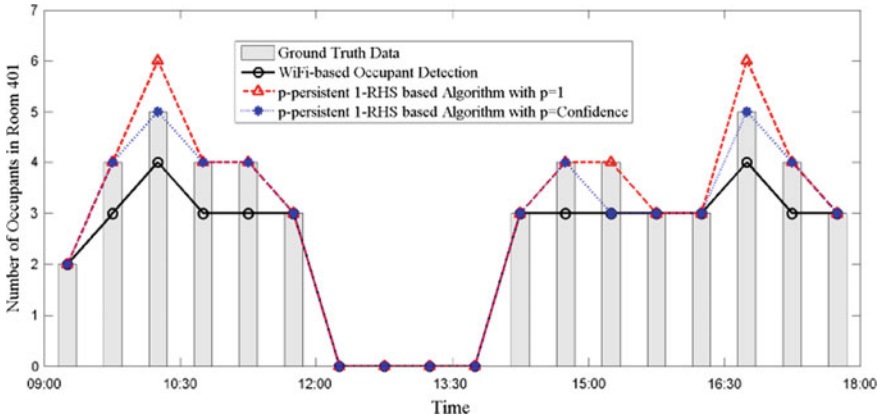


Fig. 3 Number of total people in Room 401 per half an hour on Jan. 23, 2017

threshold of 80%, and part of them are shown in Fig. 2b. Finally, with Algorithm 2, the predicted numbers of occupants with two configurations, $p = 1$ and $p = confidence$, respectively, are given in Fig. 3. Also for comparison, the result of WiFi-based occupant counting in [8] labeled with WiFi-based occupant detection is also given in Fig. 3.

It is found that for the WiFi-based occupant detection method, the occupant who turn off his smartphone's WiFi module cannot be counted always while our proposed 1-RHS-based p -persistent correction algorithm can solve this problem with a probability. When p takes the value of 1, it means that for an association rule, once its LHS is detected inside the target zone, the RHS will be counted if the RHS is not in the current detected MAC address set. In most scenarios, it is reasonable. However, there are some exceptions since an association rule describes a probabilistic event. This is why it is seen that in Fig. 3 the number of occupants predicted by 1-RHS based with $p = 1$ is mostly the same as the real number and larger than the real number sometimes. Moreover, 1-RHS based with $p = confidence$ correction algorithm is found to have a better performance than that with $p = 1$ by dealing with this problem as a probabilistic event. Therefore, the choice of the probability p is critical for the performance of our proposed correction algorithm. Further optimization and parameter tuning of the algorithm will be included in our future research.

5 Conclusions and Future Work

In this paper, a p -persistent frequent itemsets with 1-RHS-based correction algorithm is proposed to improve the performance of WiFi-based (those device-based) occupant detection methods by mining the association relationships between occupants in buildings. It can solve the undercount problem for those device-based occupant

detection methods. Our results show that, for each target area, this proposed association analysis-based correction method for occupant detection in buildings can improve the accuracy of that pure WiFi-based method. One of the challenges is the determination of probability p . This could be another research project. Besides, our future work will also include the optimization of the storage and matching for the association rules.

Acknowledgements This work was supported by National Key Research and Development Project of China, No. 2017YFC0704100 (entitled New generation intelligent building platform techniques), National Experimental Teaching Demonstration Center (entitled Building Control and Energy Saving Optimization Experiment Center, Anhui Jianzhu University), National Natural Science Foundation of China (Grant No. 11471304), and Ph.D. Research Startup Foundation of Anhui Jianzhu University (Grant No. 2017QD07).

References

1. Akkaya, K., Guvenc, I., Aygun, R., Pala, N., Kadri, A.: IoT-based occupancy monitoring techniques for energy-efficient smart buildings. In: Proceedings of IEEE Wireless Communications Networking Conference Workshops, pp. 58–63 (2015)
2. Huang, Q., Cox, R., Shaurette, M., Wang, J.: Intelligent building hazard detection using wireless sensor network and machine learning techniques. In: International Conference on Computing in Civil Engineering, pp. 485–492 (2012)
3. Labeodan, T., Zeiler, W., Boxem, G., Zhao, Y.: Occupancy measurement in commercial office buildings for demand driven control applications—a survey and detection system evaluation. *Energy Build.* **93**, 303–314 (2015)
4. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Data Bases, pp. 487–499 (1994)
5. Ryan, C., Brown, K.N.: Predicting occupant locations using association rule mining. In: 33rd SGAI International Conference on Artificial Intelligence, Cambridge, England, pp. 63–77 (2013)
6. Musa, A.B.M., Eriksson, J.: Tracking unmodified smartphones using Wi-Fi monitors. In: Proceedings of ACM Conference on Embedded Network Sensor Systems, ser. SenSys '12, New York, NY, USA, pp. 281–294. ACM (2012)
7. Kropeit, T.: Don't trust open hotspots: Wi-Fi hacker detection and privacy protection via smartphone, BS Thesis (2015)
8. Vattapparamban, E., Ciftler, B.S., Guvenc, I.G., Akkaya, K., et al.: Indoor occupancy tracking in smart buildings using passive sniffing of probe requests. In: IEEE International Conference on Communications Workshops. IEEE, pp. 38–44 (2016)
9. Ciftler, B.S., Dikmese, S., Guvenc, I.G., et al.: Occupancy counting with burst and intermittent signals in smart buildings. *IEEE Internet Things J.* 1–11 (2017)
10. Qolomany, B., Al-Fuqaha, A., Benhaddou, D., Gupta, A.: Role of deep LSTM neural networks and Wi-Fi networks in support of occupancy prediction in smart buildings. In: The 15th IEEE International Conference on Smart City (SmartCity 2017), Bangkok, Thailand, 18–20 Dec 2017
11. Nguyen, C.L., Khan, A.: WiLAD: wireless localisation through anomaly detection (2018). https://www.researchgate.net/publication/319416168_WiLAD_Wireless_Localisation_through_Anomaly_Detection

Day-Ahead Short-Term Optimization of Renewable Energy of Microgrid in Multiple Timescales



Xiaohui Wang and Shiqi Zong

Abstract Microgrid is an effective way to accept distributed renewable energy, and the development and application of renewable energy can effectively solve the current energy and environmental crisis. However, due to the uncontrollable and intermittent nature of renewable energy, coupled with the complexity of the operation modes of the microgrid, it is more difficult to optimize and control its operation, which has become a key issue in the energy management of microgrid. Considering the randomness of renewable energy, a multiple timescale optimization plan is proposed, which is a two-stage optimization scheme. The scheduling period of day-ahead optimization is 24 h. The targets of load supply and cost are selected as the objective function of an independent microgrid, and the power constraint of each distributed power source is set. The particle swarm optimization algorithm is used to optimize the system. Short-term optimization optimizes the results of day-ahead optimization for a second time, and takes 15 min as a scheduling period. The objective functions of revisions to the plan of day-ahead and the cost are selected and solved by the particle swarm optimization algorithm. The results are verified by the particle swarm optimization and show rationality and feasibility of the method proposed in the paper.

Keywords Microgrid · Energy management · Multiple time scales · Particle swarm optimization

1 Introduction

Microgrid is a small power generation and distribution system, which consists of distributed power supplies, energy conversion devices, loads, monitoring and protection devices, etc. It is an autonomous system that can realize self-control and management. It can make full use of local scattered natural resources, reduce environmental

X. Wang (✉) · S. Zong
Beijing University of Civil Engineering and Architecture, Beijing 100044, China
e-mail: wxhbucea@sina.com

pollution, slow down fossil energy shortage, and meet the diversified demands for power consumption [1].

There are still many islands and remote areas not yet powered by the power grid in China, and these areas are far away from the city, which will waste energy and cost to use the traditional power grid. An independent microgrid is an independent system which is divorced from the power grid. It relies on its power generation unit for its own power supply and is an important support for power consumption in remote areas. Independent microgrid can make full use of the local scenery resources and form a complementary, which can effectively improve the efficiency of renewable energy utilization and reduce the amount of diesel fuel to improve the environment and have greater economic and social benefits.

Because of the prominent effect of wind power and photovoltaic power generation, they are widely used in an independent microgrid. But because the output of wind and solar energy changes a lot with the weather, the power supply stability of generation units are not good. In order to make full use of these renewable resources, the scheduling of energy must be optimized, considering the microgrid stability and economy, forecasting renewable energy power generation, and scheduling reasonably and feasibly according to the load power requirements to ensure maximum load power consumption [2].

In the process of the optimal operation of a microgrid, it cannot be operated in short time, because this will not only increase the long-term operation cost of the microgrid but also have great influence on environment benefit of the microgrid. But the time scale is not the longer the better, because the prediction errors of intermittent energy such as wind and photovoltaic will increase with the timescale. Therefore, how to select the appropriate time scale and optimize the scheduling algorithm to minimize the error between the predicted results and the actual data is the key point of current research.

An independent microgrid optimization framework for multi-time scale is built in this paper, aiming at the randomness of renewable energy, and a two-level optimization scheme of day-ahead short-term is proposed. The day-ahead optimization is based on daily as a period, that is, each after 24 h; the optimization plan will conduct an automatic update to arrange in advance a combination scheme of a power unit in the next cycle. The short-term optimization is used to amend the results of day-ahead optimization with 15 min as a scheduling period. The optimal solution is calculated by the two-level optimization to select the optimal combination of the distributed power grid system and to optimize and schedule energy [3]. So that all the distributed energy sources in the microgrid can be reasonably supplied through the best unit commitment.

2 Optimization Principle

The optimization goal of the day-ahead plan is the most optimal economy and the greatest of a power supply according to solar power and load demand forecasting. It optimizes the future 24 h power generation based on the constraints and power

balance of each generation unit to achieve reasonable control operation and enhance the stability of power supply system [4–6].

Two objective functions should be met in the day-ahead optimization.

Objective function 1: The independent microgrid system can achieve the lowest total operating cost. The operation unit is 1 h, and the optimized length is 24 h.

$$Cost_{total} = Cost_{WT} + Cost_{PV} + Cost_{DG} + Cost_{Bat} + Cost_{device} \quad (1)$$

where $Cost_{total}$ is the total cost of a microgrid, yuan; $Cost_{WT}$ is the wind turbines cost, yuan; $Cost_{PV}$ is the PV array cost, yuan; $Cost_{DG}$ is the diesel generator cost, yuan; $Cost_{Bat}$ is the battery cost, yuan; $Cost_{device}$ is the microgrid device cost, yuan;

Objective function 2: The total load rate of independent microgrid is highest.

$$Load_{supply} = (W_{cl} + W_{el}) / (W_{clmax} + W_{el}) \cdot 100\% \quad (2)$$

where $Load_{supply}$ is the load supply rate; W_{cl} is the power supplied to general loads, kWh; W_{el} is the power supplied to important loads, kWh; W_{clmax} is the maximum amount of electricity supplied to general loads, kWh.

There are important loads and common loads in the independent microgrid, in which the important loads must be guaranteed not to be powered off. And because the weather changes dynamically all the time, the system power supply will be random and intermittent, which cannot guarantee the reliability of the power supply. It may produce some power off of the loads, and this effect can usually be expressed as follows:

$$C_{impact} = \frac{aW_{impact}^2}{W_m} + bW_{impact} \quad (3)$$

where C_{impact} is the effect on loads in the system, yuan; W_m is the maximum electricity demand of total loads, kWh; W_{impact} is the total power of affected loads, kWh; a, b is the constants dependent on loads, $a = 10, b = 0.5$.

The optimal principle is to minimize the total cost of $Cost_{total}$ and C_{impact} in the microgrid.

3 Microgrid Optimal Operation in Short Time

The followings are the objective functions which short-term optimization needs to satisfy.

Objective function 3: Based on the day-ahead optimization, in order to maintain the stability of the system, the scheme of minimum output power of short-term change is proposed, that is, the output power of each unit in short-term optimization should

be as close as possible to that of the day-ahead optimization, which can make stable operation of the independent microgrid system.

$$\Delta P = \Delta P_{WT} + \Delta P_{PV} + \Delta P_{DG} + \Delta P_{Bat} \quad (4)$$

where ΔP is the output variable of short time optimization, kW; ΔP_{WT} is the output power change of short-term optimization and day-ahead optimization of wind power generator, kW; ΔP_{PV} is the output power change of short-term optimization and day-ahead optimization of PV array, kW; ΔP_{DG} is the output power change of short-term optimization and day-ahead optimization of diesel generator, kW; ΔP_{Bat} is the output power change of short-term optimization and day-ahead optimization of battery, kW;

Objective function 4: The total operating cost of the independent microgrid is minimum.

$$Cost'_{total} = Cost'_{WT} + Cost'_{PV} + Cost'_{DG} + Cost'_{Bat} + Cost'_{device} \quad (5)$$

where $Cost'_{total}$ is the short-term optimization cost of independent microgrid, yuan; $Cost'_{WT}$ is the wind turbine cost of short-term optimization, yuan; $Cost'_{PV}$ is the photovoltaic array cost of short-term optimization, yuan; $Cost'_{DG}$ is the diesel generator cost of short-term optimization, yuan; $Cost'_{Bat}$ is the battery cost of short-term optimization, yuan; $Cost'_{device}$ is the microgrid cost, yuan.

According to the judgment principle of the optimal algorithm, each ΔP and the sum of corresponding $Cost'_{total}$ and C_{impact} will be calculated and the output power change should be minimum.

4 Case Study

The structure of an independent microgrid is shown in Fig. 1, and there are four groups of a photovoltaic array with total rated power of 40 kW; six wind turbine units with total rated power of 60 kW; three diesel generators with the total rated power of 80 kW, and one lead-acid battery with the rated power of 96 kW. In addition, there are important loads and general loads included in the microgrid system, so the loads are powered hierarchically and the power supply of important loads cannot be powered off.

The daily maximum output generating a capacity of wind power generator is 603.16 kWh, which fluctuates near the daily maximum power in the day-ahead optimization. The maximum output generating capacity of PV array is 223.1 kW, which as well as fluctuates near its daily maximum power in the day-ahead optimization. Therefore, the total maximum allowable power output of renewable energy is 826.26 kWh. Figure 2 is the upper limit of the output power of the PV array, and Fig. 3 is the upper limit of the output power of the wind generator.

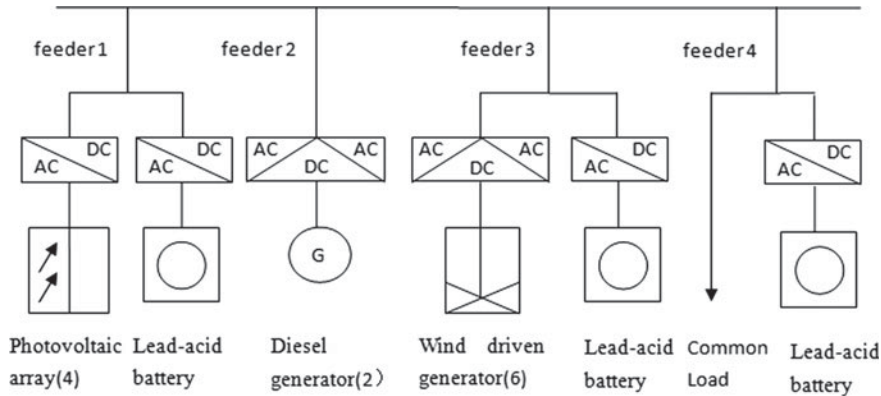


Fig. 1 Microgrid model

Fig. 2 The maximum output power of photovoltaic array

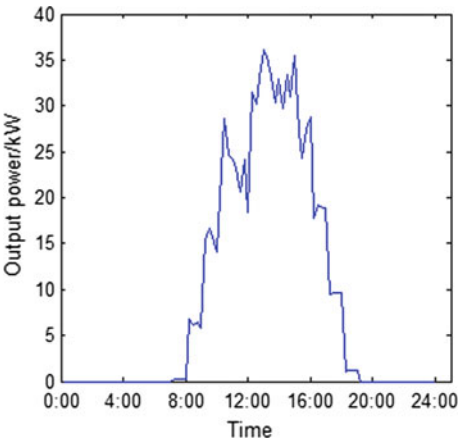
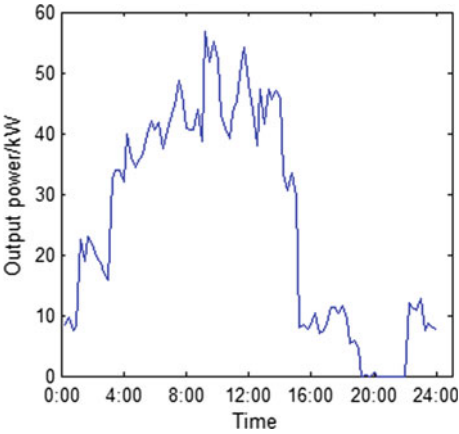


Fig. 3 The maximum output power of wind turbine



By entering the objective functions and constraints, the optimization is executed based on PSO. First, a group of random particles is initialized, and they are iterated constantly until the global optimal solution is found. The optimization group number of day-ahead optimization is 30, iteration time is 50, and the particle dimension is 24, which are used to search the global optimal.

5 Results and Discussion

Through the analysis on output power configuration of each distributed generation unit, storage devices and a diesel generator in the day-ahead optimization, the results are brought into short-term optimization scheme to take further optimization.

The comparison between the day-ahead optimized output power and the short time optimized output power of the wind turbines in the microgrid is shown in Fig. 4. It can be seen from Fig. 4a, the adjustment to day-ahead optimization scheme of wind turbine generator is large in the two periods of 7:00–8:00 and 11:00–12:00, and that of the whole day is very small, which fluctuates around the optimization scheme. And it can be seen that the adjustment to the day-ahead optimization scheme is large in five periods of 10:00–11:00, 11:00–12:00, 12:00–13:00, 13:00–14:00, and 14:00–15:00, and the sunshine of these periods is the biggest, so the short-term optimization scheme changes with fluctuation. Comparisons between output powers of short-term optimization and day-ahead optimization of PV array are shown in Fig. 4b.

Comparisons between output powers of short-term optimization and day-ahead optimization of the lead-acid battery are shown in Fig. 5. It can be found that diesel generator is enabled in the period of 00:00–1:00 in the short-term optimization, and

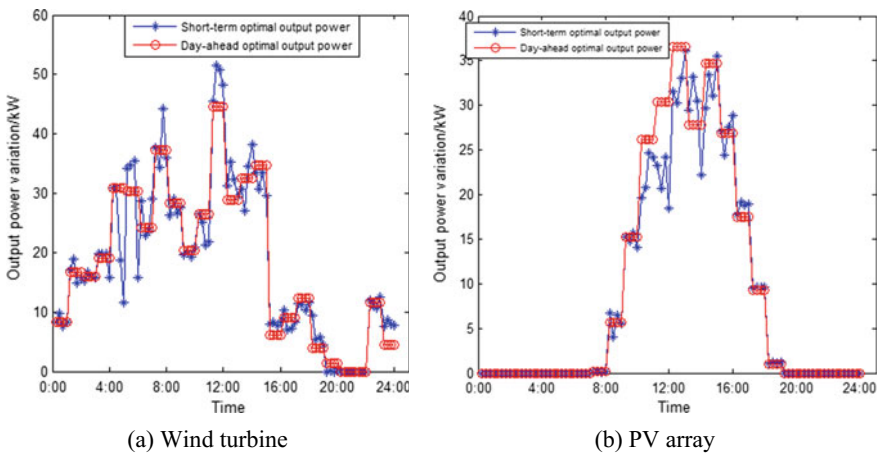
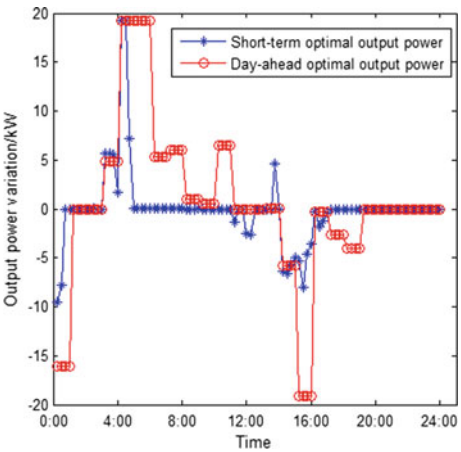


Fig. 4 Comparison of short-term optimal output power with day-ahead optimal results

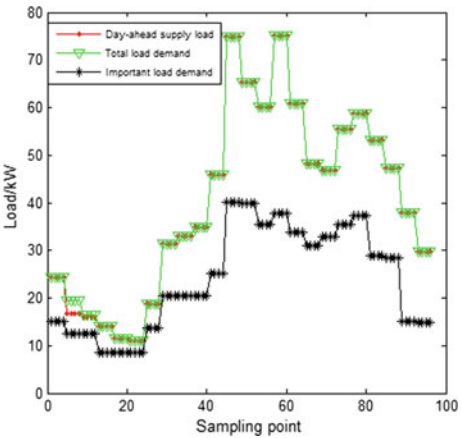
Fig. 5 Comparison of short-term optimal output power of battery with day-ahead optimal results



diesel generator output power in the period of 15:00–16:00 is higher than that of the day-ahead optimization scheme. Meanwhile, both the battery groups in the two periods are reduced output. The adjustments to the day-ahead optimization of the battery group in the periods of 00:00–1:00, 5:00–8:00, 10:00–11:00, 15:00–16:00, and 17:00–19:00 are very large, and the adjusted output curve is more smooth and coherent, so the optimization scheme is more reasonable. Comparisons among the day-ahead load supply, total load demand and significant load demand are shown in Fig. 6.

Comparisons among the short-term load supply, total load demand, and significant load demand are shown in Fig. 7. The adjustment amount to day-ahead optimization and the corresponding cost are shown in Fig. 8.

Fig. 6 Comparison among day-ahead supply load, total load demand, and important load demand



Figures 6 and 7 show the contrast among day-ahead load supply, total load demand and significant load demand, and the contrast among short-term load supply, total load demand, and significant load demand. It can be found that partial blackouts are implemented in short-term optimization at three periods of 11:00–12:00, 14:00–15:00, 15:00–16:00 based on the day-ahead optimization to satisfy the load demand, and the others are basically the same with very small change. Figure 8 shows the finally optimal result of adjustment amount of 12.2 kWh and the operation cost of 1611.8 yuan.

In summary, the short-term optimization scheme can guarantee the performance of the day-ahead adjustment smaller and can meet the stability of the system. At the same time, it is reliable for the output power of the equipment module to be further optimized and the quality of the optimization results is very high.

Fig. 7 Comparison among short-term supply load, total load demand and important load demand

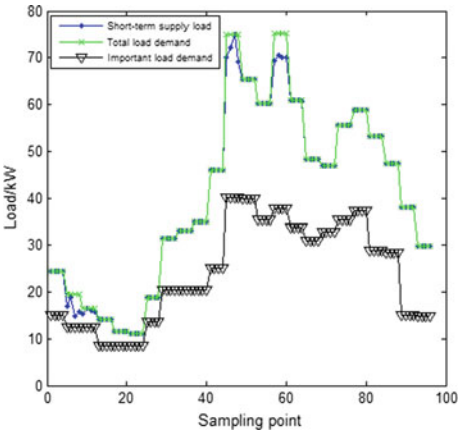
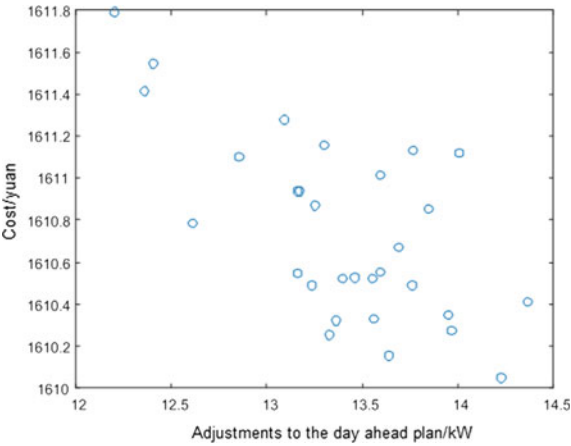


Fig. 8 Optimization adjustment and the corresponding cost



6 Conclusions

The multi-time scale optimization of an independent microgrid is executed based on the equipment model and constraint conditions, which uses the day-ahead optimization and short-term optimization scheme in follows, through the selection of objective functions and optimization process. It is shown through the analysis of the microgrid shows that the adjustment of the final optimization scheme to the day-ahead plan is relatively small, which makes the short-term optimization again based on the day-ahead optimization. It can also satisfy the economy and stability of the system, and is the optimal unit combination and energy scheduling strategy.

Acknowledgements This research is supported by the Fundamental Research Funds for Beijing University of Civil Engineering and Architecture (Project Number: X18031); National Key R&D Program of China, Research and Demonstration of Key Technology of Net-Zero Energy Building (Project Number: 2016YFE0102300); Research Fund Project of Beijing University of Civil Engineering and Architecture (Project Number: 00331616040); Research Program of Beijing City Board of Education (KM201610016002); and Science and Technology Project of Ministry of Housing and Urban-Rural Construction, China (Project Number: 2015-K1-012).

References

1. Zhang, J.H., Huang, W.: Control and protection technology of microgrid operation. China Electric Power Press, Beijing, China (2010) (in Chinese)
2. Liserre, M., Sauter, T., Hung, J.Y.: Future energy systems, integrating renewable energy sources into the smart power grid through industrial electronics. *IEEE Ind. Electron. Mag.* **4**(1), 18–37 (2010)
3. Zhao, B., Zhang, X., Chen, J., et al.: Operation optimization of standalone microgrids considering lifetime characteristics of battery energy storage system. *IEEE Trans. Sustain. Energy* **4**(4), 934–943 (2013)
4. Shouxiang, W., Dong, W., Liang, H.: Interval linear programming method for day-ahead optimal economic dispatching of microgrid considering uncertainty. *Autom. Electric Power Syst.* **38**(4), 5–11 (2014)
5. Jiang, Q.Y., Xue, M.D., Geng, G.C.: Energy management of microgrid in grid-connected and stand-alone modes. *IEEE Trans. Power Syst.* **28**(3), 3380–3389 (2013)
6. Bae, Hyangryul, Tsuji, Takao, Oyama, Tsutomu: A study on feasible region for islanded operation mode in power systems with distributed generators in emergency. *J. Int. Council Electr. Eng.* **4**(4), 275–280 (2014)

Modeling of Multiple Heating Substations Based on Long Short-Term Memory Networks



Qi Li, Bingcheng Han, Mingwei Yu and Jianglan Shang

Abstract The central heating is a complex nonlinear system. It is difficult to establish an accurate model based on multiple heating substations. In this paper, the Long Short-Term Memory (LSTM) algorithm is proposed to solve this problem. Heating substations generate data with the time series characteristics. The algorithm not only reflects the characteristics of time sequence of heating substations, but also solves the problem of long-term dependence. And, the necessary information can be saved in a limited memory capacity. Based on a large amount of historical data of the heating system of a Baotou heating company, ensuring that the total heat source is sufficient, the simulation results of the LSTM model in multiple substations show the validity, which provides the basis for the optimization of the central heating system, and a reference for LSTM to solve the complex time series modeling and prediction problems.

Keywords Time sequence · Long-term dependence · LSTM · Modeling of multiple heating substations

1 Introduction

Central heating [1] has become a popular heating method in northern China due to its advantages of energy saving, environmental protection, economy, and high efficiency. However, the central heating system is a multi-variable, strongly coupled, time-varying, and large-lag nonlinear system. It is difficult to establish a relatively accurate heating station model.

In the related research of the heating power station system, it is difficult to obtain accurate mathematical models by using the traditional mechanism modeling and experimental modeling methods for heating station modeling [2]. Heating substations

Q. Li (✉) · B. Han · M. Yu · J. Shang
College of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou, China
e-mail: richey@imust.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_48

515

generate data with a chronological character. Long Short-Term Memory (LSTM) neural network has a strong ability of fitting and processes the characteristics of time series data.

The LSTM-based system [3] can complete the tasks such as translation of language [4], image analysis [5], speech recognition, image recognition, and handwriting recognition. Vinayakumar et al. [6] used LSTM architecture to Android malware detection. These examples indicate that LSTM has good learning ability and generalization ability.

For the modeling of multiple heating substations, this paper proposes the modeling of multiple heating substations based on LSTM. When heat source total heat is sufficient, according to a large amount of historical data, the heating substations is modeled as a black box, which only considers the relationship between input and output, using the strong approximation capability of the neural network.

2 The Principle of LSTM Network

The LSTM is a neural network that is further improved on the basis of RNN. It not only solves the problem of long-term dependence, but also preserves the necessary information in a limited memory capacity.

The LSTM adds a cell to determine whether the information is useful or not. Three gates were deliberately added to the cell, including the input gate, the forget gate, and the output gate. The Sigmoid activation function is added to each gate to make the output within the range of 0–1, allowing the information to be selectively removed or added to the cell to avoid the long-term dependence problems.

The ingenuity of LSTM is that the weight of the self-loop is changed by adding the gate, so that the integral scale of different time can be changed dynamically in the case of the fixed parameters of the model, thus avoiding the problem of gradient disappearance or gradient expansion.

3 Modeling of Multiple Heating Substations Based on LSTM

3.1 Multiple Heating Substations Model Structures

In a central heating system, a heat source is responsible for supplying heat to a number of heat substations. The structure diagram of the central heating system is shown in Fig. 1. The relationship between the heating station and the heat source is primary, and the relationship between the heating station and the hot user is secondary.

Ten heat substations including Cai Hong, Huan Bao, Fu 7, Fu 10, Zi Lai Shui, Fu 8, Nong Hang, Fu 9, Xing Fu 8, and Xing Fu 6 station are chosen as the research

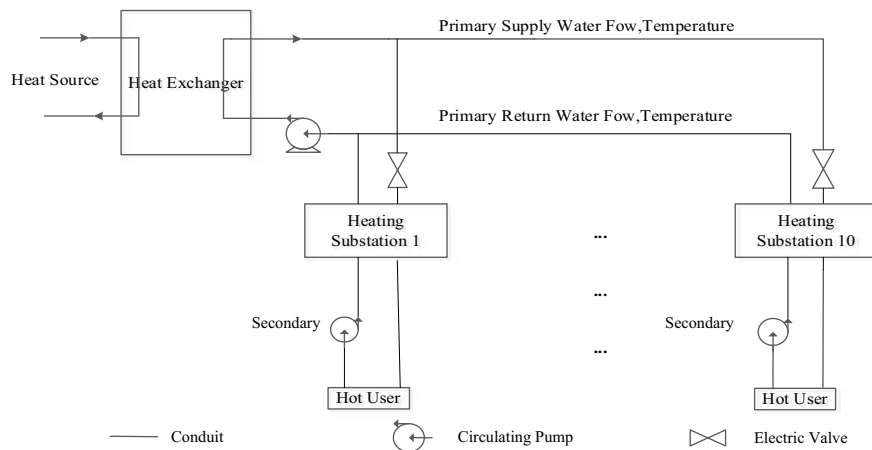


Fig. 1 Central heating system structure



Fig. 2 Distribution of heating substations

objects, which are on the main heating pipeline from Baotou Heating Company to Qingdong heat source. The distribution of the 10 heating substations is shown in Fig. 2.

The heating station is a complex sequential control system. According to the heating station system, the dynamic model expression is determined as follows:

$$y_n = \dot{f}(u_{1n}, u_{2n}, u_{3n}, t) \quad (1)$$

where n is the number of the heating substations; u_{1n} is the opening of the primary regulating valve of 10 heating substations at t time (primary supply water flow); u_{2n}

is the primary supply water temperature of 10 heating substations at t time; u_{3n} is the outdoor temperature at t time; y_n is the instantaneous heat of 10 heating substations at t time.

The weather factors, the great influence on the instantaneous heat, are added in control variables, which is more in line with the actual situation.

3.2 Data Selection and Preprocessing

The research object of this paper is 10 heat substations, simplifying the heat source and the heat users. Select 3600 sets of training data, which was sampled at intervals of 1 h, and was sampled 24 times a day from January 1, 2017 to January 15, 2015, as shown in Table 1. The input variables include the primary supply water flow, the primary supply water temperature, the outdoor temperature, and the output variable is instantaneous heat.

The data that are difference from the other collected data is removed. The input data are standardized to ensure that the characteristic data variance of each dimension is 1 and the mean value is 0. When the actual output data are standardized on the training set, the relative error of each heating substation in the test results gets smaller. The comparison results are shown in Table 2. Therefore, the data of three inputs and one output are standardized in the training set.

3.3 Model Structure and Parameters Based on LSTM

The typical LSTM neural network is composed of input layer, hidden layer, and output layer. Combining the LSTM with the heating substation principle, three variables are selected as input variables and one variable for output. The structure of multiple heating substations model based on LSTM is shown in Fig. 3.

The input and output variable is the same as in Eq. (1).

Table 1 Training datasheet

Time series (h)	Heating substation	Primary supply flow (t/h)	Primary supply temperature (°C)	Outdoor temperature (°C)	Instantaneous heat (GJ)
2017/1/1 0:00	Cai Hong	78.3	80.2	−13	1.2
2017/1/1 0:00	Huan Bao	54.42	80.7	−13	8.07
....					
2017/1/15 23:00	Xing Fu8	17.13	77.4	−1	2.45
2017/1/15 23:00	Xing Fu6	100.18	77.0	−1	13.08

Table 2 Comparison of modeling parameters based on LSTM heating substations

Training data	1510	1510	1510	1510	1510	1510	1510	3600
‘y’ normalized	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Node	20	20	50	50	50	20	30	20
Layer	3	3	3	5	2	2	2	3
CaiHong (%)	10.0	10.0	20.0	5.0	10.0	7.0	8.0	6.0
Huan Bao (%)	8.0	4.0	6.0	6.0	3.5	4.0	7.0	3.5
Fu7 (%)	5.0	5.0	12.0	10.0	6.0	3.5	6.0	4.0
Fu10 (%)	6.0	5.0	10.0	10.0	8.0	3.5	6.0	5.0
Zi Lai Shui (%)	6.0	4.0	8.0	7.0	10.0	7.0	7.0	7.0
Fu8 (%)	6.0	6.0	6.0	2.5	3.5	5.0	3.5	5.0
Nong Hang (%)	2.5	5.0	10.0	10.0	6.0	5.0	8.0	4.0
Fu9 (%)	5.0	4.0	4.0	2.0	6.0	3.0	3.5	3.5
Xing Fu8 (%)	7.0	3.5	2.5	3.5	8.0	3.0	6.0	2.5
Xing Fu6 (%)	7.0	2.0	8.0	7.0	10.0	7.0	6.0	6.0

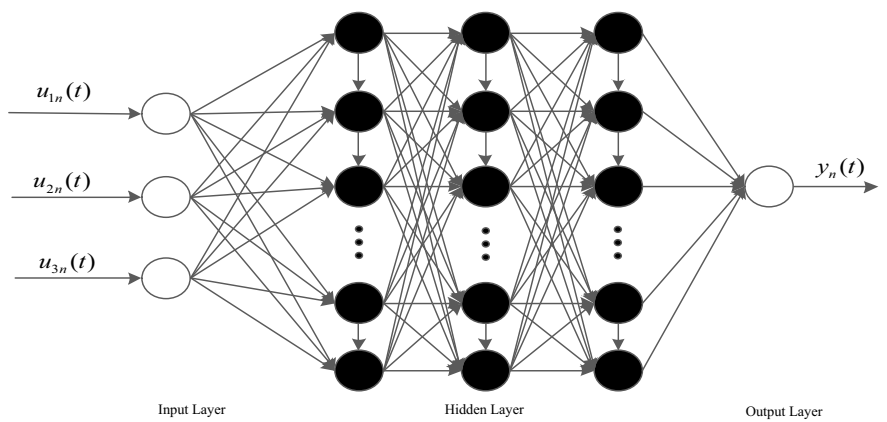


Fig. 3 LSTM-based heating substation model structure

The establishment of multiple heating substation system models based on LSTM requires the optimization of parameters. The relative error of the prediction results of the representative parameters in the optimization process is summarized in Table 2. After several adjustments, the final setting parameters of the model are shown in Table 3.

Table 3 Multiple heating substation model parameters based on LSTM

Parameter name	Parameter value	Parameter name	Parameter value
Input size	3	Node size	20
Output size	1	Layer number	3
Batch size	30	Learning rate	0.01
Time step	10	Loop times	60

4 Test Results

After several adjustments, the model structure uses a total of 3840 sets of data for modeling experiments. 240 sets of data of them are used as test sets, and the test results of the data are visualized in Figs. 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13.

The predicted instantaneous heat and the actual instantaneous heat deviation are within 0.8GJ and the relative error is within 7%, especially Xing Fu 8 stations can reach about 2.5%.

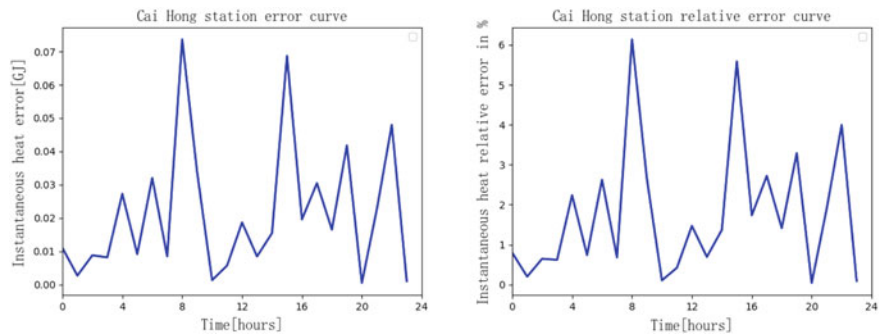


Fig. 4 Cai Hong station instantaneous heat error and relative error curve

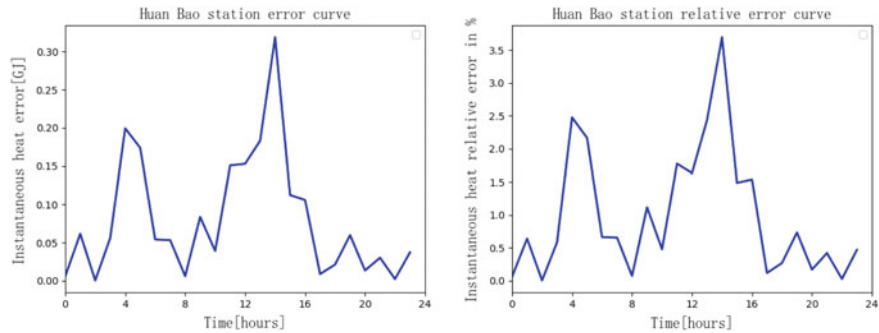


Fig. 5 Huan Bao station instantaneous heat error and relative error curve

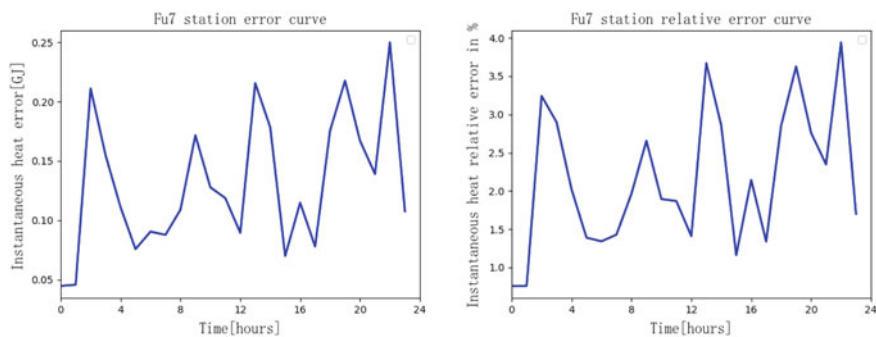


Fig. 6 Fu7 station instantaneous heat error and relative error curve

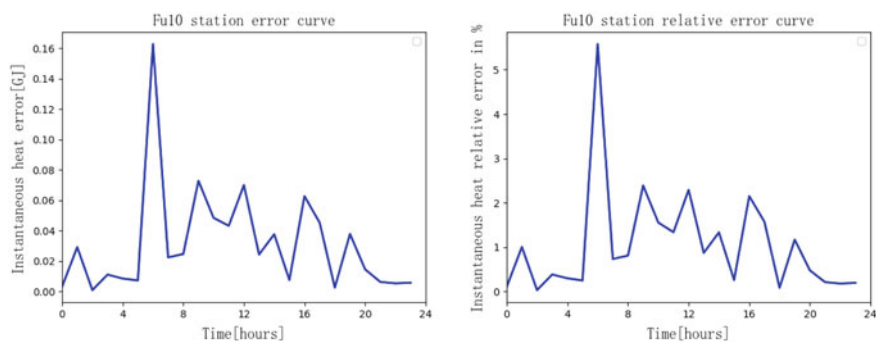


Fig. 7 Fu10 station instantaneous heat error and relative error curve

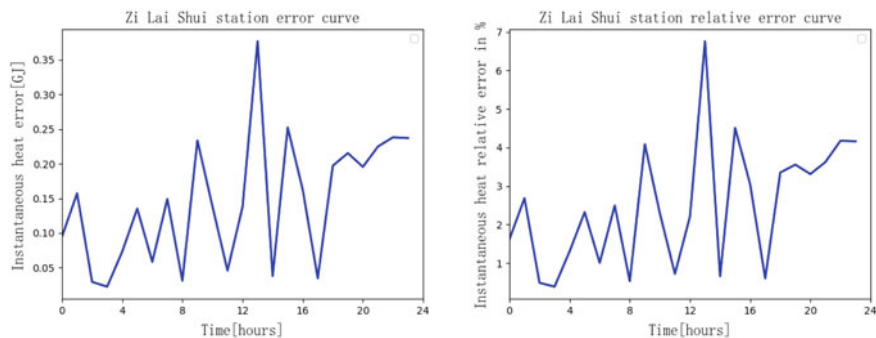


Fig. 8 Zi Lai Shui station instantaneous heat error and relative error curve

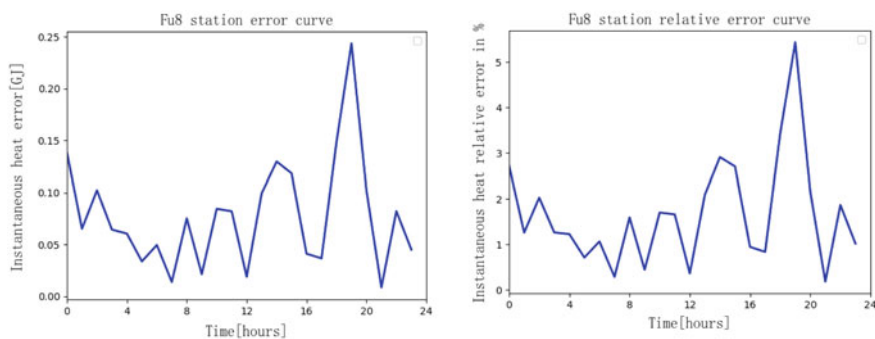


Fig. 9 Fu8 station instantaneous heat error and relative error curve

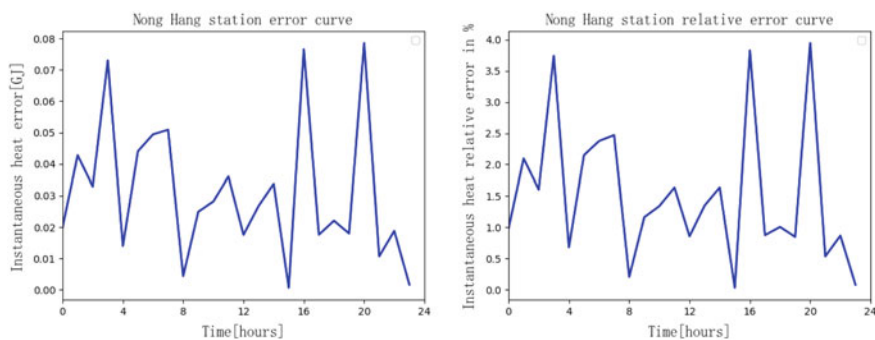


Fig. 10 Nong Hang station instantaneous heat error and relative error curve

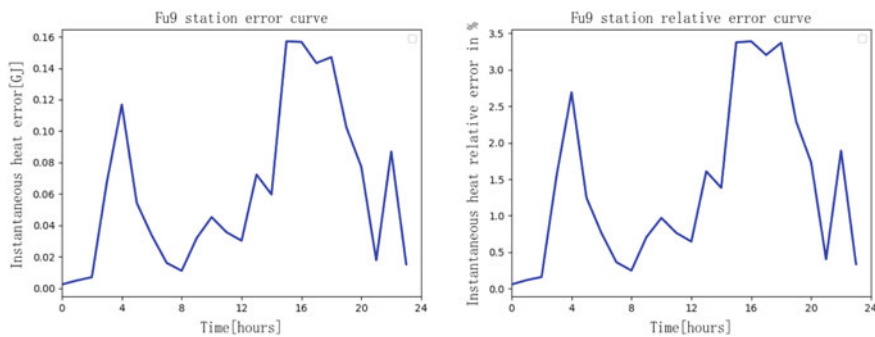


Fig. 11 Fu9 station instantaneous heat error and relative error curve

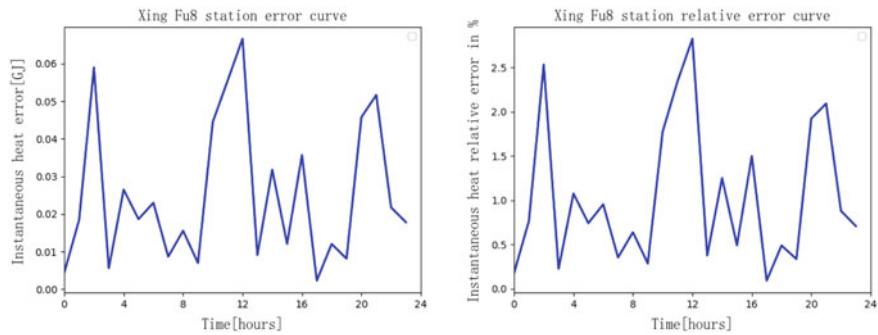


Fig. 12 Xing Fu8 station instantaneous heat error and relative error curve

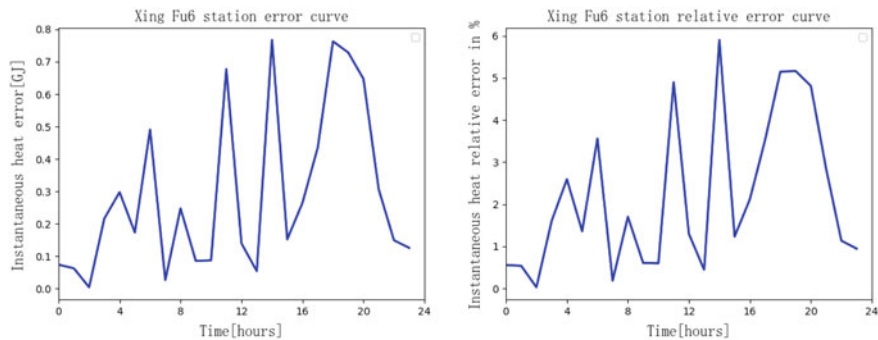


Fig. 13 Xing Fu6 station instantaneous heat error and relative error curve

From the data, it can be seen that the model of multiple heating substations based on LSTM networks meets the requirements of the instantaneous heat control accuracy required by the heating companies.

5 Conclusions

LSTM neural network can effectively deal with time series data, which reflects the data characteristics of the heating substations.

According to the model of multiple heating substations, the algorithm of LSTM is proposed to solve the problem of long-term dependence and establish a relatively accurate heating substations model. The simulation results of the LSTM model in the multiple heating substations show the validity, which provides the basis for the optimization of the central heating system, and a reference for LSTM to solve the complex time series modeling and prediction problems.

Acknowledgements The work was supported by National Natural Science Foundation of China (61463040).

References

1. Olsthoorn, D., Haghighat, F., Mirzaei, P.A.: Integration of storage and renewable energy into district heating systems: a review of modelling and optimization. *Sol. Energy* **136**, 49–64 (2016)
2. Schweiger, G., Larsson, P.O., Magnusson, F., Lauenburg, P., et al.: District heating and cooling systems—framework for modelica-based simulation and dynamic optimization. *Energy* **137**, 566–578 (2017)
3. Zhang, J.Y., Wang, H.L., Guo, Y., Hu, X.: Review of deep learning. *Appl. Res. Comput.* **35**(7), 1–12 (2018). (in Chinese)
4. Ma, N., Tian, G.D., Zhou, X.: A lip-reading recognition approach based on long short-term memory. *J. Univ. Chinese Acad. Sci.* **35**(1), 109–117 (2018). (in Chinese)
5. Song, J., Tang, S.L., Xiao, J., Wu, F., et al.: LSTM-in-LSTM for generating long descriptions of images. *Comput. Visual Media* **2**(4), 379–388 (2016)
6. Vinayakumar, R., Soman, K.P., Poornachandran, P., Kumar, S.S.: Detecting android malware using long short-term memory (LSTM). *J. Intel. Fuzzy Syst.* **34**(3), 1277–1288 (2018)

The Elman Network of Heat Load Forecast Based on the Temperature and Sunlight Factor



Qi Li, Shiqi Jiang and Xudan Wu

Abstract In urban district heating systems, the change of heat load is greatly influenced by various exterior factors. In order to meet the demand of heating system while achieving energy conservation and environmental protection, it is in this study, many kinds of artificial neural networks are compared, and a kind of Elman neural network is proposed for modeling heat load forecasting based on the temperature and the sunlight factor. The method obtains the real-time weather temperature from the Application Programming Interface (API) interface of the meteorological web site, added the illumination intensity as an input of the heat load forecasting model, and established the sample data sequence of the forecasting model. The real-time data is used to update history data and it makes up the new inputs to achieve short-term heat load rolling forecasts. The simulation results show that this method can accurately predict the future heat load, and achieve the purpose of on-demand heating, energy conservation, and environmental protection.

Keywords Heat load forecast · Neural network · Elman · Weather forecast · Sunlight factor

1 Introduction

The district heating systems make up of the heat source, heat supply network, heating stations, and heat users, and its purpose is to continuously provide the heat energy for the users during heating economically and reliably, and meet quality standards.

Q. Li · S. Jiang (✉)

College of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou, China

e-mail: 512314424@qq.com

Q. Li

e-mail: richey@imust.cn

X. Wu

Haiwan Specialty Chemical Co., Ltd, Qingdao, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_49

At present, the control technology of the district heating system in our country is still imperfect, there exist problems of oversupply at high temperature and short supply at low temperature in the users side, only after obtaining a relatively accurate heat load forecast can we adjust the flow and temperature of the water supply scientifically and reasonably, for purpose of achieving the balance between supply and demand, as well as energy conservation and environmental protection. To this problem, many scholars have put forward a mass of methods of the heat load forecast based on different types of the neural network [1–3]. There also have many scholars combined with meteorological factors based on Elman neural network in short-term power forecasting [4, 5], but in the aspect of heat load forecast is limited. The heat load is inseparable from the meteorological factors, and in order to ensure on-demand heating while achieving energy conservation and environmental protection, this study puts forward a heat load forecasting model based on Elman neural network combined with illumination intensity and air temperature. It has been verified by experiment that this method can achieve the expected effect.

2 Heat Load Forecasting Based on Elman Neural Network

2.1 Structure of Elman Neural Network

Elman network is a kind of typical local regression network, which can be understood as a forward neural network with the partial memory unit and local feedback connection. Its structure is similar to the multilayer forward network, which is generally divided into four layers. Figure 1 shows the heat load forecasting model of Elman neural network.

Specifically, the first layer is the input layer which plays the role of signal transmission; the second layer is a middle layer (also known as the hidden layer) that can use linear or nonlinear functions as the transfer function; the third layer is called

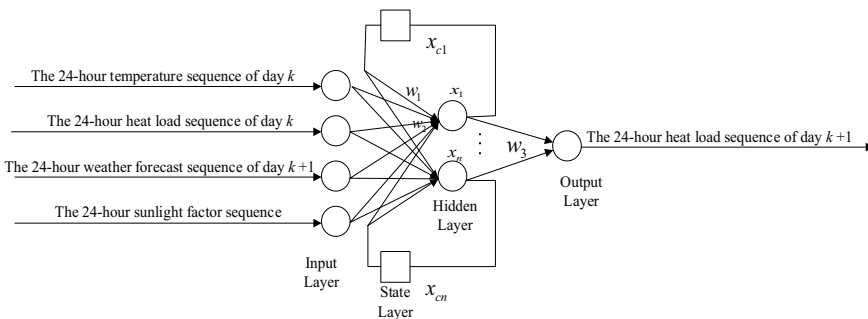


Fig. 1 The heat load forecasting model of Elman neural network

state layer (also called the context layer or the succession layer), which can remember the output value of the middle layer unit at the previous moment and return it to the input of the network; and importantly, the connection weight of the state layer is not changed during the learning process; the fourth layer called the output layer which can be understood as a linear weighting one-step delay operator. The biggest characteristic of the Elman network structure is that the output of the middle layer can be self-linked to the input of the middle layer through the delay and storage of the state layer [6].

2.2 Learning Process of Elman Neural Network

The nonlinear state space expression of Elman network is

$$y(k) = g(w^3 x(k)) \quad (1)$$

$$x(k) = f(w^1 x_c(k) + w^2(u(k-1))) \quad (2)$$

$$x_c(k) = x(k-1) \quad (3)$$

where y is one-dimensional output node vector, x is n -dimensional hidden layer node vector, u is four-dimensional input vector, x_c is n -dimensional feedback state vector, w^3 is the weight between hidden layer and output layer, w^2 is the weight between input layer and hidden layer, w^1 is the weight between state layer and hidden layer, and $g(*)$, and $f(*)$ are transfer function.

In the algorithm of weight correction, Elman neural network adopts BP algorithm, and learning index function uses the sum of squared error function.

$$\sum_{k=1}^n (y_k(w) - \tilde{y}_k(w))^2 \quad (4)$$

where $\tilde{y}_k(w)$ is the target input vector.

3 Model Establishment

According to the historical data of heat load, the model is determined as four inputs and one output system. The input node is the temperature, heat load, sunshine factor coefficient of day k , and the weather forecasting value sequence of day $k+1$. The output node is the heat load forecasting value sequence of day $k+1$. The inherent law of heat load of central heating system, temperature, and sunshine factors is found

out by the model itself, so as to achieve the purpose of forecasting the next day’s heat load.

3.1 Selection of Experimental Data

The experimental data comes from a heating company in Baotou City in China and China Weather Network to ensure the accuracy of the data. In this experiment, the actual data of the heating source were selected for 8 days in December 2014 from Baotou Aluminum Power Plant, and the heating area of the source was 1.12 million square meters. Measured data is on an hourly basis, so there are 24 groups per day with a total of 576 data. The heat load sample data are shown in Table 1. We use the data of the first 7 days as the training sample of the network. Specifically, the temperature, heat load, sunlight factor coefficient of day k and the weather forecast value of day $k + 1$ are used as the input vector for the system, and the heat load forecast value of day $k + 1$ as the target vector. The data of the eighth day is used as a verification sample of the network to test whether the network can reasonably predict the load data of day $k + 1$.

This experiment employs MATLAB invoking the API interface of Chinese WeatherNet to obtain the corresponding data, namely the temperature forecast for the next 24 h. The specific method is to use `urlread` function in MATLAB to read the web page, and then use `regexp` regular function to extract specific string that point in time and temperature. Next, extract the data into the database and make calls in the program.

Selecting the MATLAB programming as follows:

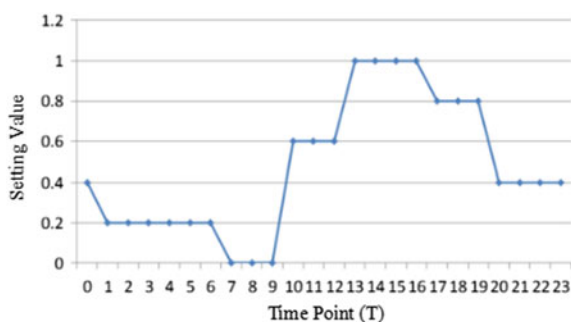
```
S = urlread('http://m.weather.com.cn/mhours/101080201.shtml','get','','GBK');
num = regexp(S, expr1, 'end');
```

In order to ensure the accuracy and real-time performance of the heat load forecasting, the training sample sequence needs to be updated by rolling. The daily updated

Table 1 Sample data of heat load

Sampling time	Real-time temperature (°C)	Sunlight factor	Heat load (GJ)	Weather forecast (°C)
0:00	−12	0.4	153.6133	−11
1:00	−12	0.2	148.4325	−14
2:00	−13	0.2	153.7009	−14
...	...	...	...	...
21:00	−10	0.4	168.6741	−8
22:00	−11	0.4	168.8279	−6
23:00	−11	0.4	168.3252	−8

Fig. 2 Setting value of sunlight factor



data is added to the end of the training sample sequence, and the head-end data is eliminated, so as to update the training sample sequence to achieve the real-time rolling prediction.

3.2 Setting of Sunshine Factors

In an actual central heating system, the heat load value at each time depends not only on the temperature at that time, but also on the environmental impact. In the course of the day, the greatest effect of heat load is the intensity of sunlight in addition to the weather forecast. At the same temperature, the higher the illumination intensity is, the higher the body temperature is, and the heat supply should be reduced appropriately. Similarly, in the absence of sunshine, the body temperature will be lower than the actual temperature, and the heat supply should be increased appropriately. The sunshine factor is set by the above theory, and the set value is shown in Fig. 2.

4 Algorithm Implementation

Create Elman network call format:

```
net = newelm(threshold, [2, 24], {'tansig', 'purelin'});
```

where threshold is the set value of the maximum value and minimum value of the input data. Because the order of magnitude of the input data varies greatly, the input data is all normalized in order to avoid the network prediction error, so the threshold is in the interval [0, 1].

The setting of the training parameters is calculated roughly according to the formula and experience, and after many experiments, the related parameters are set as follows: the training times are set to 2000 (it can train the network well and not take too long); the learning rate is set to 0.1 (the learning effect is better); the number

of hidden layer is set to (it cannot be over-fitted nor can it achieve better prediction effect); and the training target is set to 0.000001. The network training takes 46 s.

The procedure of short-term heat load forecasting algorithm is as follows:

- (1) The 24-hour heating load measured from the heat source at the scene and the corresponding 24-hour temperature obtained through the API interface of the weather station are preprocessed, and the training sample sequence of the heat load forecasting model was formed by adding the sunlight factors.
- (2) We use the newlm function provided by the toolbox in MATLAB and train the samples to create an Elman neural network.
- (3) Input the model parameters and get the 24-hour heat load forecasting value sequence for the next day.
- (4) Add the new measurement value obtained every day to the training sample sequence, eliminate the data of the earliest day, update the training sample sequence, and realize real-time rolling prediction.

Figure 3 shows the flowchart of heat load forecasting algorithm in MATLAB.

5 Analysis of Forecast Results

5.1 *Comparison of Prediction Results of Multiple Neural Networks*

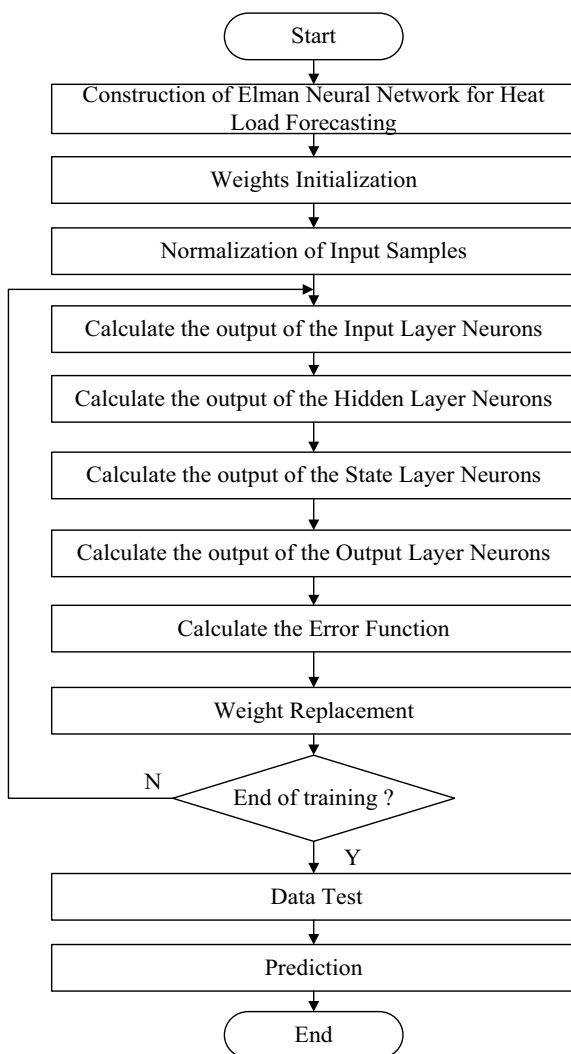
In order to find out the most reasonable and accurate heat load forecasting method, the model of heat load forecasting was established by using BP neural network, GABP neural network, RBF neural network, and Elman neural network, respectively, in this experiment. The experiment used the same sample data with the structure of three inputs and one output. The forecast results are shown in Figs. 4, 5, 6, 7, 8, 9, and 10.

It can be obtained from the result diagrams that the prediction result curve is consistent with the actual heat load change curve, and it fluctuates greatly around sunrise and when the illumination intensity is the highest. However, Elman neural network achieved a relatively good prediction result in the short-term heat load forecasting, and the maximum absolute error is less than 3GJ, as well as the relative error is calculated to be no more than 1.5%. Therefore, Elman neural network was selected as the heat load forecasting model for this experiment.

5.2 *Forecast Result of Adding Sunshine Factor into Elman Network*

In the selected Elman neural network, the forecast results are shown in Figs. 11, 12 and 13 by adding the sunlight factor set in 4.3. Specifically, Fig. 11 shows the

Fig. 3 The flowchart of heat load forecasting algorithm



prediction result, Fig. 12 shows the absolute prediction error, and Fig. 13 shows the relative prediction error.

From Figs. 11 and 12, it can be seen that the predicted value is slightly higher than the expected value before sunrise and lower than the expected output in the afternoon, which meets our experimental requirements. From Fig. 13, it can be seen that the error is still relatively small, and the relative error is only 2% at the maximum.

Fig. 4 Heat load forecast output figure of BP network

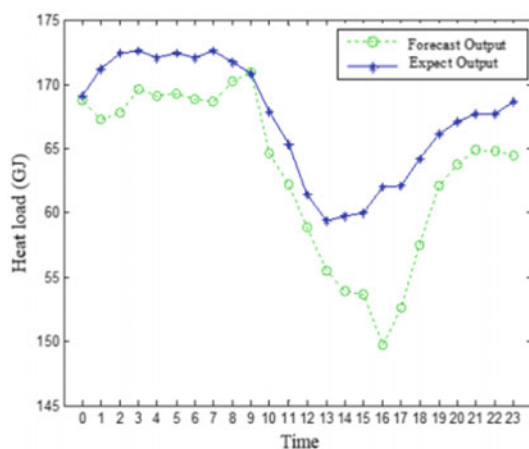


Fig. 5 Heat load forecast output figure of GABP network

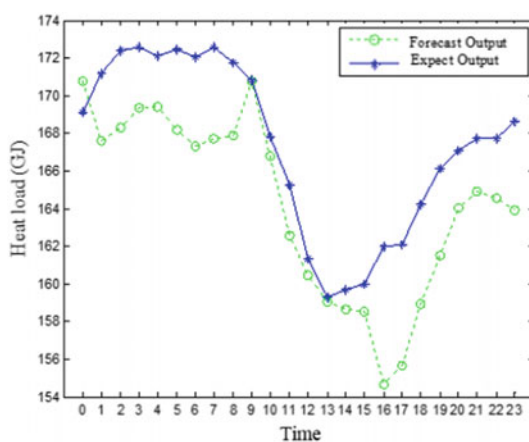


Fig. 6 Heat load forecast absolute error figure of BP network and GABP network

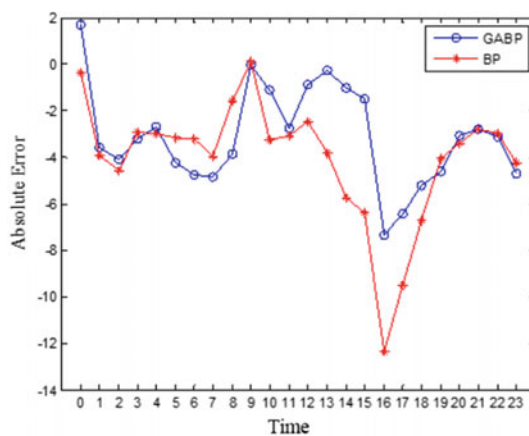


Fig. 7 Heat load forecast output figure of RBF network

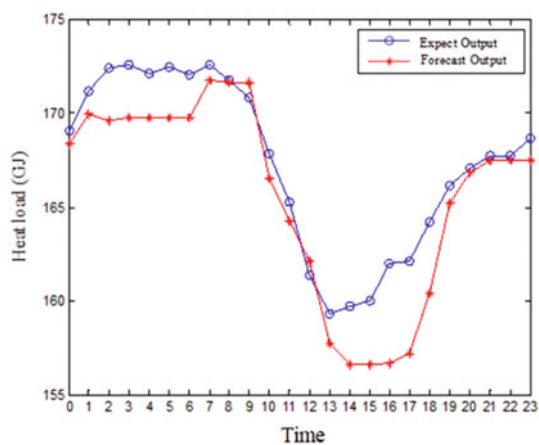


Fig. 8 Heat load forecast error figure of RBF network

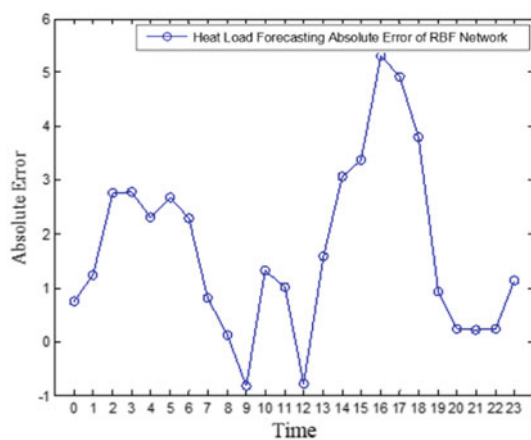


Fig. 9 Heat load forecast output figure of Elman network

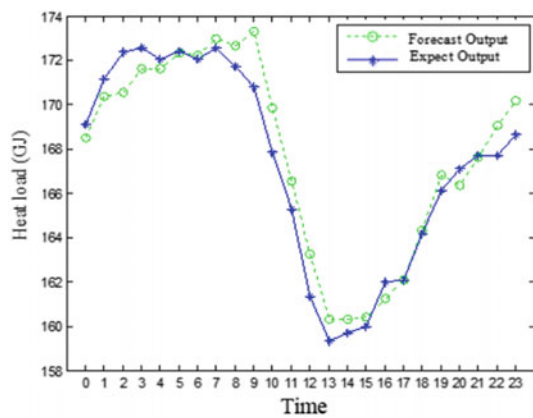


Fig. 10 Heat load forecast error figure of Elman network

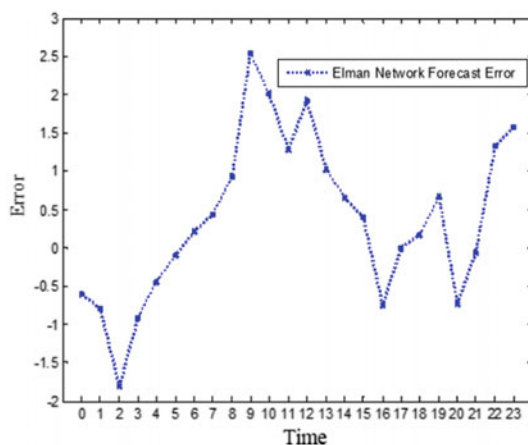
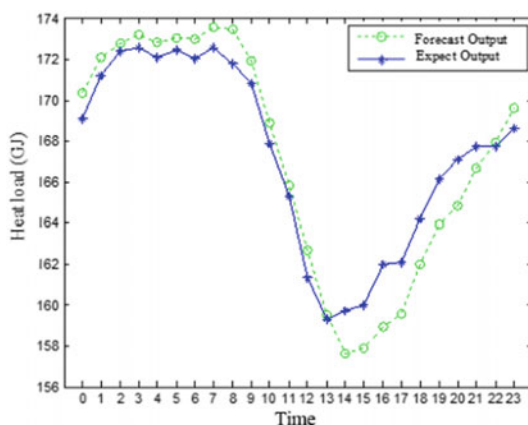


Fig. 11 Heat load forecast output figure of Elman network with the sunshine factor



6 Conclusions

The experimental results show that the heat load forecasting model based on Elman neural network with the addition of temperature and sunlight factors can achieve the effect of saving energy while satisfying the heat load demand of users. The heat load forecasting method provides practical heat load forecasting data for the subsequent district heating control system, which can bring about considerable economic benefits for the urban central heating system. At the same time, it also improves the comfort of heat users, reduces the waste of energy, and lays a good foundation for the subsequent optimization of heat production.

Fig. 12 Heat load forecast absolute error figure of Elman network with the sunshine factor

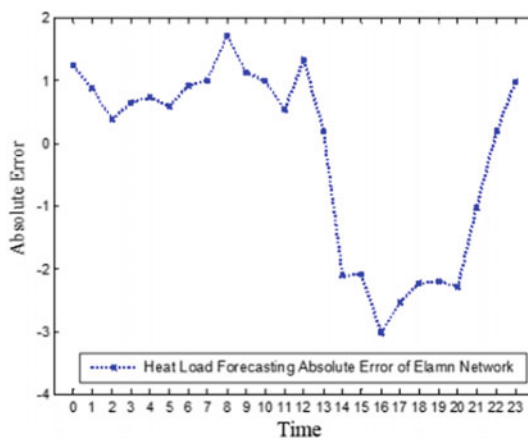
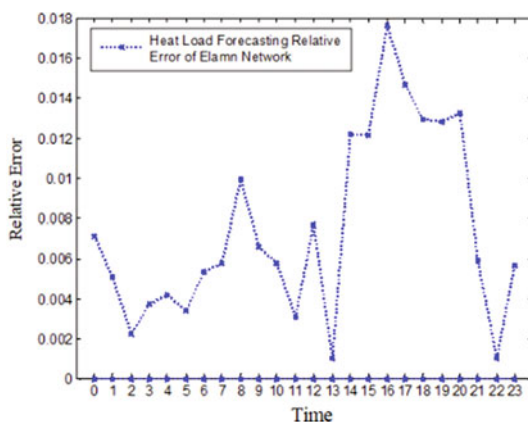


Fig. 13 Heat load forecast relative error figure of Elman network with the sunshine factor



Acknowledgements The work was supported by the National Natural Science Foundation of China (61463040).

References

1. Sakawa, M., Kato, K., Ushiro, S.: Cooling load prediction in a district heating and cooling system through simplified robust filter and multilayered neural network. *Appl. Artif. Intel.* **15**(7), 633–643 (2001)
2. Yang, W.D., Wang, J.Z., Wang, R.: Research and application of a novel hybrid model based on data selection and artificial intelligence algorithm for short term load forecasting. *Entropy* **19**(2), 52 (2017)
3. Suryanarayana, G., Lago, J., Geysen, D., Aleksiejuk, P., Johansson, C.: Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. *Energy* **157**, 141–149 (2018)

4. Dou, C.X., Qi, H., Luo, W., Zhang, Y.M.: Elman neural network based short-term photovoltaic power forecasting using association rules and kernel principal component analysis. *J. Renew. Sustain. Energy* **10** (2018)
5. Zhang, H., Tao, H.Q.: Power load forecasting based on Elman neural network. In: 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, vol. Iv, pp. 374–376 (2009)
6. Cheng, Y.C., Qi, W.M., Cai, W.Y.: Dynamic properties of Elman and modified Elman neural network. In: 2002 International Conference on Machine Learning and Cybernetics, Proceedings, vols. 1–4, pp. 637–640 (2002)

Theoretical Study on Even Heating of Single Pipe Heating System



Xiaoli Yin, Mingsheng Liu, Zhixian Ma and Jili Zhang

Abstract In order to solve the problem of thermodynamic imbalance commonly existing in the heating system, an adjustment method that can realize the uniform heating to a building is put forward. In the single pipe system, it regulates the flow based on the energy balance between the heat of system supply for users and the heat dissipation of radiator to indoor, reversing the direction of the supply water and return water. By means of theoretical analysis, the results indicate that this method can solve the problem of building overheating, and the energy saving rate is higher than 30%. This method can also solve the problem of uneven heat and cold of the building by adjusting the ratio of positive and reverse times.

Keywords Heating system · Energy balance · Reverse direction · Overheating · Uneven cold and heat

Notations

Q_e	the heat supply for users, W
Q_E	the heat supply for users at the set room temperature, W
Q_1-Q_7	the thermal load from the first to the seventh floor, W
N_1-N_2	the number of radiators from the first to the seventh floor, W
Q_r	the heat dissipation of a single chip radiator, W
A_r	the area of dissipation of a single chip radiator, W
$t_{o,d}$	the outdoor calculation temperature of heating, °C
$t_{i,d}$	the room design temperature of heating, °C
Q_n	the heat dissipation of the radiator to the room, W
Q_N	the heat dissipation of the radiator to the room at the set temperature, W

X. Yin · M. Liu · J. Zhang

Institute of Building Energy, Dalian University of Technology, Dalian, China

Z. Ma (✉)

Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian, China

e-mail: mazhixian@dlut.edu.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,

Advances in Intelligent Systems and Computing 890,

https://doi.org/10.1007/978-981-13-6733-5_50

Δt	the average water temperature in the radiator, °C
ρ	the density of water, kg/m ³
c	the specific heat capacity of water, J/(kg °C)
G	the volume flow of water, m ³ /h
T_1	the inlet water temperature, °C
T_2	the outlet water temperature, °C
T_n	the room temperature, °C
T_N	the set room temperature, °C
K	the heat transfer coefficient of radiator, W/(m ² °C)
A	the heat exchanger area, m ²
B	the energy regulation coefficient, W/°C
a	the time of up-supply and down-circle
b	the time of down-supply and up-circle

1 Introduction

The hydraulic and thermal conditions of the central heating system are very complicated. In order to realize heating as expectation and energy saving, the central heating system should not only guarantee the hydraulic balance of the heating system, but also guarantee the thermal balance of the heating system. In the viewpoint of the heat imbalance of the central heating, many people have studied the thermodynamic modeling of the central heating systems [1–3], the operation adjustment [4, 5], and the solution of the hydraulic imbalance [6–8]. The energy balance method came up with [9], but the problem of uneven heat and cold from bottom to the top of the building remained unsolved. This paper briefly discusses the causes of uneven heat and cold in single pipe system buildings. On this basis, a commutation model is proposed for a specific building. Finally, the theoretical calculation results are given.

2 Heating System Model

2.1 Physical Model

The building adopts radiator heating systems, whose form is a single pipe heating system of up-supply and down-circle. The design condition: the inlet water temperature is 80 °C and the backwater temperature is 60 °C. A steel cylindrical radiator with the model of GZ-3-10-1.0 is selected. The heat radiation area of the radiator is 0.281 m² per piece, and the heat dissipation of the monolithic radiator is

$$Q_r = 0.47 \Delta t^{1.39} \quad (1)$$

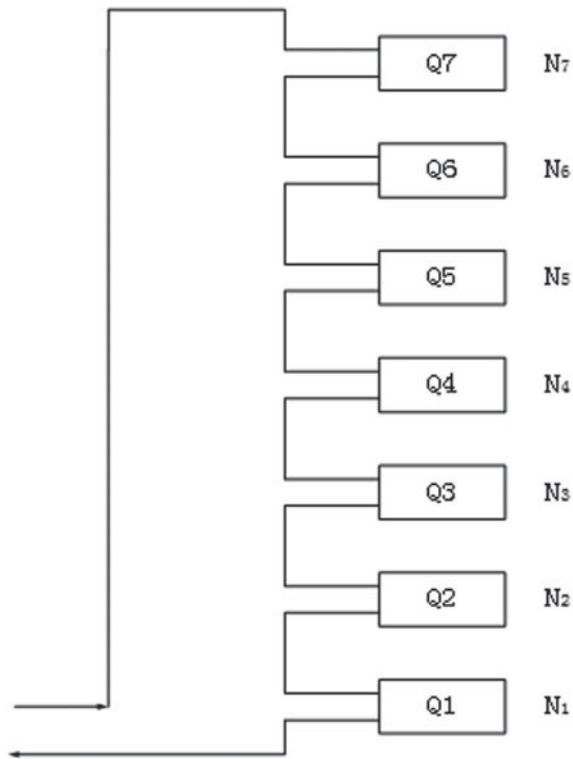


Fig. 1 Schematic diagram of the single pipe heating system

Table 1 Heat load of each floor

Floor	Heat load (W)
Seventh floor	2098.53
Sixth floor	1867.43
Fifth floor	1867.43
Fourth floor	1867.43
Third floor	1867.43
Second floor	1867.43
First floor	2721.41

The schematic diagram of the single pipe heating system is shown in Fig. 1. In the model, all radiators are connected in series with the same flow rate but different inlet temperatures. The number of radiators varies from room to room, matching the heat load. We choose a building in Beijing, $t_{o,d}$ is 18 °C and $t_{i,d}$ is −9 °C, Table 1 shows the heat load of each floor.

2.2 Mathematical Model

The cycle of the heating system has first arrived at the radiator through the heat-supply pipeline, and the heat transferred is

$$Q_e = \rho c G(T_1 - T_2). \quad (2)$$

The heat transfer of radiator to the room is

$$Q_n = \frac{KA}{2}(T_1 + T_2 - 2T_n). \quad (3)$$

Ignoring the heat dissipation into the surrounding environment, Q_e is equal to Q_n .

$$\rho c G(T_1 - T_2) = \frac{KA}{2}(T_1 + T_2 - 2T_n). \quad (4)$$

Let

$$B = \frac{KA}{2}. \quad (5)$$

In the control process, T_N is fixed, the heat transfer from radiators to the room is

$$Q_N = B(T_1 + T_2 - 2T_N) \quad (6)$$

and the heat-supply system needs to provide the heat $Q_E = Q_N$. In the actual process of the system operation, the heat provided by the system is $Q_e = \rho c G(T_1 - T_2)$. If the two parameters are not equal, the actual temperature of the room is not equal to the set temperature.

If $Q_e < Q_N$, that is,

$$\rho c G(T_1 - T_2) < B(T_1 + T_2 - 2T_N)$$

and

$$B(T_1 + T_2 - 2T_n) < B(T_1 + T_2 - 2T_N), \quad T_n > T_N.$$

In summary:

$$Q_e < Q_N, \quad T_n > T_N; \quad Q_e > Q_N, \quad T_n < T_N; \quad Q_e = Q_N, \quad T_n = T_N.$$

When the energy balance method is adopted to control the heating system, there is no need to monitor the room temperature directly. As long as we find the rela-

tion between the heat supplied by the system to the user and the heat dissipation of the radiator to the room at the standard set temperature, we can know that the indoor temperature is higher or lower than the set temperature, and then make some adjustments.

The key is how to acquire the value of Q_N . In our calculation, all radiators for the building are regarded as a whole.

$$Q_r = \alpha \Delta t^\beta \quad (7)$$

$$Q_N = \frac{A}{A_r} \alpha \Delta t^\beta. \quad (8)$$

Extensive test on the actual building heating system can fit the function relation between Q_N and Δt . Based on the method of undetermined coefficients, figure out α and β .

In this paper, α and β separately take 0.47 and 1.39, and the heat dissipation area is 45.7 m².

The backwater temperature can be obtained from (4).

$$T_2 = \frac{2\rho cG - KA}{2\rho cG + KA} T_1 + \frac{2KA}{2\rho cG + KA} T_n. \quad (9)$$

In the reverse process, the supply and return water temperatures of each floor can be obtained by (9), changing the flow until the total heat supply is equal to the demand heat supply, and then the heat supply in each floor is calculated.

3 Theoretical Analysis and Result Discussion

In the single pipe system of up-supply and down-circle, the inlet water temperature of the seventh floor and the room set temperature of each floor are kept at 18 °C and the flow rate is adjusted. In practice, the heat demand of the most unfavorable heat users should be met, that is, the room temperature of the first floor can reach 18 °C. According to the heating load of each floor, the outdoor calculation temperature, and the calculation equation of room load, the room temperature of the other floors is calculated.

It is shown in Table 2, the results of up-supply and down-circle systems.

Under the design conditions, the hot water flow is 0.45 m³/h. From the above table, it can be seen that in the actual situation, the thermodynamic imbalance is serious in the building because of increasing the area of radiators in each floor.

For the traditional up-supply and down-circle system, the average temperature of the building is 21.9 °C, which is higher than the design temperature. Figure 2 shows the unbalance rate.

Table 2 Results of up-supply and down-circle systems

Floor	Water supply temperature (°C)	Backwater temperature (°C)	Temperature difference (°C)	Heating load (W)	Room temperature (°C)
Seventh	80.00	74.86	5.14	2679.17	25.48
Sixth	74.86	70.46	4.40	2299.17	24.25
Fifth	70.46	66.21	4.25	2218.47	23.07
Fourth	66.21	62.11	4.10	2136.85	21.89
Third	62.11	58.17	3.94	2054.25	20.70
Second	58.17	54.40	3.78	1970.61	19.49
First	54.40	49.18	5.22	2721.51	18.00

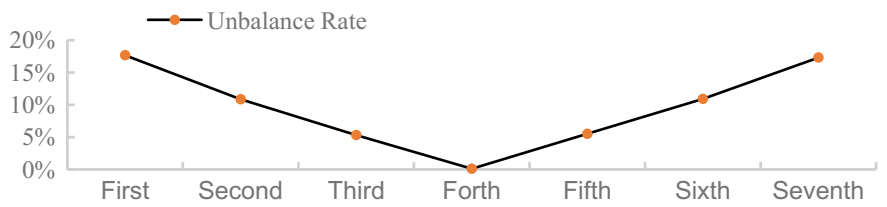


Fig. 2 The unbalance rate of each floor

Table 3 Results of down-supply and up-circle systems

Floor	Water supply temperature (°C)	Backwater temperature (°C)	Temperature difference (°C)	Heating load (W)	Room temperature (°C)
Seventh	65.45	63.66	1.79	2098.57	18.00
Sixth	67.18	65.45	1.74	2032.29	20.38
Fifth	69.07	67.18	1.89	2209.45	22.94
Fourth	71.13	69.07	2.06	2411.94	25.87
Third	73.39	71.13	2.26	2644.95	29.24
Second	75.88	73.39	2.49	2915.08	33.15
First	80.00	75.88	4.12	4827.92	38.89

Figure 2 shows that the unbalance rate is exactly symmetrical. And, there is the lowest rate of imbalance in the middle layer.

In the single pipe system of down-supply and up-circle, we keep the inlet water temperature of the first floor and the room set temperature of each floor constant, and adjust the flow rate, and keep the room temperature of the seventh floor 18 °C, and calculate the room temperature of the other floors according to Q_1-Q_7 , $t_{o,d}$ and $t_{i,d}$.

Table 3 shows the results of down-supply and up-circle systems.

Under this design condition, the hot water flow is 1.0 m³/h. From the above table, it can be seen that the uneven heat and cold are more serious.

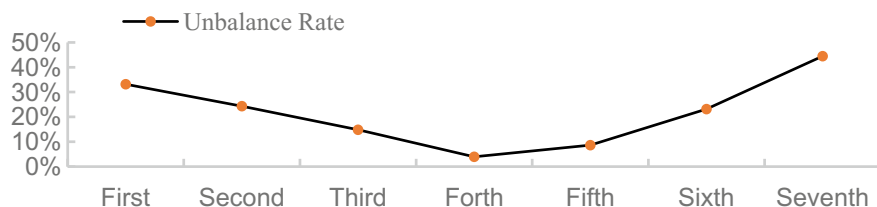


Fig. 3 The unbalance rate of each floor

Table 4 Results of up-supply and down-circle systems

Floor	Water supply temperature (°C)	Backwater temperature (°C)	Temperature difference (°C)	Heating load (W)	Multiple load
Seventh	80	72.78	7.22	2632.11	1.25
Sixth	72.78	66.80	5.98	2178.32	1.17
Fifth	66.80	61.23	5.57	2028.25	1.09
Fourth	61.23	56.07	5.16	1881.92	1.01
Third	56.07	51.30	4.77	1739.40	0.93
Second	51.30	46.90	4.4	1600.80	0.86
First	46.90	41.15	5.75	2096.40	0.77

For the down-supply and up-circle system, the average temperature of the building is 26.9 °C, which is higher than design temperature. Figure 3 shows the unbalance rate.

Figure 3 shows that the unbalance rate is big. And still, there is the lowest rate of imbalance in the middle layer.

Using the energy balance method, Eq. (8) and the average temperature of building is set as 18 °C. The final outlet water temperature is 41.58 °C and the flow is 0.315 m³/h. Assuming that the room temperature of each floor can reach 18 °C, the heat transfer coefficient is calculated on design conditions, and Eq. (9) is used to calculate the water supply temperature from the first to the sixth floor.

Table 4 shows the results of up-supply and down-circle systems.

The flow rate is 0.312 m³/h, reverse the flow and maintain the same flow rate.

Table 5 shows the results of down-supply and up-circle systems.

The heating load of each floor is calculated, we aim to keep the heat supply in each floor be equal to the load, changing the time ratio of the two kinds of system.

Take the time ratio of seventh floor as an example, the calculation equation is given as follows:

$$\frac{a}{b} = \frac{2098.53 - 1112.41}{2632.11 - 2098.53} \quad (10)$$

Table 5 Results of down-supply and up-circle systems

Floor	Water supply temperature (°C)	Backwater temperature (°C)	Temperature difference (°C)	Heating load (W)	Multiple of load
Seventh	44.20	41.15	3.05	1112.41	0.53
Sixth	47.41	44.20	3.21	1169.65	0.63
Fifth	51.20	47.41	3.79	1379.87	0.74
Fourth	55.70	51.20	4.50	1641.18	0.88
Third	61.11	55.70	5.41	1969.62	1.05
Second	67.66	61.11	6.55	2387.52	1.28
First	80	67.66	12.34	4496.95	1.65

Table 6 Time ratio forward and reverse

Heating load for up-supply and down-circle systems (W)	Heating load for down-supply and up-circle systems (W)	a/b
2632.10	1112.38	1.85
2178.31	1169.63	2.24
2028.24	1379.84	3.03
1881.90	1641.15	15.64
1739.38	1969.60	0.80
1600.77	2387.50	1.95
2096.34	4496.94	2.84

Because the heat supply of the fourth floor is almost equal to the heat load, that is, the time ratio of is approximated to an arbitrary value, so the time ratio of the fourth floor can be removed when the mean value is obtained, so the time ratio of reverse the flow is 2.36 (Table 6).

4 Conclusions

Using the method of energy balance and reversing the flow, the rate of flow decrease is more than 31% for up-supply and down-circle systems, the energy saving is 13%. For the down-supply and up-circle systems, the flow is reduced by about 70%, and the energy saving rate is about 26%. Furthermore, the thermal load is designed according to the most unfavorable conditions, because of the internal heat source and solar radiation, it cannot reach the design load in most cases, so the energy saving ratio of the energy balance method is higher than the calculated. Based on the balance of energy, reversing the flow is used to make the energy distribution adapt to the thermal load and make the heating uniform.

References

1. Wang, J.D.: Comparative analysis of pipe flow-resistance characteristic coefficient identification methods in heating network. Science and Engineering Research Center. In: Proceedings of 2015 International Conference on Computer Science and Environmental Engineering (CSEE 2015), pp. 7. Science and Engineering Research Center, Beijing (2015)
2. Lin, F.: Study on peak load operation mode of thermoelectric (cooling) cogeneration system. Tsinghua University, Beijing (1999). (in Chinese)
3. Palsson, O.P., Madsen, H., Sogaard, H.T.: Predictor-based optimal control of supply temperature in district heating systems. *Control Eng. Pract.* (1), 81–85 (1993)
4. Mijakovski, V.: Possible ways of regulation for branched heating systems. *Appl. Thermal Eng.* **29**(11), 2579–2582 (2008)
5. Rauh, A., Aschemann, H.: Parameter identification and observer-based control for distributed heating systems-the basis for temperature control of solid oxide fuel cell stacks. *Math. Comput. Model. Dyn. Syst.* **18**(4), 329–353 (2012)
6. George, Jr., R.C.: The history of primary/secondary pumping system [OL]. FHRA Netw. (2003)
7. You, T., Shi, W., et al.: A new ground-coupled heat pump system integrated with a multi-mode air-source heat compensator to eliminate thermal imbalance in cold regions. *Energy Build.* **107** (2015)
8. Linuan, L.: Single tube heating system and its application research (Dissertation for Master's Degree). Taiyuan University of Technology, vol. 9 (2012). (in Chinese)
9. Lei, Y.: Energy balance method and its application in floor radiant heating control system (Dissertation for Master's Degree). Harbin Institute of Technology, vol. 3 (2012). (in Chinese)

Illumination Variation Similarity Based Fault Diagnosis for HV-LED Lamp Driven by Segmented Linear Driver



Fukang Sun, Shaofeng Zhu and Ye Wang

Abstract In this paper, fault diagnosis methods based on illumination variation similarity analysis for High-voltage Light Emitting Diode (HV-LED) lamp driven by segmented linear driver are proposed. The proposed methods assess the illumination variation similarity between the diagnosed lamp and the different fault-type lamp to diagnose whether the diagnosed lamp occurs fault or not and confirm the fault type. Euclidean distance is applied for calculating illumination variation similarity. The proposed fault diagnosis methods contain four parts: illumination signal smoothing, similarity calculation, similarity assessment, and fault recognition. The illumination variation data of the fault-free and three different faulty lamps, which are based on one HV-LED lamp driven by four-segment linear driver, are investigated for method verification experiments. The experimental results and analysis are given to demonstrate the validity and effectiveness of the proposed fault diagnosis methods in test chamber environment.

Keywords HV-LED lamp · Segmented linear driver · Fault diagnosis method · Illumination variation characteristics · Similarity analysis · Euclidean distance

F. Sun (✉)

Anhui Province Key Laboratory of Intelligent Building and Building Energy-Saving, Anhui Jianzhu University, Hefei, China
e-mail: fukang0731@126.com

F. Sun · S. Zhu · Y. Wang

Engineering Research Center for Building Energy Efficiency Control and Evaluation, Ministry of Education, Hefei, China

F. Sun · S. Zhu · Y. Wang

Anhui Engineering Technology Research Center for Building Energy-Saving, Hefei, China

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*, Advances in Intelligent Systems and Computing 890, https://doi.org/10.1007/978-981-13-6733-5_51

1 Introduction

Compared with the conventional light source, LED lamp has the advantages of higher Color Rendering Index (CRI), higher light efficiency, smaller size, longer service life, and more eco-friendly [1]. Segment linear solution is a novel technology that is applied for driving HV-LED chips [2–5]. HV-LED lamp driven by segment linear solutions is composed of only rectifier bridge and segmented linear driver and several HV-LED strings without the use of passive components [6–9]. Hence, compared with other LED drive schemes, the HV-LED lamp based on segment linear driver has better performance, such as higher power factor, lower total harmonic distortion, higher reliability, and longer service life [10, 11]. However, few literatures concern on the fault diagnosis methods of HV-LED lamp based on segment linear solution.

The damage of the HV-LED chips is the main failure of the HV-LED lamp driven by segment linear solution, and can directly cause the illumination value of the lamp to drop and even not to light. Nevertheless, the damage of few HV-LED chips may not lead to a big drop in illumination, and it is difficult to identify whether the HV-LED lamp occurs fault or not, especially in the applications of outdoor high-power street lamp and high-power lamp in the factories. Therefore, in this paper, the fault diagnosis methods based on illumination variation similarity analysis for HV-LED lamp driven by segmented linear driver are proposed. The core idea of the proposed method is to judge whether the diagnosed lamp is in trouble by calculating the illumination variation similarity between the diagnosed lamp and the faulty lamps. Moreover, in the research field of circuit fault detection, the method of analyzing the similarity between the diagnosed samples and the fault samples has been widely applied [12, 13].

The remaining parts of the paper are organized as follows: Sect. 2 presents the circuit principle and illumination variation characteristics and fault type of HV-LED lamp driven by segmented linear driver. Section 3 describes the proposed fault diagnosis methods. In Sect. 4, an illustrative case of the HV-LED lamp driven by four-segment linear driver is used to demonstrate the effectiveness and practicability of the proposed fault diagnosis methods in Sect. 3. Conclusions are drawn in Sect. 5.

2 Segmented Linear Solution for HV-LED Lamp

Figure 1 illustrates the circuit structure of HV-LED lamp based on segmented linear solution. The circuit includes several HV-LED sub-strings ($Str_1, Str_2, \dots, Str_n$). According to the corresponding threshold voltage point ($V_1, V_2, \dots, V_n$), the controller generates the control signals to the switch component ($S_1, S_2, \dots, S_n$). When the voltage of AC source rises, HV-LED sub-strings are lightened sequentially. And, when the voltage of AC source drops, HV-LED sub-strings are closed sequentially.

Hence, the illumination of HV-LED lamp continuously fluctuates with the sub-strings light and closes sequentially. Figure 2 illustrates the illumination variation of fault-free and faulty HV-LED lamp driven by four-segment linear driver. Figure 2a

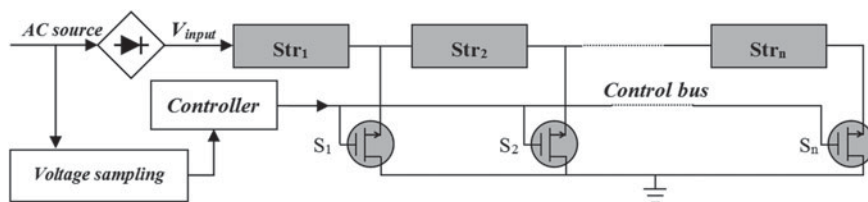


Fig. 1 Schematic diagram of the segmented linear solution

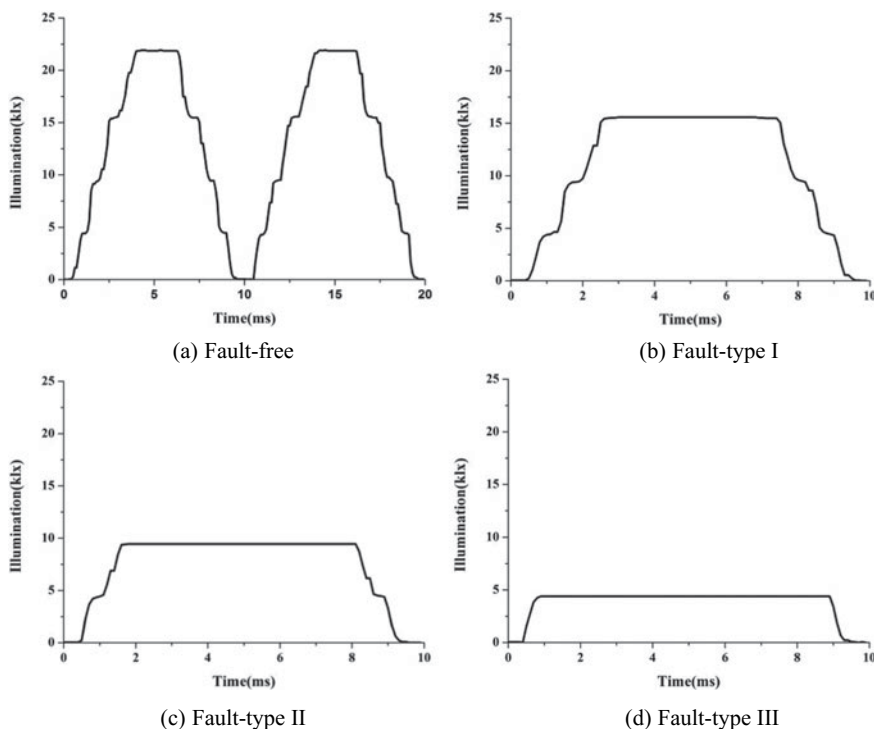


Fig. 2 Diagram of illumination curve of HV-LED lamp driven by four-segment linear driver

illustrates illumination variation of fault-free HV-LED lamp during two work period. There are four segments in illumination curve. As shown in Fig. 2b, in the condition of fault-type I, Str_4 cannot be lightened, and illumination of HV-LED lamp drops, and illumination curve lacks one segment. As shown in Fig. 2c, in the condition of fault-type II, Str_3 and Str_4 cannot be lightened, and illumination curve lacks two segments. As shown in Fig. 2d, in the condition of fault-type III, Str_2 and Str_3 and Str_4 cannot be lightened, and illumination curve lacks three segments. Obviously, there is significant illumination difference between normal lamp and faulty lamp.

3 Proposed Methods

In this paper, as shown in Fig. 3, there are four steps in the procedure of the proposed fault diagnosis methods: (1) illumination signal smoothing, (2) similarity calculation, (3) similarity assessment, and (4) fault recognition.

Assume that $X = \{x_1, x_2, \dots, x_n\}$ represents original illumination value sequence of the HV-LED lamp base on segmented linear driver during single work period, which are collected with high-speed illumination sensor.

Step 1: Illumination signal smoothing

First, the average filtering method is applied to reduce noise in the original illumination value collected from the HV-LED lamp, and the size of filtering window is three. The average filtering method is defined as Eqs. (1) and (2).

$$x_i = \frac{x_{i-l} + x_i + x_{i+l}}{2l + 1} \quad (1)$$

$$w = 2l + 1, \quad (2)$$

where w represents the filter window width of average filtering method, and the l equals 1 in this paper.

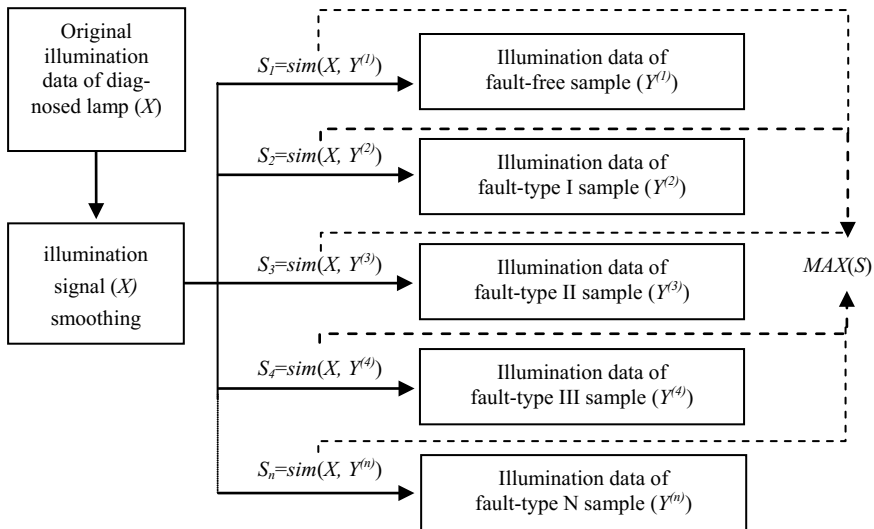


Fig. 3 Flowchart of the proposed fault diagnosis methods

Step 2: Similarity calculation

Next, assume that $S = \{S_1, S_2, \dots, S_n\}$ represents the illumination variation similarity set between the diagnosed lamp and the different fault-type lamps, and the illumination variation similarity calculation is defined in Eqs. (3) and (4).

$$D(X, Y^{(i)}) = \sqrt{(x_1 - y_1^{(i)})^2 + (x_2 - y_2^{(i)})^2 + \dots + (x_n - y_n^{(i)})^2} \quad (3)$$

$$S_i = \text{sim}(X, Y^{(i)}) = \frac{1}{1 + D(X, Y^{(i)})}, \quad (4)$$

where $Y^{(i)}$ represents standard illumination sequence of fault-type i HV-LED lamps. The function $D(X, Y^{(i)})$ is applied to calculate the Euclidean distance between sequence X and sequence $Y^{(i)}$. The $\text{sim}(X, Y^{(i)})$ is applied to calculate the similarity between X and $Y^{(i)}$.

Step 3: Similarity assessment

Then, the $\text{MAX}(S)$ is used to calculate the maximum m of illumination variation similarity set, and defined in Eq. (5).

$$m = \text{MAX}(S) \quad (5)$$

Step 4: Fault recognition

Finally, based on the maximum m calculated in Step 3, the fault-type i can be confirmed with the mapping relation between m and i .

4 Experimental Results

To prove the performance of the proposed fault diagnosis methods, five 8 W HV-LED lamps based on four-segmented linear driver are chosen as the diagnosed samples. Sample 1 is a fault-free lamp. Sample 2 is a fault-type I lamp, and Str_4 cannot be lightened. Sample 3 is a fault-type II lamp, and Str_3 and Str_4 cannot be lightened. Sample 4 is a fault-type III lamp, and Str_2 and Str_3 and Str_4 cannot be lightened. Sample 5 is a fault-type IV lamp, and all sub-strings cannot be lightened. During the experiments, the illumination data sample of the HV-LED lamps was collected with LFA-2000 light flicker analyzer.

Furthermore, standard illumination dataset of various faulty HV-LED lamps driven by four-segmented linear driver is established. The standard dataset contains the illumination data of the normal lamp and the four different types of faulty lamps. Standard illumination datasets are implemented by three steps: (1) prepare fault-free and various faulty lamps, (2) collect illumination samples of various, and (3) illumination data filtering and smoothing. Besides, it should be noted that data collection

Table 1 Experimental results

Sample	$Y^{(1)}$	$Y^{(2)}$	$Y^{(3)}$	$Y^{(4)}$	$Y^{(5)}$	$MAX(S)$	Fault type
1	0.122	0.028	0.013	0.009	0.007	0.122	Fault free
2	0.029	0.245	0.022	0.012	0.008	0.245	I
3	0.013	0.022	0.153	0.023	0.012	0.153	II
4	0.009	0.012	0.023	0.206	0.024	0.206	III
5	0.007	0.008	0.012	0.024	0.762	0.762	IV

methods and the experimental environment of aforementioned two datasets are the same. In particular, there is no natural light and external light in the experimental environment.

The experimental results are shown in Table 1. In Table 1, $Y^{(1)}$ and $Y^{(2)}$ and $Y^{(3)}$ and $Y^{(4)}$ and $Y^{(5)}$ denote the illumination data of the normal lamp and the four faulty lamps, respectively. According to the experimental results, it can be concluded that sample 1 is a fault-free lamp, sample 2 is a fault-type I lamp, sample 3 is a fault-type II lamp, sample 4 is a fault-type III lamp, and sample 5 is a fault-type IV lamp. And then, the experimental results demonstrate the effectiveness of the proposed fault diagnosis methods for fault diagnosis of HV-LED lamp based on segmented linear driver.

5 Conclusions

In the paper, the fault diagnosis methods are proposed for identifying the HV-LED lamp driven by segmented linear driver using illumination variation geometrical similarity. Experimental results demonstrate that the proposed diagnosis methods are effective methods, and can confirm the fault status of the HV-LED lamp. However, the proposed methods cannot deal with the distance effect, and cannot effectively calculate the similarity between the illumination of diagnosed lamp and the standard illumination dataset in different test distances. Hence, in the future, we will continue to study the fault diagnosis methods for the more complicated light environment.

Acknowledgements This work was supported by the Natural Science Research Project of Anhui Education Department (KJ2015JD24 and KJ2016A814 project), the Research Project of Anhui Provincial Housing and Urban Construction Department (2015YF-18 project), and the Research Project of Anhui Jianzhu University (2014XQZ03 project and 2017XQZ01 project).

See Tables 2 and 3.

[illegible]

References

1. Steigerwald, D.A., Bhat, J.C., et al.: Illumination with solid state lighting technology. *IEEE J. Sel. Top. Quantum Electron.* **8**(2), 310–320 (2002)
2. Hwu, K.I., Tu, W.C.: A high brightness light-emitting diode driver with power factor and total harmonic distortion improved. In: 26th IEEE Applied Power Electronics Conference & Exposition, pp. 713–717. IEEE, USA (2011)
3. Jong-hyun, K.I.M., Myung-hyo, R.Y.U., et al.: A new dimmer for alternating-current directly driven light-emitting-diode lamp. *J. Cent. South Univ.* **19**(2), 374–379 (2012)
4. Dayal, R., Modepalli, K., Parsa, L.: A direct AC LED driver with high power factor without the use of passive components. In: 2012 IEEE Energy Conversion Congress and Exposition, pp. 4230–4234. IEEE, USA (2012)
5. Park, C., Rim, C.T.: Filter-free AC direct LED driver with unity power factor and low input current THD using binary segmented switched LED strings and linear current regulator. In: 28th Applied Power Electronics Conference and Exposition, pp. 870–874. IEEE, USA (2013)
6. Ning, N., Chen, W.B., Yu, D.J., et al.: Self-adaptive load technology for multiple-string LED drivers. *Electron. Lett.* **49**(18), 1170–1171 (2013)
7. Seo, K., Nguyen, V.H., Jung, J., et al.: Multi-string AC-powered LED driver with current regulation reduction based on simple circuitry. *IEICE Electron. Express* **11**(19), 1–8 (2014)
8. Li, Y., Guo, W., Zhu, Z.: A high efficiency and power factor segmented linear constant current LED driver. *China J. Semicond.* **36**(4), 165–171 (2015)
9. Liu, C., Lai, X., Hanxiao, D.: Improvements in performance and reliability for segmented linear LED drivers. *China J. Semicond.* **37**(7), 41–47 (2016)
10. Noge, Y.: Linear AC LED driver with the multi-level structure and variable current regulator. In: 9th International Conference on Power Electronics and Ecce Asia, pp. 964–969. IEEE, South Korea (2015)
11. Noge, Y., Fuse, H., Shimizu, T.: Experimental validation of linear AC LED driver with quantitative design method. In: 2017 Applied Power Electronics Conference & Exposition, pp. 1484–1491. IEEE, USA (2017)
12. Huang, H.P., Li, C., Jeng, J.C.: Multiple multiplicative fault diagnosis for dynamic processes via parameter similarity measures. *Ind. Eng. Chem. Res.* **46**(13), 4517–4530 (2007)
13. Liu, C., Jiang, D., Yang, W.: Global geometric similarity scheme for feature selection in fault diagnosis. *Expert Syst. Appl.* **41**(8), 3585–3595 (2014)

Point Illumination Calculation Method in Special-Shaped Space



Jundong Fu, Qing Chen, Yunxia Qiu and Li Chen

Abstract This paper proposes a novel approach to spatial illumination calculation by experiments. Existing researches on illumination calculations have these problems such as large errors in the calculation of optical metrics, and there are great limitations in the spatial form and layout of devices. In order to solve these problems, based on the light distribution curve and the law of Lambertian reflectance, the spatial illumination can be obtained, and the mathematical model of a complex spatial light environment with more accurate calculation results and wider application range was established and implemented by MATLAB. Finally, it is used in the illumination calculation simulation of different spatial models. After calculation, the average illumination intensity calculation of space work surface is compared with that of DIALux results. The error is within $\pm 3\%$, the distribution of plane illumination curves is basically the same, and the solution to the circular bottom surface space, the unequal height of the bottom spatial light environment cannot directly describe the problem.

Keywords Lambert radiation law · Illumination calculation · Discretization of energy

1 Introduction

When designing a space for lighting, it is necessary to evaluate the average illuminance value in the space, the illuminance uniformity, and other indicators. In the engineering lighting design using a wide range of software such as Hao

J. Fu (✉) · Q. Chen · Y. Qiu · L. Chen
School of Electrical and Automation Engineering, East China Jiaotong University,
Nanchang, China
e-mail: 8755915@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_52

Chen Electric, Tianzheng Electric, and others only the spatial illuminance value is estimated, it is difficult to get the rest of the parameter values; professional lighting design software DIALux has a very high authority in the lighting calculation, and its simulation calculation. The error between the measured value and the real environment is about 5% [1], but it has certain limitations when designing irregular spaces. Therefore, designing a universal and accurate point illumination calculation method is of great significance to the development of the lighting design industry.

There are many methods that can be used to calculate the illuminance of space, such as the widely used coefficient of use method and point-by-point calculation method, as well as the high accuracy of nonsequential ray tracing [2] and the establishment of luminous flux transfer function [3, 4]. Wait. In [5], it is proposed that the luminescent light source is approximated by a Lambertian body, and through the diffuse reflection of the hemispherical surface, the illuminance distribution on the detection plane is obtained. Simulation and comparison of this method in [6] verified the reliability of the method. Based on this, in the literature [7], when designing the reflector, the concept of energy discretization was proposed, the reflected light was quantified, and the reflected illuminance of the detection surface was obtained.

Combined with the above studies, the concept of energy discretization was extended from the design of reflectors to the calculation of indoor illumination. This paper studies the description of the complex light environment above the most basic light metric theory in [8] and [9]. And apply it to the field of lighting design that describes the illumination distribution of indoor space, which is mainly the measurement of spatial illumination.

2 Problem Description

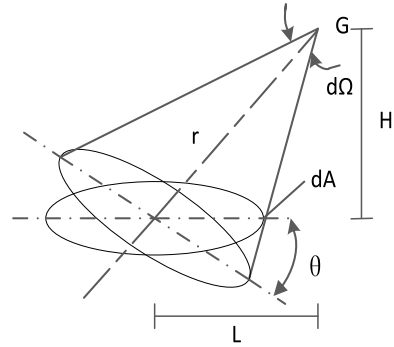
The total illumination at a certain point i in the room can be expressed as $E_i = E_{di} + E_{ri}$, the direct illumination E_{di} which is calculated by invoking the light intensity parameter of the light distribution curve, the reflected illuminance E_{ri} is obtained according to the Lambert radiation law, and then the optical illuminance transfer matrix is established, and the illuminance of reflection at each point is calculated on the basis of the direct illuminance.

2.1 Calculation of Direct Illumination

As shown in Fig. 1, according to the definitions of illuminance and light intensity, the illuminance on the dA surface can be expressed as

$$E_{di} = \frac{I}{r^2} \cos \theta = \frac{IH}{(L^2 + H^2)^{\frac{3}{2}}}, \quad (1)$$

Fig. 1 Direct illumination metering



where

- I the light intensity of the light source at the illuminated point;
- H the relative height between light source and illuminated point;
- L the horizontal distance between the light source and the illuminated point;
- r the distance between the light source and the illuminated spot; and
- θ the angle between the light source surface and the vertical axis.

Assume that in the three-dimensional coordinate system, the coordinates of the illuminated point are (x, y, z) and the coordinates of the light source are (x_0, y_0, z_0) . Substituting into Eq. (1), it can be deduced that the direct illuminance is

$$E_{di} = \frac{I_i(z_0 - z)}{((x_0 - x)^2 + (y_0 - y)^2 + (z_0 - z)^2)^{1.5}}. \tag{2}$$

2.2 Reflectivity Illuminance Calculation

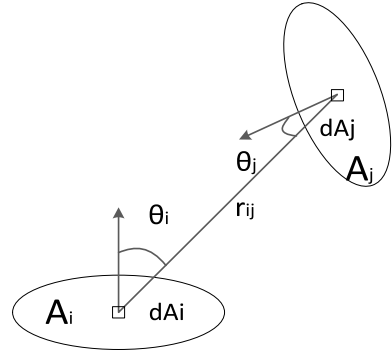
As shown in Fig. 2, the reflection calculation part regards the unit plane where the point j is located as a Lambertian micro-surface, and the radiation light flux exiting from the j plane is equivalent to the light flux received at the i plane, which is expressed as

$$\Phi_i = M_j = \pi L_i. \tag{3}$$

Reflected from j to i surface luminous flux can be expressed as

$$d\Phi_j = \rho E_j dA_j. \tag{4}$$

Fig. 2 Location of illuminated and illuminated point



According to the definition of radiance, illuminance can be described as

$$E_{ij} = \frac{d\Phi_i}{dA_i} = \frac{L dA_i \cos \theta_i \cos \theta_j}{r_{ij}^2}. \quad (5)$$

Substituting (3) and (4) into (5), it can be derived that the reflected illuminance is

$$E_{ij} = \int_{A_j} \frac{\rho_j E_j}{\pi} \frac{\cos \theta_i \cos \theta_j}{r_{ij}^2} dA_j. \quad (6)$$

Consider the reflection surface j as a surface light source consisting of k areas. Then, the reflection illuminance of j toward i is expressed like this.

$$E_{ij} = \sum_{n=1}^k \frac{\rho_j E_{jn}}{\pi} \frac{\cos \theta_i \cos \theta_{jn}}{r_{ij}^2} \Delta A_{jn}. \quad (7)$$

Assuming that the vector between the center of the reflective surface j and the center of the illuminated surface i is $\vec{ij}(x, y, z)$, the illumination on the i surface can be deduced therefrom.

$$E_{ij} = \sum_{n=1}^k \frac{-\rho_j E_{jn}}{\pi r_{ijn}^2} \frac{|\vec{n}_i| \cdot |\vec{i}|}{|\vec{n}_i \cdot \vec{i}|} \frac{|\vec{n}_j| \cdot |\vec{j}|}{|\vec{n}_j \cdot \vec{j}|} \Delta A_{jn}, \quad (8)$$

where $\vec{n}_i$ is the normal vector of the plane i and $\vec{n}_j$ is the normal vector of the plane j .

2.3 Spatial Model Establishment

In order to confirm the feasibility and accuracy of the research method, this article will calculate the illuminance of the space of four different types of rooms or lighting arrangements, as shown in Table 1.

After constructing the spatial light environment model, the distribution of the illuminance of the working surface and the average illuminance value are output, and the results are compared with the results of the DIALux operation. In order to simplify the calculation, all spaces are assumed to have no windows, the heights of the ceiling and work surface are 2.8 and 0.75 m. The reflection ratio of the wall is $\rho_w = 0.5$, the ceiling is $\rho_c = 0.2$, and the floor is $\rho_f = 0.7$. Using a fluorescent lamp power of 30 W and a luminous flux of 1350 lm, the light intensity distribution table was input into the model.

Case 1 (A1): Assuming that the bottom surface is a 6 and 6 m rectangular space, the lighting equipment is laid out as shown in Fig. 3 and the equipment is ceiling mounted.

Case 2 (A2): The layout of lamps and lanterns is shown in Fig. 3.
The floor of the space is unequal height plane, and the side sectional view of the ground is shown in Fig. 4. It consists of two planes. Because the DIALux software

Table 1 Light environment layout case table

Number	Space bottom shape	Lighting arrangement
Case 1	Conventional rectangular space	Uniform arrangement
Case 2	Unequal space on the ground	Uniform arrangement
Case 3	Conventional rectangular space	Unequal heights
Case 4	Round floor space	Ring arrangement

Fig. 3 Distribution plan of indoor lamps

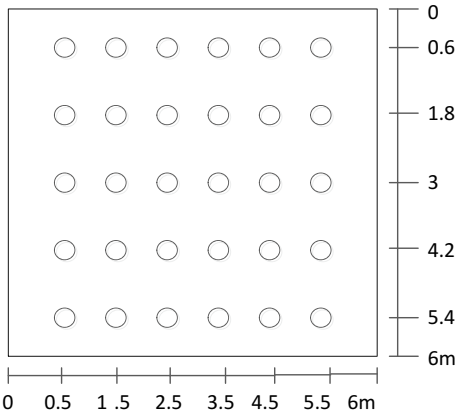


Fig. 4 Unequal ground xz profile

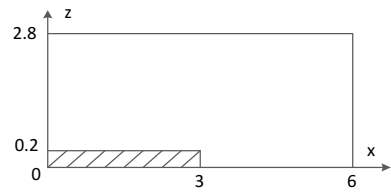
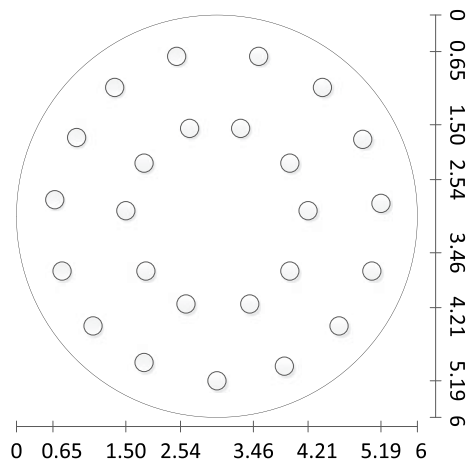


Fig. 5 Case layout of lamp 4



cannot directly set the unequal floor space, an opaque block is placed on the basis of the Case 1 space. Its size is $3 \times 6 \times 0.2$ (m), and the reflectivity is $\rho = 0.2$, and thus to achieve the setting of the spatial light environment.

Case 3 (A3): The layout of lamps is shown in Fig. 3, the left three columns of lamps are ceiling mounted and their coordinates are 2.8 m in the z-axis. The right three columns of light fixtures are suspended at a coordinate of 2.5 m in the z-axis.

Case 4 (A4): The layout of lamps and lanterns is shown in Fig. 5. The space on the ground is a circle with a radius of 3 m. A total of 15 units of equipment are evenly arranged at a distance of 1.5 and 2 m from the center of the circle.

3 Parameter Assignment Analysis

Set the input parameters of the model before performing the simulation, such as the number of reflection illumination calculations (rn) and the calculation step size (cs). Before the evaluation of rn, its influence on the calculation result should be studied, keeping the calculation step constant ($cs = 0.01$), and assigning the reflection calculation times to $rn = (1, 2, 3, 4, 5)$, and then calculate the illuminance of each of the above four case spaces. Similarly, keep $rn = 3$ and let $cs = 0.1, 0.2, 0.3, 0.5, 1$ (m) and calculate the illuminance of the above four case spaces, respectively. Observe

the changes in the calculation times of each calculation index with the reflectance illuminance, compare it with the calculation results of DIALux software, and perform error analysis.

3.1 Reflection Calculations

As shown in Fig. 6, the abscissa represents the number of reflection illuminance calculations (the same below). It can be seen that as the number of reflection calculations increases, the calculation error gradually tends to 0.

As shown in Fig. 7, the calculation error is relatively large, but the error value gradually decreases with the increase of reflection calculation times.

As shown in Fig. 8, with the increase of the number of reflection calculations, the average illuminance calculation result will be closer to the true value, but the program operation time length is relatively increased.

Therefore, when the number of calculations r_n is greater than two, the descending trend of each curve is slowed; when the number of calculations r_n is greater than three, the values of the curves are almost constant. To ensure the accuracy of the model under the premise of minimizing the running time, the number of reflection calculations in the model assignment $r_n = 3$.

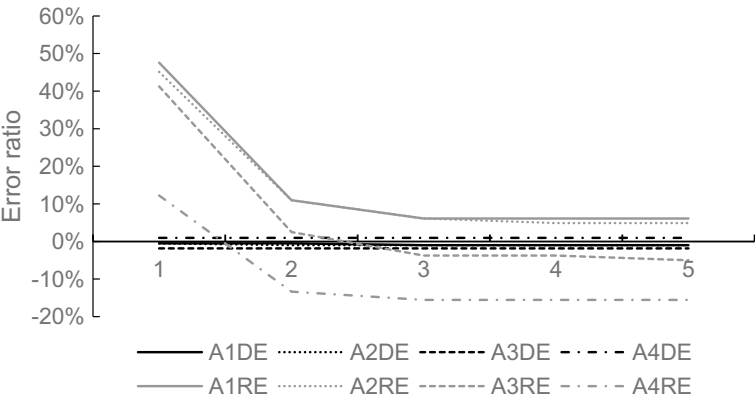


Fig. 6 The error rate of direct irradiance and reflected illuminance varies with the number of reflection calculations

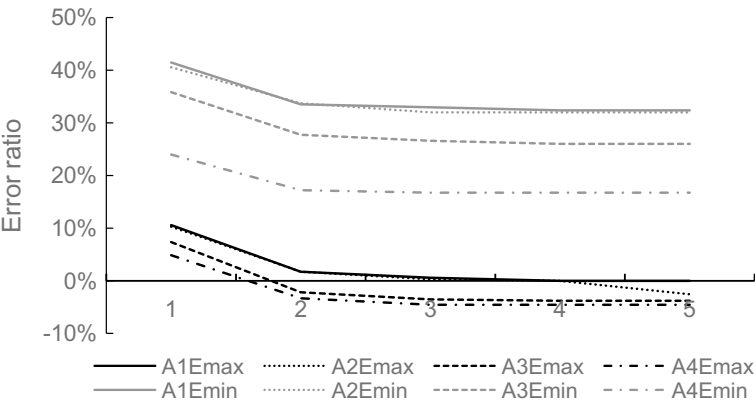


Fig. 7 The error rate of maximum illumination and minimum illumination changes with the number of reflection calculations

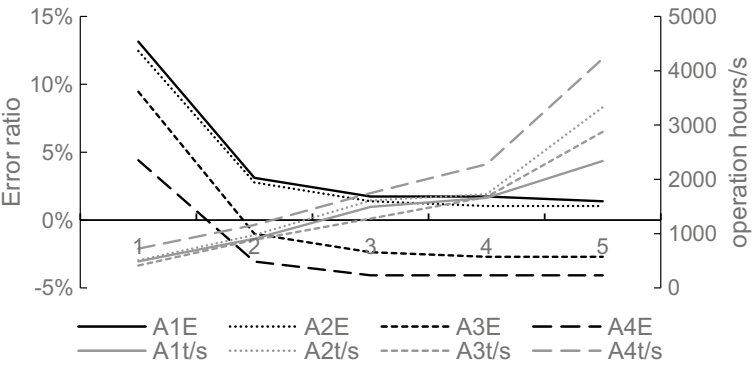


Fig. 8 Average illuminance error rate and program run time vary with reflection calculation times

3.2 Calculation Step

As shown in Fig. 9, the reflectance error ratio of Case 2 increases violently with the increase of the calculation step size. The rest of the illuminance error ratio oscillates with the increase of the calculation step size. After the calculation step cs is greater than 1 m, the divergence condition appears.

As shown in Figs. 10 and 11, the same situation as in Fig. 9 has occurred. The running time decreases with the increase of the calculation step, and the other parameters appear irregular, but the error rate of the illuminance values will increase as a whole, so calculate the step size to take $cs = 0.1$ m.

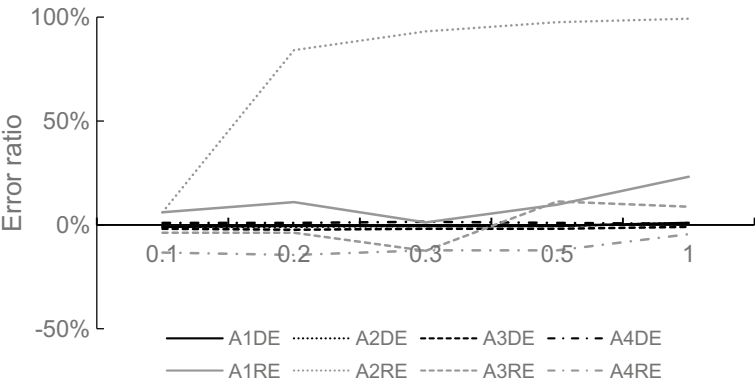


Fig. 9 The error rate of direct irradiance and reflected illuminance varies with different calculation steps

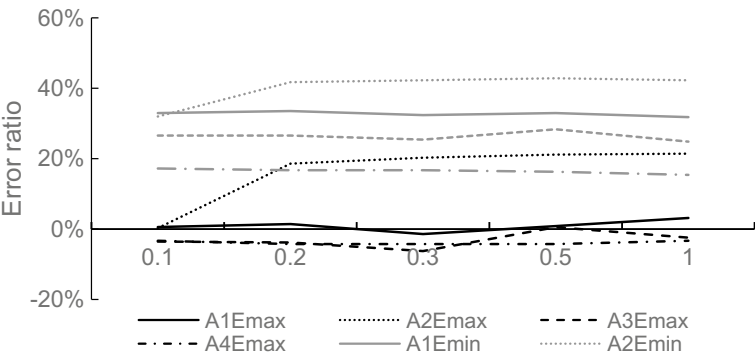


Fig. 10 The error rate of maximum illumination and minimum illumination changes with different calculation steps

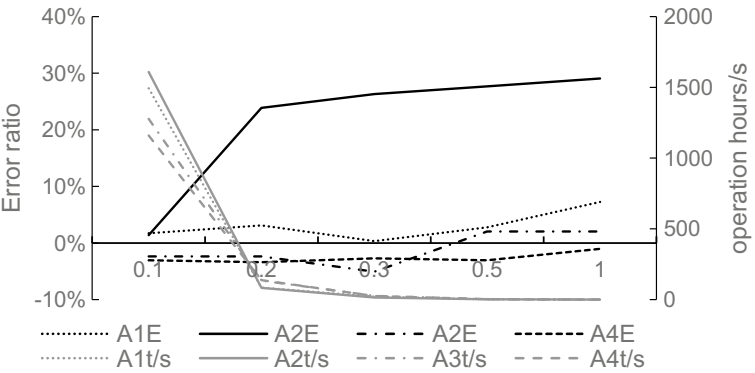


Fig. 11 Average illuminance error rate and program run time vary with different calculation steps

4 Analysis of Simulation Results

The uniformity of illuminance can be defined in this way [10] as

$$U = \frac{E_{\min}}{E_{\max}}, \tag{9}$$

where $E_{\max}$ is the minimum illuminance and $E_{\min}$ is the maximum illuminance. After calculating the illuminance of each point in the space, the average illuminance of the working surface is output as the data result, as shown in Tables 2 and 3.

The illuminance curve, such as the space work surface, is used as a graphic output result and compared with the DIALux output graphic results, as shown in Figs. 12, 13, 14, and 15. When DIALux was used to calculate the space in Case 4, the cylindrical wall was divided into 36 planes. Therefore, in the case of the four-space construction, the calculation step is adjusted to 0.6 m to correspond to the DIALux calculation result.

Comparing the average illuminance values of the work surface calculated based on the complex light environment description method and the DIALux software, the average direct illuminance error is about 2%, the error of reflected illuminance is relatively large, between −7 and 4%, the total average illuminance value calculation result is good, and the error is within ±3%.

Work surface illumination distribution is basically the same, with little error.

However, the error of the minimum illuminance value is relatively large. The possible reason is that when the reflection illuminance is calculated, the light reflection condition is too ideal and is different from the actual situation.

Table 2 Comparison of simulation results

		Direct illumination (lx)	Reflected illumination (lx)	Total illumination (lx)
Case 1	Simulation value	208	77	284
	DIALux	206	82	289
	Error (%)	−0.76	−6.48	0.76
Case 2	Simulation value	208	77	285
	DIALux	206	82	288
	Error (%)	0.77	−5.74	−1.08
Case 3	Simulation value	220	83	303
	DIALux	216	80	296
	Error (%)	2.05	3.76	2.52
Case 4	Simulation value	202	92	293
	DIALux	205	90	295
	Error (%)	−1.63	1.78	−0.60

Table 3 Comparison of illumination uniformity calculation results

		Minimum illumination (lx)	Maximum illumination (lx)	Uniformity
Case 1	Simulation value	119	348	0.34
	DIALux	176	350	0.50
	Error (%)	33.9	0.6	32
Case 2	Simulation value	119	349	0.34
	DIALux	175	350	0.50
	Error (%)	34.1	0.3	32
Case 3	Simulation value	170	381	0.33
	DIALux	173	368	0.47
	Error (%)	26.6	3.5	29.8
Case 4	Simulation value	184	344	0.54
	DIALux	221	329	0.67
	Error (%)	16.7	−4.6	19.4

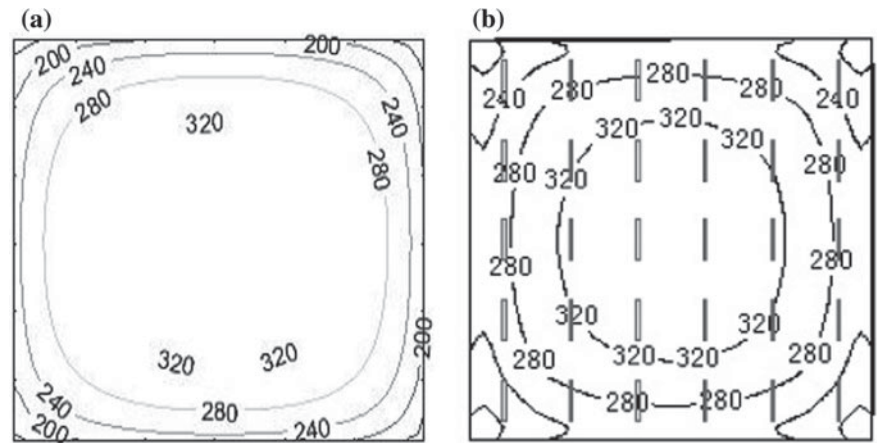


Fig. 12 Case 1: working face illumination distribution of simulation calculation (a) and DIALux software calculation (b)

5 Concluding Remarks

First, this paper proposes a research method for the construction of complex spatial light environment, describes the spatial light environment based on the basic equation of light metrics, and derives Eqs. (2) and (8) to calculate the spatial illuminance. Then, it is used in the field of lighting design, the spatial illumination surface illumination calculation model. Finally, four special cases are listed, and by inputting their spatial parameters into the model, the output results can be obtained and compared with the

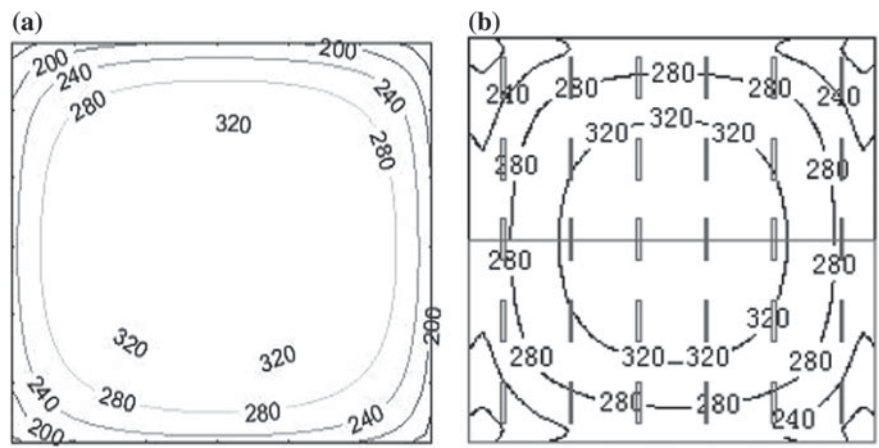


Fig. 13 Case 2: working face illumination distribution of simulation calculation (a) and DIALux software calculation (b)

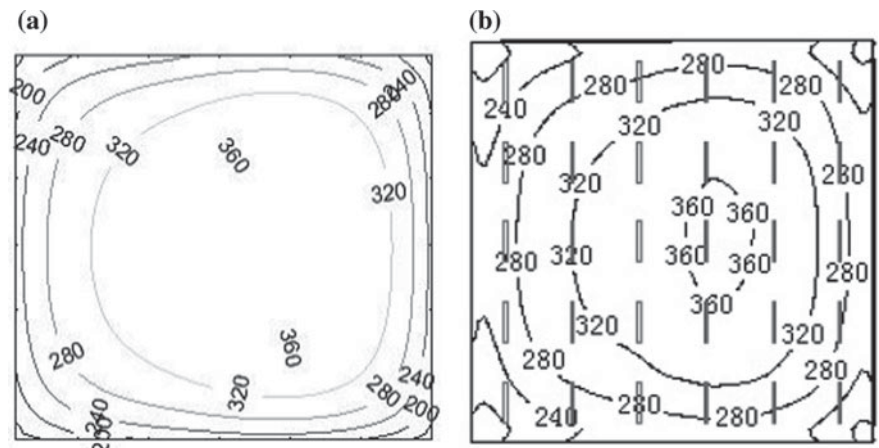


Fig. 14 Case 3: working face illumination distribution of simulation calculation (a) and DIALux software calculation (b)

DIALux software calculation results. After analysis and comparison, the illuminance calculation model proposed in this paper has a small calculation error in the output of the results, and the distribution of illuminance in the work surface is basically the same. It can be applied to the field of light environment design, energy conservation research, and other fields.

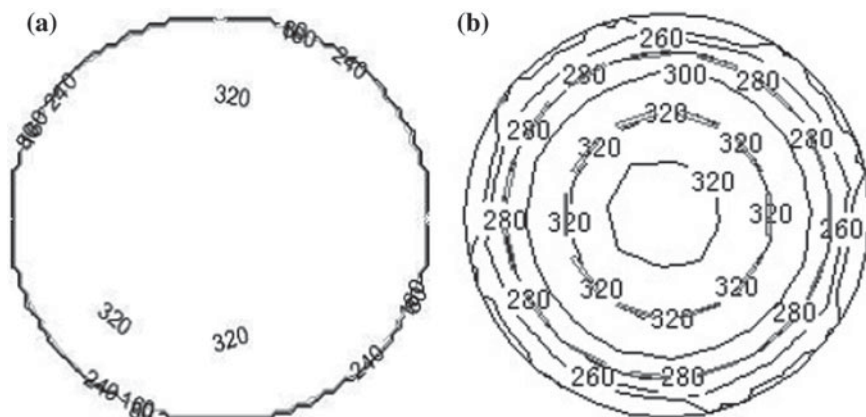


Fig. 15 Case 4: working face illumination distribution of simulation calculation (a) and DIALux software calculation (b)

References

1. Tagliabue, L.C., Buzzetti, M., Arosio, B.: Energy saving through the sun: analysis of visual comfort and energy consumption in office space. *Energy Proc.* **30**(30), 693–703 (2012)
2. An, L.S., Li, G.D.: Computer simulation analysis of illuminance distribution of illumination optical system. *Opt. Technol.* **6**, 45–47 (1998). (in Chinese)
3. Zeng, L.Q., et al.: Research on illumination calculation method based on optical transfer function matrix. *Acta Photo. Sinica.* **17**(4), 5–7 (2006). (in Chinese)
4. Ding, X.D., et al.: A point illumination calculation method. *J. Light. Eng.* **18**(4), 43–49 (2007). (in Chinese)
5. Liu, B.F., Jin, X.L., Zhu, Z.M.: Uniform illumination design based on LED array and diffuse reflection freeform surface. *J. Appl. Opt.* **458**(458), 119–132 (2014)
6. Liu, B.F., et al.: Design of uniform illumination light source based on LED array and diffuse free surface. *Appl. Opt.* **35**(4), 598–602 (2014). (in Chinese)
7. Xu, Y.Y., Jing, X.L., et al.: Design of LED reflector based on uniform illumination. *Opt. Opt. Technol.* **38**(3), 345–349 (2012). (in Chinese)
8. Zhang, J.Q., Fang, X.P.: *Infrared Physics*. Xidian University Press, Xi'an (2004). (in Chinese)
9. Yu, D.Y.: *Engineering Optics*. Beijing: Machinery Industry Press (2015). (in Chinese)
10. Wang, A.J.W., Chen, Y.Y., Teng, Y.T.: Designing uniform illumination systems by surface-tailored lenses and configurations of LED arrays. *J. Display Technol.* **5**(3), 94–103 (2009)

Part VIII

Smart Underground Space

Device-Free Activity Recognition for Underground Spaces Based on Convolutional Neural Network



Qizhen Zhou, Jianchun Xing, Xuewei Zhang and Wei Chen

Abstract Inspired by the excellent work of wireless sensing, we propose a non-invasive activity recognition system, Under-Sense, for underground space sensing with a pair of commodity Wi-Fi devices. Firstly, by extracting relative phase information from all 90 subcarriers, we construct fine-grained images and then compress the rectangle images into k -dimension by singular value decomposition (SVD). A nine-layer convolutional neural network (CNN) is designed to automatically extract important features from constructed images and classify five human activities. Our results show Under-Sense could achieve 99.5% average accuracy in the empty meeting room and 96.7% in complex student studio environment.

Keywords Smart underground space · Wi-Fi-based sensing · Activity recognition · Convolutional neural network

1 Introduction

In the upcoming era of smart underground space, underground Internet-of-Things (UIoT) system is doomed to exert more important influence on the current society, which enables everything communicating through information sensing devices in underground space [1]. The thrust of UIoT needs the integration of comprehensive sensing, reliable transmitting and intelligent processing [2]. In this paper, we mainly alleviate the sensing-related issues with commodity Wi-Fi devices for their low-cost and easy-deployed advantages. Wi-Fi-based activity sensing has become a research hot-pot, including intrusion detection [3], personnel localization [4], path tracking [5] and abnormality monitoring [6]. We first apply this technology in underground space for its promising prospects. For example, we could realize the device-free monitoring of coal miners' location, routine work, and fatigue level to get rid of the potential risks.

Q. Zhou · J. Xing (✉) · X. Zhang · W. Chen
Army Engineering University of PLA, Jiangsu, China
e-mail: xjc@893.com.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_53

In this paper, we elaborately design a non-invasive activity recognition system named Under-Sense, which has following contributions: (1) we utilize robust and sensitive relative phase information and construct images for more detailed feature; (2) we unify the varying size of images through SVD; (3) we design a nine-layer CNN architecture to automatically extract interesting features and recognize labeled activities.

2 Preliminaries

When Wi-Fi signals propagate through indoor spaces, any environment variations could affect the received signals and generate distinctive waveforms. The finer the signal granularity is, the more information could be exploited to recognize the patterns. Owing to the NICs like Intel 5300, channel state information (CSI) could be easily extracted to depict the detailed properties of each subcarrier and refer the physical characteristics of current scenario. Specifically, the fine-grained channel frequency response of subcarrier n between a pair of transmitting antenna (TX) i and receiving antenna (RX) j could be defined as

$$H_{i,j}^n = |H_{i,j}^n|e^{\angle H_{i,j}^n} \quad (1)$$

where $|H_{i,j}^n|$ and $\angle H_{i,j}^n$ represent the amplitude and phase responses of the n -th subcarrier, respectively. Some papers [7] adopt amplitude information for indoor activity recognition but we doubt that amplitude-based features can be easily affected by long-distance attenuations and physical obstacles in underground spaces [8]. In order to capture human dynamic more efficiently, we resort to extract relative phase from conjugate multiplication of CSI of one pair of antennas:

$$h_m^n = H_{i,j}^n \times \text{conj}(H_{i,k}^n) \quad (2)$$

where the synchronization between Wi-Fi NICs can be achieved and thus the unknown phase shifts can be reduced. In this paper, as we extract CSI for 30 subcarriers of a 20 MHz channel and use 1 and 3 omnidirectional antenna as TX and RX respectively, $m = 1, 2, 3$. We reconstructed the response matrix h_{res} as:

$$h_{res} = [h_1^1, \dots, h_1^{30}, h_2^1, \dots, h_2^{30}, h_m^1, \dots, h_m^{30}]^T = |h_{res}|e^{\angle h_{res}} \quad (3)$$

considering $\angle h_{res} = -\frac{2\pi}{\lambda} \Delta d$, λ is the wavelength and Δd denotes the changing path length between two antennas, we sensitively detect micro-movements caused by human interferences in a cm-level.

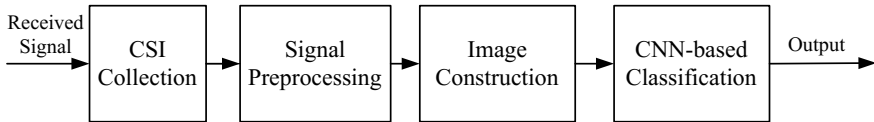


Fig. 1 The under-sense framework

3 The Under-Sense System

3.1 System Overview

Figure 1 illustrates the framework of proposed activity recognition system. Basically, Under-Sense is divided into 4 parts: CSI collection, signal preprocessing, image construction and CNN-based activity classification. Two ThinkPad laptops equipped NICs are used as transmitter (set to the injection modes) and receiver (set to the monitor modes), respectively. The sampling rate is set as 50 packet/s to acquire sufficient motion information and the links are set up to work on Chaneel 165 at 5.825 GHZ for pure channel properties. In the following step, we extract relative phase information by the conjugate multiplication of CSI of antenna pairs and then calculate the cross-correlation between adjacent subcarriers to indicate the trans-points during human activities. Next, we construct the two-dimensional relative phase measurement matrix with the same k -dimension by using SVD to reduce dimension. A 9-layer CNN architecture is then adopted to automatically extract distinctive features from constructed images and realize robust classification. The detailed implementation of each module is presented in the following sections.

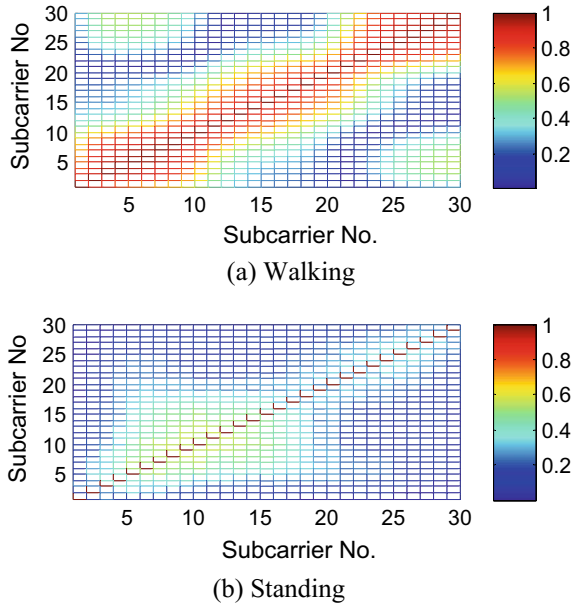
3.2 Signal Preprocessing and Image Construction

To ensure the real-time and reliable activity detection, the signal sanitation and segmentation methods are critical to the overall performance of proposed system. After phase conjugate multiplication, we obtain the robust relative phase information and then utilize the cross-correlation between relative phases of subcarriers in the same antenna pair. The rationale is that when people perform actions in wireless environment, the correlation between adjacent subcarriers could increase [5].

In Fig. 2, we carry out a simple experimental study by calculating the cross-correlation of all 30 subcarriers in $i \leftrightarrow j$ antenna pair. Noticeably, when human walk, the cross-correlation between adjacent 5 subcarriers increase. Therefore, we design a motion indicator based on cross-correlation of 5 continuous subcarriers to detect the start and end points of human motion.

After signal segmentation step, we obtain signal slices with different sizes and try to construct squared images for further CNN-based feature extraction. In this paper,

Fig. 2 Cross-correlation between relative phases of subcarriers in the same antenna pair



we alleviate the problem through SVD to approximate the original images with fewer dimensions. Specifically, SVD is a factorization of complex matrix in linear algebra, whose mathematical equation can be written as:

$$h_{resm \times n} = U_{m \times m} \sum_{m \times n} V_{n \times n}^T \quad (4)$$

where $h_{resm \times n}$ denotes the m relative phase data acquired in n subcarriers, U and V represents m -dimension and n -dimension unitary matrix, respectively. $\sum_{m \times n}$ means a $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal.

The Eq. (4) can also be approximated as (5) if the singular value $[\sigma_{k+1} \dots \sigma_m]$ is small enough:

$$h_{resm \times n} = [u_1 \dots u_k | u_{k+1} \dots u_m] \begin{bmatrix} \sigma_1 & & \vdots \\ & \ddots & 0 \\ & & \sigma_k \\ & & \vdots \\ \dots & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \\ v_{k+1}^T \\ \vdots \\ v_n^T \end{bmatrix}$$

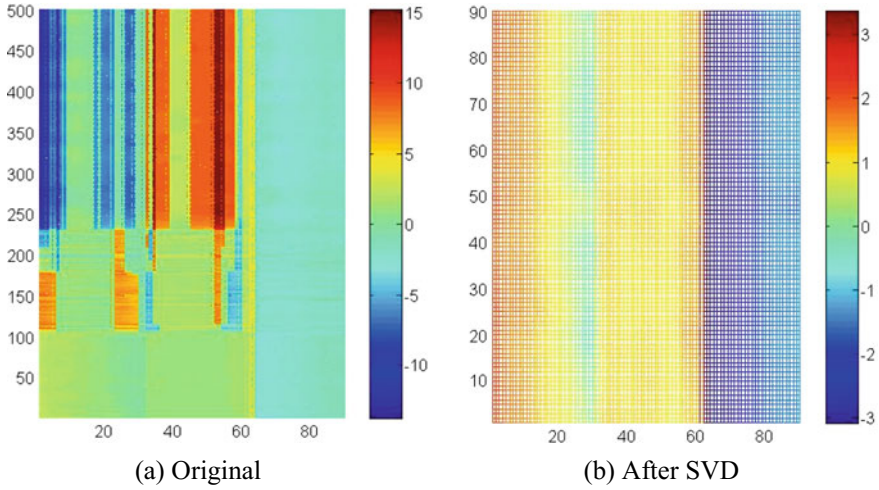


Fig. 3 Comparison between **a** original 500×90 and **b** processed 90×90 image

$$\approx [u_1 \dots u_k] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} \quad (5)$$

where $[u_1 \dots u_k]$ and $\begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix}$ are called left-singular and right-singular vectors for σ_k , respectively. By discarding those slight diagonal entries of $\sum_{m \times n}$, we reconstruct and unify the k -dimension state matrix.

In Fig. 3a, we can see that the original image cannot be directly put into the CNN because of the non-square shape. After SVD, though the resolution ratio of image decreases, we reduce the dimension of pictures for rapid online recognition and reserve the major motion information.

3.3 CNN-Based Feature Extraction and Classification

In the following step, we utilize a nine-layer CNN [9] to automatically extract features of interest. Our CNN consists of one input layer, four convolutional layers, three subsampling layers, one fully-connected layer, and one output layer. Generally, convolutional layers apply a convolution operation to the input and allow for a deeper network with fewer parameters based on the shared weights. Subsampling act as a form of non-linear down-sampling, which partition the input images into non-

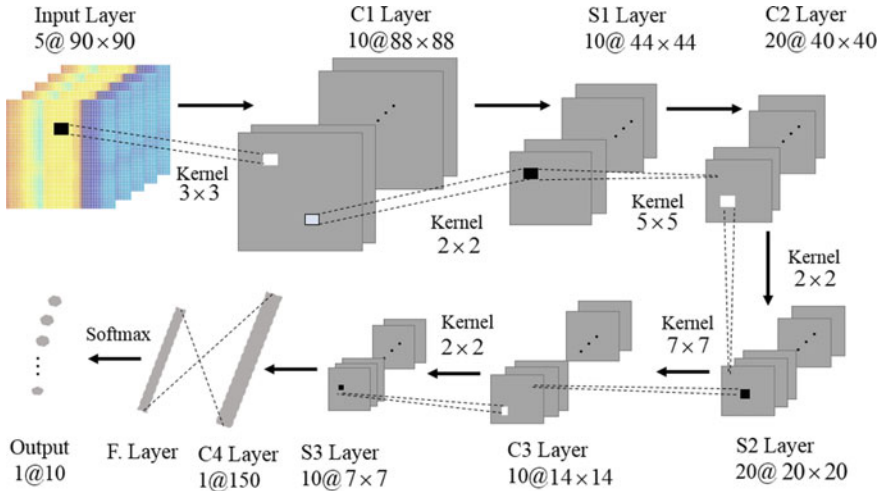


Fig. 4 CNN-based activity recognition architecture

overlapping rectangles and serve to reduce the computation complexity and control overfitting. After a period of superposition of convolution layers and subsampling layers, the high level reasoning in deep learning network is done by a fully-connected layer.

In Fig. 4, we can see the architecture and parameter settings of the nine-layer CNN. For starter, we construct five relative phase based images with size 90×90 from the repetitive-motion recordings. For each images in the first convolutional and subsampling layer, we adopt 10 3×3 convolutional kernels and stride of 1 in both vertical and horizontal directions to obtain the same number of feature map and then we adopt a 2×2 kernel to reduce the dimension of feature map and control overfitting. Next, there are two kernels with size 5×5 to extract more abstract features in C2 layer and then subsampling with size 2×2 in S2 layer to improve generalization of the CNN. We finally get a single convolutional layer with 150 neurons and send the feature maps into fully connected layer to combine all extracted features. Softmax function is applied to classify 10 predefined activities.

4 Experimental Results

In this section, we elaborate the experiment setup and present the overall performance of Under-Sense.

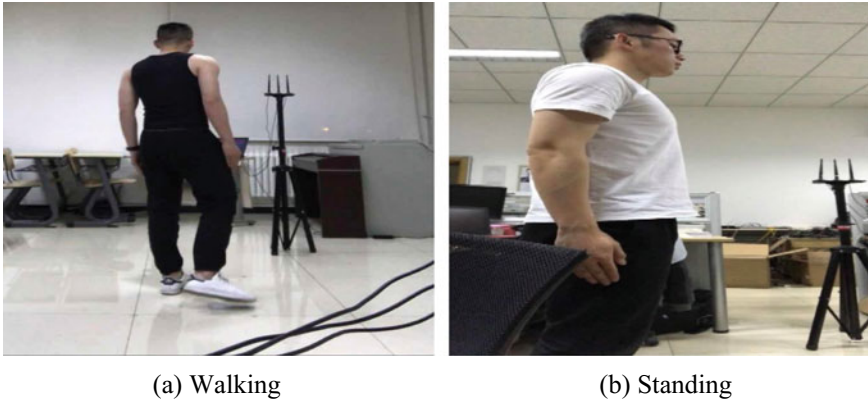


Fig. 5 Experiment setups

4.1 Experiment Setup

We implement Under-Sense with a pair of 5 GHz commodity Wi-Fi devices which run Ubuntu 11.04 operating system. The packet transmission rate in TX is fixed at 50 packet/s to enable daily activity recognition because normal human activities are no more than 1 m/s ($V_{upper} = \frac{f\lambda}{2} \approx \frac{50 \text{ Hz} \times 0.05 \text{ m}}{2} = 1.25 \text{ m/s}$) [10]. The antennas of RX are placed loosely in a line with a half wavelength (2.6 cm) for 5.825 GHz Wi-Fi. In Fig. 5, we emulate the wireless environment for underground spaces in two typical indoor environment (a) empty meeting room and (b) complex student studio. We ask 5 volunteers (3 males and 2 females) to repeat 5 predefined daily activities (i.e., standing, sitting, waving, waking, jumping) 20 times every day in a two-week period.

4.2 Overall Performance

We use confusion matrix to evaluate our system performance. Each column is the confusion matrix represents the actual instances and each row denotes the classified activities. All correct classifications are located in the diagonal of the matrix, while the errors are represented by values outside the diagonal. As can be seen in the Fig. 6, our system achieves average accuracy of 99.5% in the empty meeting room and 96.7% in the complex student studio. We find that the performance has varying but acceptable decreases in 5 different activities, especially in standing and sitting. The reason is that relatively stationary human motions are more difficult to extract features and vulnerable to the environment interferences. Based on CNN-based feature extraction and classification module, we could automatically obtain distinctive features and realize efficient recognition.

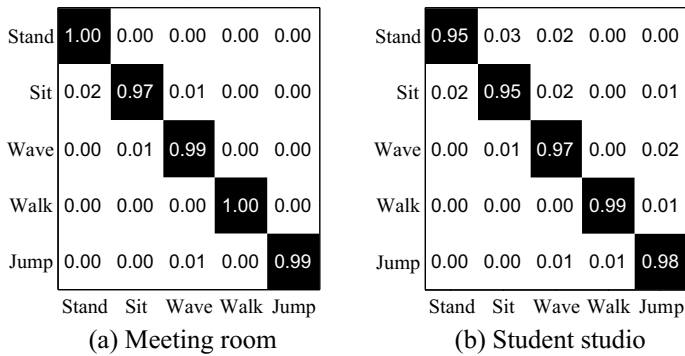


Fig. 6 Confusion matrix of activity recognition in two typical scenarios

5 Conclusions

In this paper, we propose a CNN-based activity recognition approach for underground spaces with commodity Wi-Fi devices. First, relative phase information is extracted from raw CSI data and then cross-correlation values between adjacent subcarriers are utilized to indicate human motions. In order to obtain sufficient features for accurate recognition, we reconstruct the k -dimension phase images by SVD for the input of the deep learning network and resort to a nine-layer CNN to automatically extract and recognize activity features. Multiple experiments are implemented in two typical indoor scenarios. The results demonstrate the effectiveness of our proposed activity recognition system. In the future, we would further explore the potentials of other wireless signals to achieve contact-free sensing for specific use.

Acknowledgements, Ethical Approval and Informed Consent This work is supported by the National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques).

This study has been conducted with ethics approval obtained from the ethics committee: Prof. Qiliang Yang and Prof. Ping Wang of Department of National Defense Engineering, Army Engineering University of PLA, Jiangsu, China.

Participants are recruited from Army Engineering University of PLA, Jiangsu, China, and all participants have given their consent to use the dataset and disclose information relevant for research in this study.

References

1. Stankovic, J.: Research directions for the internet of things. *IEEE Internet Things J.* **1**(1), 3–9 (2014)
2. Liu, T., Lu, D.: The application and development of IoT. In: *Information Technology in Medicine and Education (ITME)*, 2012 International Symposium, vol. 2, pp. 991–994. IEEE (2012)

3. Qian, K., Wu, C., Yang, Z., Liu, Y., Zhou, Z.: PADS: passive detection of moving targets with dynamic speed using PHY layer information. In: *Parallel and Distributed System, 2014 20th IEEE International Conference*, pp. 1–8. IEEE (2014)
4. Wang, X., Wang, X., Mao, S.: Cifi: deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi. In: *Communications (ICC), 2017 International Conference*, pp. 1–6. IEEE (2017)
5. Qian, K., Wu, C., Yang, Z., Liu, Y., Jamieson, K.: Widar: decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi. In: *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM (2017)
6. Zhou, Q., Wu, C., Xing, J., Li, J., Yang, Z., Yang, Q. Wi-Dog: Monitoring school violence with commodity Wi-Fi devices. In: *International Conference on Wireless Algorithms, Systems, and Applications*, pp. 47–59. Springer, Cham (2017)
7. Wang, X., Gao, L., Mao, S., Pandey, S.: CSI-based fingerprinting for indoor localization: a deep learning approach. *IEEE Trans. Veh. Technol.* **66**(1), 763–776 (2017)
8. Wu, C., Yang, Z., Zhou, Z., Qian, K. PhaseU: real-time LOS identification with WiFi. In: *Computer Communications*, pp. 2038–2046. IEEE (2015)
9. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: *International Conference on Neural Information Processing Systems*, vol. 60, pp. 1097–1105. Curran Associates Inc. (2012)
10. Qian, K., Wu, C., Zhou, Z., Zheng, Y., Yang, Z., Liu, Y.: Inferring motion direction using commodity Wi-Fi for interactive exergames. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1961–1972 (2017)

A Decentralized Parallel Kalman Filter in Multi-sensor System for Data Verification



Guoping Li and Shiqiang Wang

Abstract In order to ensure the successful completion of the tasks of a system, the accuracy of data is the foundation and indemnification. To complete the data verification on the decentralized computing platform, a decentralized Kalman filter with state constraints is presented in this paper. The decentralized sensing architecture takes the form of a network of transputer-based sensor nodes, each with its own processing system. So it does not require any central processor or common clock. Based on that, this new algorithm can allow fully decentralization of the multisensory Kalman filter equations with state equality constraints among a number of sensing nodes to verify the data. The algorithm is developed from a centralized method named projection method to minimized the communication among nodes and can take place without any prior synchronization between nodes. Theoretical derivation is provided to the decentralized algorithm. Finally, the case study of the secondary chilled water pump system illustrates the effectiveness of the proposed method.

Keywords Decentralized computing platform · Decentralized Kalman filter · Data verification · State equality constraints · Projection estimation

1 Introduction

Nowadays, the Internet of Things is playing an increasingly important role. The intelligent building is an important part of that. With the world's first intelligent building has been built in the United States since the early 1980s, IB (intelligent building) has got the world's attention [1]. Although the IB has developed from a single building to smart city with complex building facilities, the measurement signal of the sensor is the basis to control the IB. The higher the level of automation and intelligence of IB, the greater the degree of dependence on the sensor. Sensor validity

G. Li (✉) · S. Wang
College of Defense Engineering, Army Engineering University of PLA,
Nanjing, Jiangsu, China
e-mail: 215506349@qq.com

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_54

can play a crucial role in improving the system availability, safety, and reliability. When the sensor's performance reduced or breaks down, it will bring serious impact to IB, such as increasing energy consumption, or even causing Equipment failure [1].

Thus, a significant component of sensor fault detection and diagnosis (SFDD) method has been investigated [2]. Kalman filter is the basic approach for model-based SFDD and can be used to estimate the states of a dynamic system [3]. Many researchers have improved the Kalman filter. Zhang et al. [4] studied the state estimation problem of a networked linear system where a sensor and an estimator are connected via a lossy network with the measurement loss by the intermittent Kalman filter (IKF) and further address the nonlinearity of the networked system. Geng et al. [5] presented a scheme for integrated navigation systems with applications of sensors data fusion and fault-tolerant Kalman filter (FTKF). Through a series of extended Kalman filters (EKF) using an innovative initialization of covariance matrix in system dynamics, Jayaram S. presented an approach to detect and isolate the sensor failures [6]. An unscented Kalman filter (UKF) has been developed in [7] to utilize multi-sensor data fusion techniques for process fault monitoring.

However, these approaches have been applied to the measurement of individual variable. As the development of measuring technology, the observed information becomes diverse and abundant [8]. More and more constraints among numerous variables can be established. To make full use of the analytic redundancy to further improve the accuracy of the state estimation of the system, Simon D. presented a method for incorporating linear state constraints in a Kalman filter [9].

Besides, as the broad use of sensors, modern engineering systems are ever increasing in complexity and tend to be distributed in wide areas. On one hand, the sensor network system scale also has been increasing. On the other hand, a sensor network always consists of spatially distributed sensors to monitor various status of a system, events phenomena or environmental variables such as temperature, vibration, pressure, motion and so on. Thus, centralized monitor cannot satisfy the need of system.

Based on the above analysis, a fully decentralized sensor diagnostic method is required for the distributed sensor network. A new and entirely different system with decentralized architecture has been developed by the Building Energy Research Center in Tsinghua University [1]. There is not any central node. Thus the data measured by sensors will be able to self-organized to check and exchange, avoiding the cumbersome delivery and storage tasks. In addition, the constraints between the nodes are based on the objective law of physical constraints and energy conservation, without setting the detection threshold. It has strong adaptability and can detect the sensor failure in real time. Under this structure, each sensor is also fully distributed. Each sensor is fitted with a microprocessor unit such that it becomes a smart sensor [10]. Each smart sensor can communicate with other nodes and operate collaboratively to complete sensor fault detection and self-repair tasks.

The main contribution of this paper is to propose a decentralized parallel Kalman filter with local state constraints under the innovative sensor network structure. Section 2 shows the problem formulation and the instruction of this new structure.

Section 3 presents a brief summary of the Kalman filter solution without any state constraints, then, infers a new Kalman filter. Section 4 presents some simulation results and Sect. 5 closes with some concluding remarks and suggestions for further work.

2 Problem Formulation

First, it is considered a connected sensor network composed of n sensor nodes that are responsible for collection of interested measurements. The Kalman filter is to obtain the satisfactory estimate $\mathbf{x}$ by time iteration. In the traditional centralized steepest-descent solution, $\mathbf{x}$ can be computed by the fusion center or sink node.

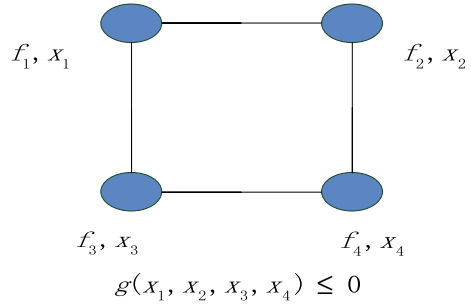
In the above entirely decentralized structure, it consists of a large number of “intelligent nodes”, each of which corresponds to a building subspace or an electromechanical device unit. The “intelligent node” in the centerless system is called CPN (Computing Process Node). The CPN only communicates with its neighbors, and the CPNs are connected to each other according to their space distribution in buildings.

In this system, each CPN is seen as a sensor which can form a decentralized network. The sensor network can be simplified as an undirected connected graph $G = (V, E)$, the vertex set V contains nodes and E is the edge set. Each vertex of the graph is referred to as a node and endowed with a state $\mathbf{x}_i(k)$, and every $i \in V$. The edge $(i, j) \in E$ refer that node i and node j can establish a communication link between them.

The neighbors of node i is denoted as $N_i = \{j | (i, j) \in E\}$. The set $\mathbf{x}_{ij}, j \in N_i$ is introduced to denote the coupled variable, which is the decision variable shared by nodes i and j .

The advantage of decentralized approaches is that they remove the requirement of monitoring host and enable each updated sensor-node to complete sensor fault detection and self-repair task based on local and neighboring information. Different from conventional mode where global information can be available by a central point, each node only has a local copy of the neighbors’ decision variable. A simple model is shown in Fig. 1.

Fig. 1 Decentralized and flat sensor network



3 Unconstrained and Decentralized Constrained Kalman Filter

3.1 Unconstrained Kalman Filter

This section reviews unconstrained state estimation via the Kalman filter and some important equations of the filter. These will be used in this paper. The results and notation are taken from [11]. More details can be seen in [11]. We are not talking more about it.

It is the basis of the new method in this paper.

3.2 Decentralized Constrained Kalman Filter

The DCKF (decentralized constrained Kalman filter) is derived from a centralized constrained Kalman filter that developed onto a decentralized architecture. This method is based on another method which is called projection method [12]. In this section, we develop the constrained Kalman filter by directly projecting the unconstrained state estimate $\mathbf{x}$ onto the constraint surface.

But there are some deficiencies in this algorithm. First, this centralized method does not take the effects of constraints on noise into account. Then, this projection is only acting on the estimate, the covariance matrix $\mathbf{P}$ and Kalman Gain $\mathbf{K}$ is not been projected. Thus, the accuracy would be affected.

Based on the above, this paper proposes a decentralized whole process projection. This algorithm is based on the above system, after the preliminary state prediction, project the estimate and covariance matrix onto the constraint surface immediately. Thus, the next prediction and iteration are fully based on the constraints. More importantly, this method fully applicable to the decentralized network.

The observation vector of $\mathbf{z}(k)$ is unstacked into n subvectors corresponding to each individual sensor

$$\mathbf{z}(k) = [\mathbf{z}_1^T(k), \dots, \mathbf{z}_n^T(k)]$$

and partition the observation matrix into submatrices corresponding to these observations

$$\mathbf{H}(k) = [\mathbf{H}_1^T(k), \dots, \mathbf{H}_n^T(k)]$$

The noisy vector is also decoupled,

$$\mathbf{v}(k) = [\mathbf{v}_1^T(k), \dots, \mathbf{v}_n^T(k)]$$

and the partition of observation noise covariance is

$$\mathbb{E}[\mathbf{v}(k)\mathbf{v}^T(k)] = \text{diag}[\mathbf{R}_1(k), \dots, \mathbf{R}_n(k)]$$

and these partitions are uncorrelated.

The system model $\mathbf{F}(k)$ and the state $\mathbf{x}(k)$ can, in general, be portioned in the same way as the observations.

Letting $\mathbf{F}_i(k)$ be the sensors local system model, $\mathbf{G}_i(k)$ is the local noise model, $\mathbf{w}_i(k)$ is the local input noise, and $\mathbf{Q}_i(k)$ is the covariance of local input noise. Note that both $\mathbf{w}_i(k)$ and $\mathbf{v}_i(k)$, $i = 1, 2, \dots, n$ are uncorrelated.

Thus, each individual sensing node can have its own model of the system. The local state transition and observation equations of node i are shown as

$$\begin{aligned}\mathbf{x}_i(k) &= \mathbf{F}_i\mathbf{x}_i(k-1) + \mathbf{G}_i\mathbf{w}_i(k-1) \\ \mathbf{z}_i(k) &= \mathbf{H}_i\mathbf{x}_i(k) + \mathbf{v}_i(k)\end{aligned}\tag{1}$$

Besides, the consensus constraints of sensors i and j ($i \neq j$) are expressed as

$$A_i\hat{\mathbf{x}}_i - B_j\hat{\mathbf{x}}_j = \mathbf{d}_i, j \in N_i, i = 0, 1, 2, \dots, n,\tag{2}$$

where A_i and B_i are the partial adjacency matrix.

First, in order to solve the state estimation with the equation constraints of node i at time k , using the standard KF algorithm to find the priori state estimation of the unconstrained linear system of node i .

$$\hat{\mathbf{x}}_i^{\text{upe}}(k) = \mathbf{F}_i(k-1)\hat{\mathbf{x}}_i^{\text{pe}}(k-1) + \mathbf{G}_i(k-1)\mathbf{u}_i(k-1)\tag{3}$$

where $\hat{\mathbf{x}}_i^{\text{upe}}(k)$ is the estimation of node i at time k with no projection, and $\hat{\mathbf{x}}_i^{\text{pe}}(k-1)$ is the estimation of node i at last time with projection. To get the estimation at time k with constraints, which is $\hat{\mathbf{x}}_i^{\text{pe}}(k)$, the equation are shown as

$$\begin{aligned}
\min J(\hat{\mathbf{x}}_i^{pe}k) &= \arg \min [\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}k]^T \mathbf{W}(k) \\
&\quad \cdot [\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}k] \\
\text{s.t } A_i(k)\mathbf{x}_i(k) - B_i(k)\mathbf{x}_j(k) &= \mathbf{l}_i(k)
\end{aligned} \tag{4}$$

where $\mathbf{d}_i \in R^{m \times 1}$ is constant matrices with suitable dimensions.

The Lagrangian of the constrained problem is

$$\begin{aligned}
L[\hat{\mathbf{x}}_i^{pe}k, \boldsymbol{\lambda}(k)] &= [\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}k]^T \mathbf{W}(k) [\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}(k) \\
&\quad + \boldsymbol{\lambda}(k) [A_i(k)\mathbf{x}_i(k) - B_i(k)\mathbf{x}_j(k) - \mathbf{l}_i(k)]
\end{aligned} \tag{5}$$

where $\boldsymbol{\lambda}(k) \in \mathbb{R}^k$ is the Lagrangian multiplier.

$$2\mathbf{W}_i(k)[\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}k] + A_i(k)\boldsymbol{\lambda}(k) = 0 \tag{6}$$

$$A_i(k)\mathbf{x}_i(k) - B_i(k)\mathbf{x}_j(k) - \mathbf{l}_i(k) = 0 \tag{7}$$

The Lagrangian multiplier can be obtained by combing Eqs. (8) and (9).

$$\begin{aligned}
\boldsymbol{\lambda}(k) &= 2[A_i(k)\mathbf{W}_i(k)^{-1}A_i(k)^T]^{-1} \\
&\quad \cdot [A_i(k)\hat{\mathbf{x}}_i^{pe}k - B_i(k)\mathbf{x}_j(k) - \mathbf{l}_i(k)]
\end{aligned} \tag{8}$$

Further, by substituting these results into Eq. (8), the optimal estimate can be got as

$$\begin{aligned}
\hat{\mathbf{x}}_i^{pe}(k-1) &= \hat{\mathbf{x}}_i^{upe}(k-1) - \mathbf{W}_i^{-1}(k)A_i^T(k) \\
&\quad [A_i(k)\mathbf{W}_i^{-1}A_i^T(k)]^{-1} [A_i(k)\hat{\mathbf{x}}_i^{pe}k - B_i(k)\mathbf{x}_j(k) - \mathbf{l}_i(k)]
\end{aligned} \tag{9}$$

To simplify the representation, let

$$\boldsymbol{\Theta}_i(k) = \mathbf{W}_i^{-1}(k)A_i^T(k)[A_i(k)\mathbf{W}_i^{-1}(k)A_i^T(k)]^{-1} \tag{10}$$

$$\boldsymbol{\Upsilon}_i(k) = \mathbf{I} - \boldsymbol{\Theta}_i(k)A_i(k) \tag{11}$$

Thus, the (11) can be represented as follow

$$\begin{aligned}
\hat{\mathbf{x}}_i^{pe}(k-1) &= \hat{\mathbf{x}}_i^{upe}(k-1) - \boldsymbol{\Theta}_i(k) \\
&\quad [A_i(k)\hat{\mathbf{x}}_i^{pe}k - B_i(k)\mathbf{x}_j(k) - \mathbf{l}_i(k)]
\end{aligned} \tag{12}$$

So, the covariance matrix after projection can be shown as

$$\mathbf{P}_i^{pe}(k-1) = E[\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}(k-1)][\mathbf{x}_i(k) - \hat{\mathbf{x}}_i^{pe}(k-1)]^T$$

$$\begin{aligned}
&= E\{\Upsilon_i(k)[x_i(k) - \hat{x}_i^{upe}(k-1)] \\
&\quad [x_i(k) - \hat{x}_i^{upe}(k-1)]^T \Upsilon_i^T(k)\} \\
&= \Upsilon_i(k) P_i^{upe}(k-1) \Upsilon_i^T(k)
\end{aligned} \tag{13}$$

Then, update the status, the equations shown as

$$\hat{x}_i^{pe}(k) = \hat{x}_i^{pe}(k-1) + K^{pe} [y_i(k) - H_i(k) \hat{x}_i^{pe}(k-1)] \tag{14}$$

$$K^{pe} = P_i^{pe}(k-1) H_i(k)^T [H_i(k) P_i^{pe}(k-1) H_i(k)^T + R_i]^{-1} \tag{15}$$

$$P_i^{pe}(k) = [I - K^{pe} H_i(k)] P_i^{pe}(k-1) + Q_i(k) \tag{16}$$

The main processes are as follows:

Step 1: Each node i , $i = 1, \dots, n$, can pass their own measurements to neighboring node j , $j \in N_i$ and receive its neighbors' evaluations;

Step 2: System state initialization, define the initial value, where $= 1$;

Step 3: Calculate the priori predicts as (2);

Step 4: Choose a suitable weighted matrix W_i . Usually, let $W_i = I$ to get optimal state estimate in a sense of minimum mean square error [12];

Step 5: Calculate the state predicts in constrains surface by (8), (11), (12);

Step 6: Update the measure by (14)–(16); and

Step 7: Let $k = k + 1$, jump to step 3.

4 Simulation Results

In order to verify the effectiveness of the algorithm, in this section, we use the case that the secondary chilled water pump system to illustrate the efficacy of the DCKF. The air conditioning system model is shown in Fig. 2.

As is in Fig. 2, the system architecture and sensor layout are as shown in Fig. 2. 16 temperature sensors T1 to T16 of this system are set up for data estimate.

The decentralized network model based on the constraint relationship between sensor measurement variables is shown in Fig. 3.

According to the physical knowledge of the air conditioning system, establish the following constraints

$$\begin{aligned}
t_{14} &= t_{15} = t_{16} \\
t_3 &= t_4 = t_5 = t_6 = t_7 = t_8
\end{aligned}$$

The supply water temperature of chilled water that $t_1 \sim t_8$ is between 5 and 10 °C, the return water temperature of chilled water that $t_9 \sim t_{16}$ is between 8 and 15 °C.

Fig. 2 Air conditioning system model

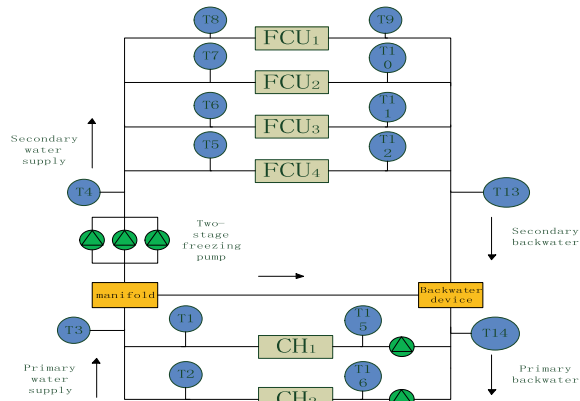
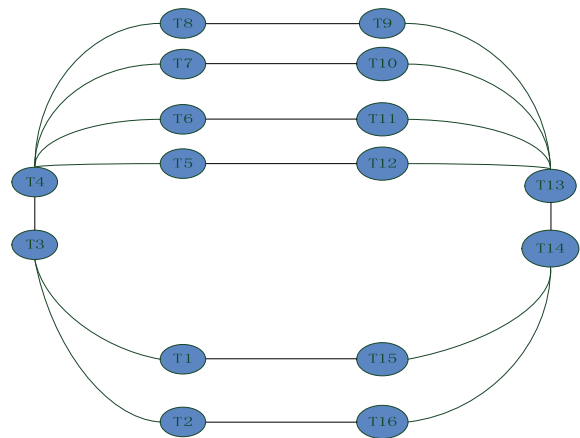
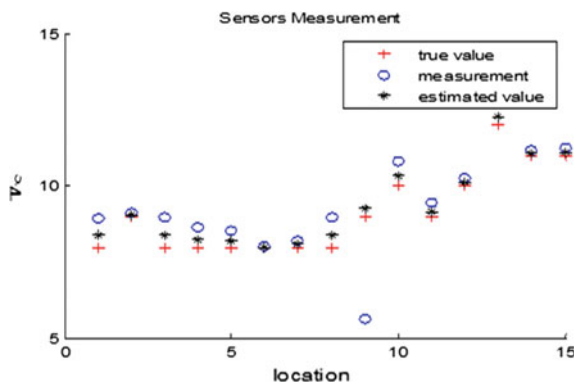


Fig. 3 Decentralized sensor network model



In the simulation study, we add a noise value to the true value as a measurement value. Because the noise is ubiquitous in actual projects. Assuming that the measurement of T_9 takes a large drift. According to the constraints between sensors, the simulation result is as shown in Fig. 4.

It can be seen that, according to the measurement value, estimate value has a higher accuracy. It can be calculated that the error is between 0.1 and 0.5, the percentage of error is less than 2%. As to T_9 , the DCKF can get a better-estimated value to satisfy the need of temperature sensors.

Fig. 4 Simulation results

5 Conclusions

In this paper, the novel sensor fault detection and self-repair method based on a decentralized network structure is proposed to solve actual engineering problems. It can do good in data verification. Its advantages are as follows:

- (1) It is based on a decentralized architecture for sensor diagnosis network. This intelligent system which is no center is showing great potential and superiority. The sensor-node communicates with neighboring nodes such that they can work collaboratively to achieve the sensor fault detection and self-repair task.
- (2) Decentralized optimization is used to solve the task of data verification under the novel network structure to achieve effectiveness of global goals. It uses a method which is called projection to improve the KF with the equality state constraints. Simulation results presented in the preceding paragraph indicate that the proposed method can perform good in this aspect.

In this paper, the results of the decentralized Kalman filter algorithm are only based on equality state constraints, and the cases that involve inequality constraints are not studied. But inequality constraints relations are common in actual situation, how to deal with this kind of problem, is quite significant for theory and engineering research. That will be studied in depth.

Acknowledgements This work is supported by National Key Research and Development Project of China No. 2017YFC0704100 (entitled New generation Intelligent building platform techniques)

References

1. Qi, S.: Studies on architecture of decentralized system in intelligent building. Tsinghua University (in Chinese) (2015)
2. Lin, et al.: Unbalanced distribution network fault analysis with hybrid compensation. IET Gener. Transm. Distrib. **5**(1), 92–100(2010)

3. Basar, T.: A new approach to linear filtering and prediction problems. Wiley-IEEE Press, pp. 35–45 (2009)
4. Zhang, J., You, K., Xie, L.: Bayesian Filtering with unknown sensor measurement losses. *IEEE Trans. Control. Netw. Syst.* **49**(22), 315–320 (2016)
5. Geng, K.K., Chu lin, N.A.: Applications of multi-height sensors data fusion and fault-tolerant Kalman filter in integrated navigation system of UAV. Elsevier Science Publishers B. V (2017)
6. Jayaram, S.: A new fast converging Kalman filter for sensor fault detection and isolation. *Sens. Rev.* **16**(3), 219–224 (2013)
7. Mahmoud, M.S., Khalid, H.M.: Fault estimation and monitoring with multi-sensor data fusion: an unscented Kalman filter approach. *Dyn. Syst.* (2011)
8. Wei, X., Verhaegen, M., Engelen, T.V.: Sensor fault detection and isolation for wind turbines based on subspace identification and Kalman filter techniques. *Int. J. Adapt. Control. Signal Process.* **24**(8), 687–707 (2010)
9. Simon, D., Chia, T.L.: Kalman filtering with state equality constraints. *IEEE Trans. Aerosp. Electr. Syst.* 128–136 (2002)
10. Wang, S., et al.: Decentralized optimization for a novel control structure of HVAC system. *Mathematical problems in engineering* (2016-5-9) **12**, 1–13 (2016)
11. Mendel, J.: Optimal filtering. *IEEE Trans. Autom. Control.* **25**(3), 615–616 (1980)
12. Jian, L.: Constrained information-based state estimation and fusion. Dissertation, (in Chinese), Shanghai Jiaotong University, China (2014)

DXF File Topological Information Extraction and Storage for Decentralized Distribution Network



Yuhan Zhang and Shiqiang Wang

Abstract In this paper, DXF (Data Exchange File) file is used to extract the topological information of a distribution network. The basic information is configured to each CPN (Computing Process Node) and the local topology is obtained through broadcast communication. According to the topological sorting algorithm of the adjacency list, the global topology is obtained. Taking a typical building distribution network as an example, the feasibility of the method is analyzed and demonstrated.

Keywords Data exchange file · Topological information configuration · Adjacency list

1 Introduction

The topological recognition of the power grid is the basis for the safety and energy-saving operation. Under the framework of the traditional intelligent system, a large number of sensors need to be installed to recognize the terminals of a distribution network, so as to obtain the global topological information. But as its higher hardware and maintenance cost, it is seldom used in practical engineering.

Decentralized intelligent system architecture, on the whole structure level to solve many problems in the application practice [1]. It can meet the needs of the development in a modern intelligent building distribution network.

The decentralized system recognizes the overall topology through the basic information configuration and the collaboration between CPNs. No longer dependent on the central computer station, all algorithms are coordinated in the CPN network. All CPNs are independent and equal to each other, and all kinds of work are completed in a coordinated manner.

Y. Zhang (✉) · S. Wang

College of Defense Engineering, Army Engineering University of PLA, Nanjing, China
e-mail: Hansel_Cheung@163.com

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_55

593

2 Extraction of AutoCAD Map Metadata Repository

The CAD diagram of the distribution network consists of several basic elements, and the set of data items contained in the picture element is called the map metadata repository. The common method of extracting the metadata repository is to use the corresponding instructions of AutoLISP. The disadvantage is that the extracted information for the CPN configuration is redundant, and needs to be preprocessed. But in this paper, DXF file is used for extracting map metadata repository [2]. DXF is a textual document, and it can meet that requirement of the development of the algorithm. The extraction procedure of AutoCAD metadata is: preprocessing CAD drawings, converting DWG (Drawing) format into DXF (Data Exchange File) format, and finally extracting the information of the graph by programming on VC++6.0 platform [3].

For example, the procedure of the algorithm for obtaining the linear map metadata repository is as follows:

- (1) Read the DXF file from top to bottom to determine whether the group code meets the requirements;
- (2) Determine whether the group code is 0, and whether the group value is LINE. Indicating that it has entered the LINE chart element and is ready to store data;
- (3) Determine whether the group code is 10, store its group value, and the group value is the starting X coordinate of the line;
- (4) Determine whether the group code is 20, store its group value, and the group value is the starting point Y coordinate of the line;
- (5) Determine whether the group code is 11, store its group value, and the group value is the terminal X coordinate of the line;
- (6) Determine whether the group code is 21, store its group value, and the group value is the end point Y coordinate of the line.

3 Topological Information Configuration and Adjacent Node Identification Rules

After the extraction of the map metadata repository, the basic topological information such as the location coordinates of the nodes and the number of branches is configured into the CPN [4]. The adjacent nodes are identified by the broadcast communication between CPNs and the local topology can be obtained.

3.1 CPN Topological Information Configuration

The topological information of CPN is the basis of topological identification. The CPN topological information configuration table is global, and its main contents

Table 1 CPN topological information configuration table

Class	Name	Type	Remark
Node information	Number	Int	The unique number of the node
	Attribute	Int	Node type
	Location	Float	The position of the node
	Logical layer	Int	The number of logical layers
Branch information	Number	Int	The unique number of the branch
	Quantity	Int	Number of branches
	Attribute	Int	Branch type
	Direction	Boolean	The direction of the branch

include CPN node information and branch information [5]. The CPN topological information configuration table is shown in Table 1.

Above topological information only configures the data of the node itself and does not need to configure the information of adjacent nodes. Such configuration mode has good flexibility and extensibility.

3.2 Identification Rules of Adjacent Nodes

When any node in the CPN network initiates a query instruction of the distribution network topological structure, the node broadcasts its own topological information. Other CPNs in the distribution network can receive this information and check the branch numbers in the topological information table. Judge the relative position between CPNs based on the direction of the branch route.

When the receiving node judges its own topological relationship with the broadcast node, the topological information is fed back to the broadcast node in a feedback-table manner. After a broadcast, the topological query initiator gets a local topology including the upstream and downstream of the adjacent node. General local topologies contain only broadcast nodes with one hop or two-hop nodes. If the local topology includes a two-hop node, the one hop node is needed to re-broadcast. And the topological information of the two-hop node can be acquired.

4 Storage Structure of Global Topology

Any CPN can get to the global topological information. The commonly used global topological storage structures are associated matrix method, adjacency matrix method, adjacency list method, etc.

The correlation matrix is one of the common topological storage structures, which is mainly used to represent the relationship between nodes and branches [6]. The rows and columns of the associated matrix correspond to nodes and branches in the construction distribution network.

The adjacency matrix method is similar to the associative matrix method. The adjacency matrix takes nodes as rows and columns respectively, and each element in the matrix is an array, representing the adjacent relation between nodes and the information of the edges between the connecting nodes.

The complicated distribution network uses the correlation matrix structure to be a great waste of storage space, because there are many “0” elements in the correlation matrix. So we consider another kind of storage structure: the adjacency list method.

In graph theory, the adjacency list method is a storage structure combining array and linked list. The adjacency list contains a vertex order table and a single linked list for each vertex. The vertices in the vertex order table correspond to the vertices in the graph, and the single linked list of each vertex contains the critical area of the node. Figure 2 is the adjacency list storage form of Fig. 1.

The local topological information of each node is stored by the adjacency list, and the global topology can be obtained by the topological sorting algorithm of the adjacency list. The global topology information provides data basis for further energy-saving control and electrical safety warning, and also provides the basis for troubleshooting and further maintenance.

Fig. 1 Topology diagram

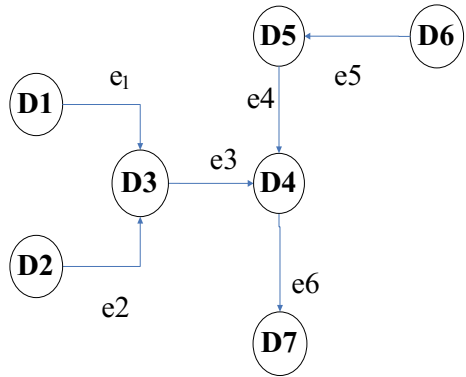
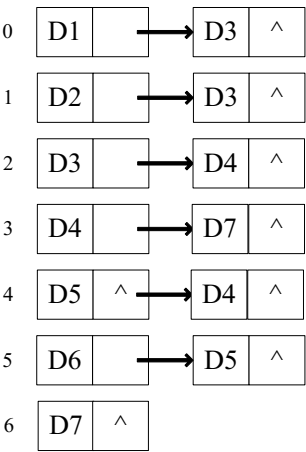


Fig. 2 Storage of adjacency list



5 Extracting and Storing the Topological Information of Typical Distribution Network

Take the typical building distribution network with UPS (Uninterruptible Power System) as an example, as shown in Fig. 3, the topological information is extracted and stored [7].

5.1 The Project Design of Typical Distribution Network

According to the above system, the CAD drawing of the distribution network is pre-processed and DXF files of the construction distribution network are obtained. Based on the above mentioned method, the location information and the branch information of each node in the distribution network are extracted. Then, the information is configured into the configuration table of the corresponding nodes, and each node begins to broadcast its own topology information at the end of configuration. Local topological relations are determined according to the feedback information of the adjacent nodes, and upstream and downstream relations with the adjacent nodes are judged. Then, based on the local topology, the logic layer of each node is marked by using depth-first search and setting broadcast permission. As shown in Fig. 3, D1 node is the logical “1”, D2, D3, D4, D5 node for logic layer “2”, D6, D7, D8, D9, D10 nodes for logic layer “3”, D11, D12, D13, D14, D15 nodes for logic layer “4”, D16, D17, D18, D19 and D20 and D21 node for logic layer “5”. The logical layer partitioning results are shown in Fig. 4.

The global topology can be obtained from any node in the distribution network. The node which initiates the query collects the local topological information of each node and converts it to the adjacency list method store.

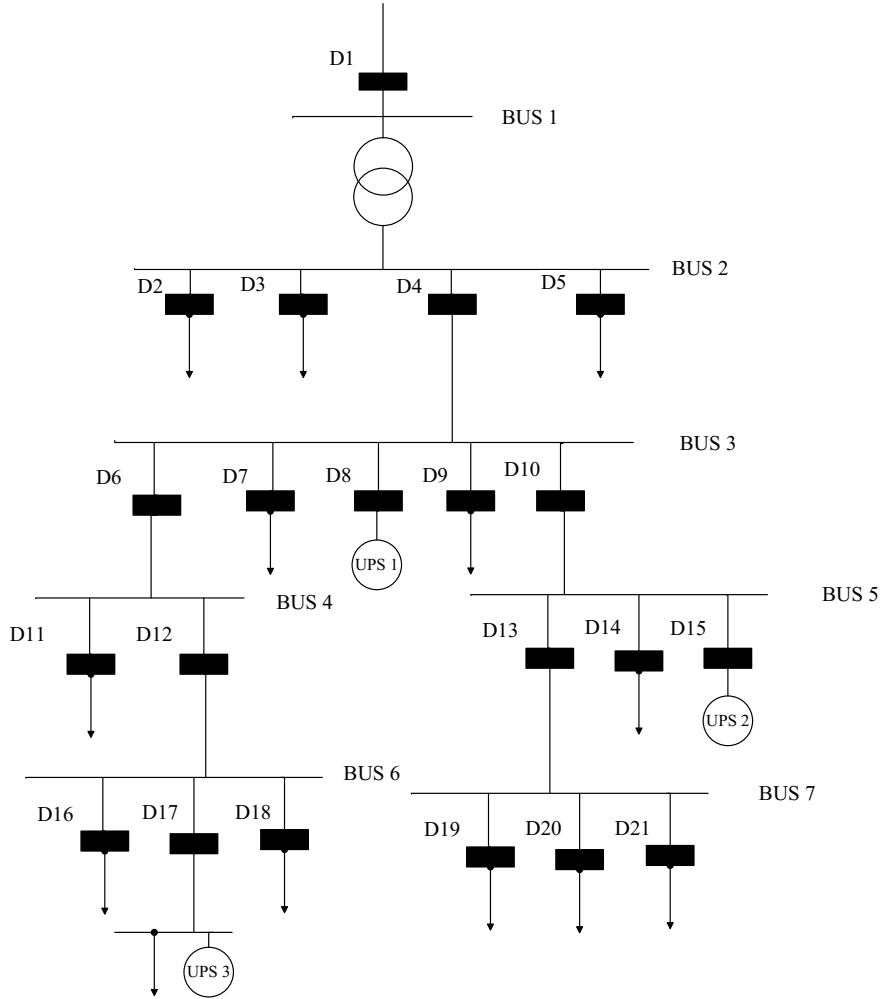


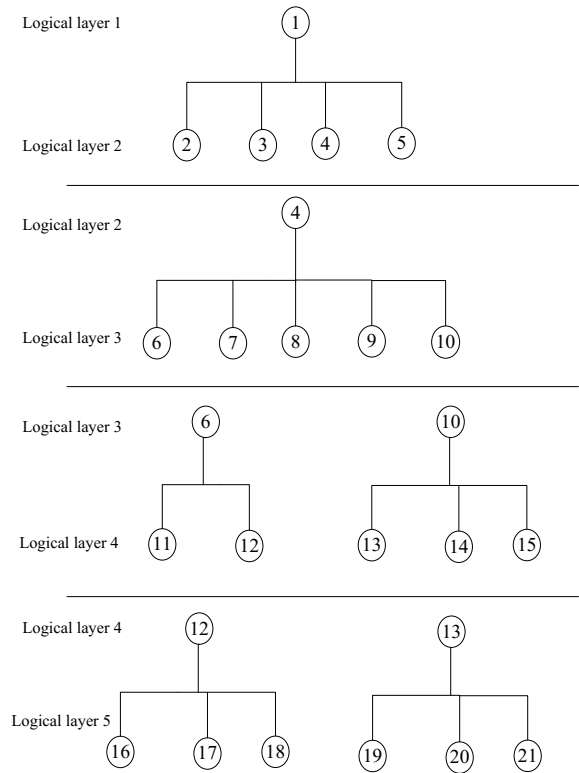
Fig. 3 Typical building distribution network

The usual topological sorting method is: Kahn algorithm and DFS (Depth-first Search) depth priority search algorithm [8]. According to the topological information adjacency list of the distribution network, it is sorted by Kahn algorithm to obtain the linear sequence of global topology.

The global topology is

D1	D2	D3	D5	D4	D7	D8	D9	D6	D12	D10	D14	D15	D11	D16	D17	D18	D13	D19	D20	D21
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Fig. 4 Logical layer partitioning



The corresponding logical layer is

1	2	2	2	2	3	3	3	3	4	3	4	4	4	5	5	5	4	5	5	5
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

5.2 Result Analysis

As shown in Table 2, the comparison of storage space for each storage structure. Table 2 shows that under the same conditions, compared to the adjacency matrix storage structure, space of correlation matrix storage structure is greatly reduced, decreased by 52%, significantly improve the utilization rate of storage space. Adjacency list method takes up the minimum space, reduces 91%, and also has good storage space utilization. Therefore, the adjacency list method has better practicability and scalability in topological information storage.

Table 2 Comparison of three storage structures in Fig. 3

Storage structure	Space usage
Correlation matrix	$21 * 20 = 420$
Adjacency matrix	$21 * 21 * 2 = 882$
Adjacency list	$5 + 1 + 1 + 6 + 1 + 3 + 1 + 1 + 1 + 4 + 1 + 4 + 4 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 41$

6 Conclusions

In this paper, the topological information extraction and storage method of distribution network based on DXF file is designed at the top level. A topological information configuration format of CPN nodes and identification rules of the adjacent nodes are proposed. Compared the three kinds of topological space, storage structure occupies in the same network topology model and node size, the method of adjacency table takes up minimal space.

Limited by capacity and time, the problems that need to be further studied and solved in the future: the global topology information extracted can be extended to other applications of distribution network, such as fault diagnosis of the distribution network, local optimization of the distribution network, etc.

Acknowledgements This work is financially supported by the National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques).

References

1. Dai, Y., Jiang, Z., et al.: A decentralized algorithm for optimal distribution in HAVC system. Build. Environ. **95**, 21–31 (2016)
2. Atia, M.: Classification and elimination of overlapped entities in DXF files. Ain Shams Eng. **5**(3), 851–860 (2014)
3. Yang, H., Yan, J.: DXF file identification with C# for CNC engraving machine system. Intell. Control Autom. **6**(1), 20–28 (2015)
4. Basu, A., Blanning, R.W.: Metagraphs and Their Applications (2007). https://doi.org/10.1007/978-0-387-37234-1_1
5. Luan, W., Peng, J., et al.: Distribution network topology error correction using smart meter data analytics. In: Proceedings of Power & Energy Society General Meeting, pp. 1–5 (2013)
6. Gaur, D., Shastri, A., Biswas, R.: Metagraph-based substructure pattern mining. In: International Conference on Advanced Computer Theory and Engineering, pp. 865–869 (2008)
7. Kochukov, O.: Model for technical-economic analysis of ENTSO-E and IPS/UPS power system optimal interconnection concepts. Sci. J. Riga Tech. Univ. Power Electr. Eng. **28**(1), 15–18 (2011)
8. David, J., Paul, H.: A dynamic topological sort algorithm for directed acyclic graphs. J. Exp. Algorithmics **11**, 1–7 (2006)

Research on Underground Device Operation and Maintenance Management System Based on BIMserver



Mengli Ding, Qiliang Yang, Jianchun Xing and Liqiang Xie

Abstract With the wide application of BIM, the technology of Operation and Maintenance (O&M) management is developing. By combining BIM with facility management, it can visually manage the model and highly integrate information of the model. As a common data standard of BIM, IFC (Industry Foundation Classes) standard can avoid the problems caused by the mismatch of information between different models. This paper proposes a method of BIM-based facility management and describes the key technical issues in the development of O&M systems. The O&M system leverages BIMserver and its public plug-in interface as a system framework to analyze IFC data. Through the study in this paper, a new method can be developed for future underground space O&M management.

Keywords BIM · Operation and maintenance (O&M) · BIM server · Underground engineering

1 Introduction

The life cycle of a building is usually divided into four phases: the planning and design phase, the construction phase, the Operation and Maintenance (O&M) management phase, and the maintenance and demolition phase. Among them, the O&M phase occupies more than 90% of the building's total life cycle and more than 80% of construction cost [1]. The O&M management determines whether buildings and device can perform their intended functions efficiently and accurately. With the continuous complication of the construction and functional requirements of underground engineering, the number and function of devices are increasing. At the same time, the expansion of underground space has increased the dispersion of devices in space, which has brought new challenges to the O&M management.

M. Ding · Q. Yang (✉) · J. Xing · L. Xie
College of Defense Engineering, Army Engineering University of PLA, Nanjing, China
e-mail: yql@893.com.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_56

The emergence of BIM has provided the possibility to construct a new visualized O&M platform. The BIM-based O&M management provides a new way of visual management. It also enables intuitive visualization of all aspects of building information and visual control of device operations based on the integration and display of information that uses BIM as an important carrier. The system improves the efficiency of O&M management and intelligent management.

The existing BIM-based management systems are mostly used in the construction stage or theoretical research. Few of them describe the complete process from research and development of O&M system. There is a lack of research on underground space O&M Management based on BIM.

This paper analyzes the existing problems of the O&M management, and proposes a method for constructing an underground engineering O&M management system. Through the research, analysis and other aspects were discussed on defining and extending EXPRESS from the geometry and attribute information of the device. And the key technologies for building the system such as how to analysis the IFC files is formulated.

2 Overview of O&M Management Based on BIM

2.1 BIM-Based Visual O&M Management

BIM-based visual O&M management system can realize information integration and sharing, achieving visual management of device. Compared with the existing O&M methods, the BIM-based global integrated management method not only improves the efficiency of design, construction, and O&M stages. It reduces the loss of information at all stages of delivery.

The visualization model can realize the accurate geometric representation of building components to achieve effective management of space. It can quickly generate an O&M management database to achieve faster and more effective information sharing to save time and cost. Using the model for simulation and debugging of building performance, the system can predict energy consumption performance and calculate the cost of the life cycle [2].

Applying BIM to underground engineering O&M management can realize the dynamic view of the whole life cycle information. The underground environmental conditions vary greatly with time, and it is necessary to monitor the safety factors such as temperature, humidity, seepage, and support force of the underground space structure in real time to improve risk response capabilities. At the same time, the system needs to have the ability to analyze the information and find faults and dangers in time like people.

Table 1 Device information management

Labeling	Category	Element
1	Basic information	Construction category, device number, device name, installation location, installation time, service life, manufacturer, rated voltage, rated power, fire rating
2	O&M information	Device operation status, maintenance information
3	O&M process information	Status records, maintenance records
4	Alarm information	Alarm value, sampling interval, monitoring value, upper threshold, lower threshold, measurement accuracy

2.2 O&M Information Management

The BIM model in the O&M platform can be used to describe the geometric information of buildings and device, and it generally will not be changed. Therefore, the information that needs to be managed is mainly the attribute information of the device and the information generated during the O&M. As shown in Table 1, the information of the devices that need to be managed in the underground project is divided into the following types:

Among them, the device information is divided into three categories: static attributes, dynamic attributes, and historical record information. A static attribute is the unchangeable attribute that comes with the device during production. The status information generated during the O&M of the dynamic attribute. In addition, the system also needs a database for recording O&M information. MySQL is used here as the information storage repository for the administrator to view.

3 O&M Platform Development Technology

3.1 O&M System Construction Principle

The development of the underground engineering O&M management system based on BIM can express the building space structure, the relative position relation between the device and the building, and the device running state intuitively in the three-dimensional form [3]. It can solve the problem that the model is not universal between the modeling software and the O&M system. And it provides a visual operation platform for device managers to add and query device information and real-time access. It is conducive to achieving efficient and reliable maintenance management [4].

As shown, the overall architecture of the system is divided into four layers: model layer, application layer, data layer, and [BIMserver.org](#) platform. First, the basic build-

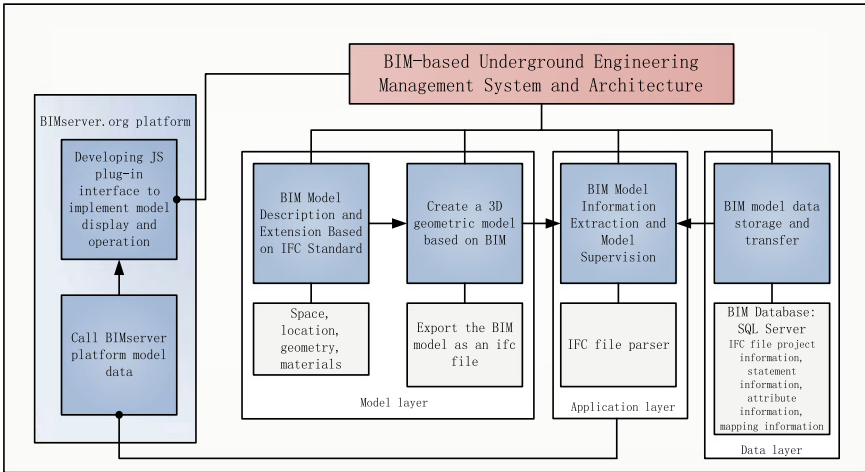


Fig. 1 O&M system principle

ing and device information of the model layer is constructed, and the basic geometry and attribute information are added to the 3D model by BIM modeling software. Thus the 3D geometric model based on BIM will be created. The attribute information added at this time is static information, and it cannot be changed with the change of the running state of the device. Export the BIM model into IFC format file and enter it into BIMserver.org platform, so that the platform can invoke the information of BIMserver platform to extract the model information and visualize the model supervision display. The manager analyzes and controls the O&M status according to the platform model information, and feedback the information to the platform model. The specific process is shown in Fig. 1.

3.2 The Display Technology of 3D Model

In this study, BIMserver.org is used as a platform for BIM model management, data exchange, and integration. BIMserver.org is an open BIM server platform written in Java language. It is the open source software that can be downloaded and installed on any computer. It can upload, display, operate, and manage the BIM model of IFC format directly through the browser. In addition, BIMserver.org has the characteristics of open source, and it can be developed two times according to the needs of users. Users can download the required resources in <https://github.com/opensourceBIM/BIMserver/releases> [5] and run it in Java environment. The operation environment of jar package requires Windows, JDK7 and above, tomcat7.056 and above. Click “Launch Webbrowser” into the web side, look at its page in the browser, and register to create new users and download plug-ins (including bimserversapi, bimsurfer,

bimviews, binaryserializers, ifcengine, ifcopenshellplugin, ifcplugins). IFC related plug-ins are installed on the local file installation mode after the system starts, and you can create a server that can use IFC data locally.

Through the web side of BIMserver, you can upload and download IFC files, 3D model display, query and other functions. Users can upload IFC data through a browser. BIMserver.org parses the IFC data and threejs rendering through IFCOpenShellEngine, and stores the data in the file in the underlying database to display the BIM model in the browser. The BIMviews integrated in BIMserver can provide BIM model browsing interface. Users can perform simple scaling, translation, rotation and other operations on the BIM model. In addition, you can view attribute information such as component type, cross section area, and so on.

The core function of BIMserver software is to understand the IFC structure and support users to use the BIM format in IFC format.

When developing an O&M management platform based on BIM, you can directly call BIMserver.org's BIM model uploading, browsing and simple operation functions through the JS plug-in interface. Before importing (check in) IFC model, you need to set the default rendering mode to ifcOpenShellEnginePlugin, as shown below, otherwise ifcengine related errors will be reported.

BIMserver based O&M management system development mainly has two aspects: the original plug-in function optimization and development of new functional interface. By modifying the BIMviews code, the display function of the model can be optimized. The source code of plug-in development is compiled using IDEA as the debugging environment, importing bimserver source code and using Maven architecture. Because of networking and other reasons, the code that is cloned through git often makes the plug-in version unsuccessful, so the source code of the corresponding version is downloaded and the remote debugging port is configured.

4 IFC-Based Entity Expression and Analysis

4.1 *The Analysis of IFC Files*

As IFC standard is a common data exchange standard in the BIM field and can describe information such as geometry and attributes of building entities. As shown in Fig. 2, the entity IfcRoot is the basis for all entities that have a unique identifier (GlobalId) and their information is stored, and is an abstract base entity. IfcRoot is defined and described by three IFC abstract classes and their derived entities. IfcObjectDefinition is used to define the object class of the object, IfcRelationship is used to implement the abstract class of the spatial. The entity framework of the shared structure is further defined from the entity type. The framework layer is to implement the description of the shared product in the entire construction industry. Subsequent derivation results in the formation of entity types in the domain layer for

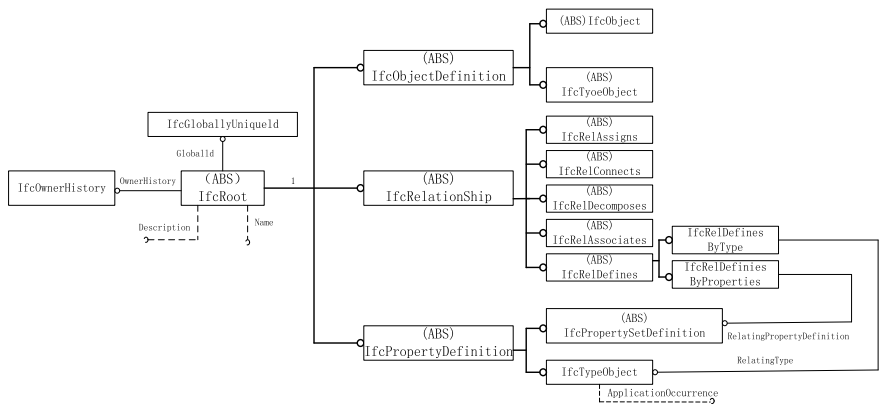


Fig. 2 Entity attributes and inheritance

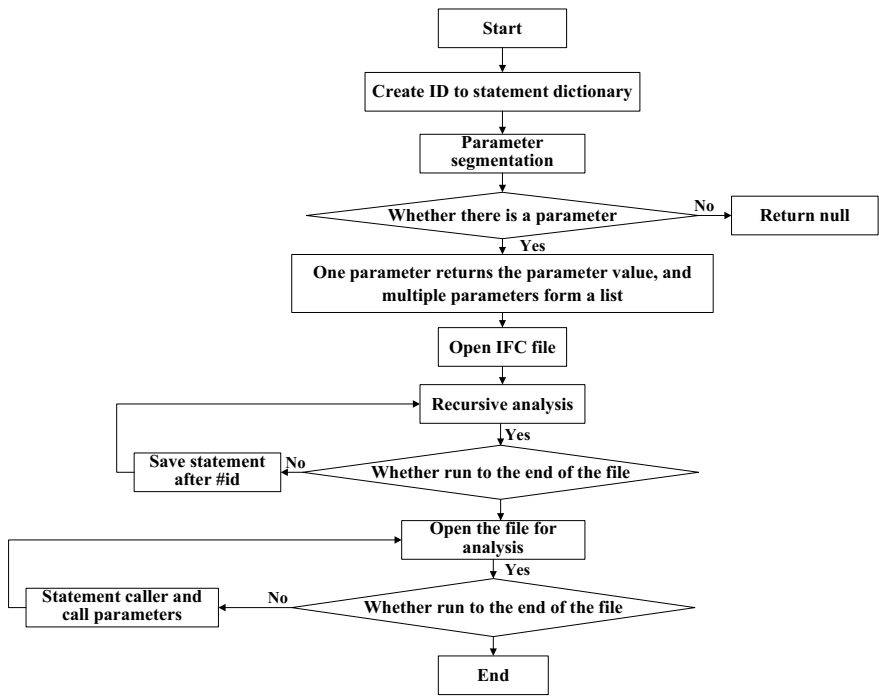


Fig. 3 IFC file processing

the description of specific fields and relationships in different fields and professions in the construction industry [6].

The IFC file is written by EXPRESS language. It is hard to be read and understood by managers. To extract the model information in the O&M platform, the IFC file

needs to be processed. First, we should extract the entity reference relationship, such as IfcDoor inherits the IfcBuildingElement content (attributes, constraints, etc.). In accordance with this inheritance relationship, it is seven layers of inheritance relation from the top IfcRoot all the way to IfcDoor. As shown in Fig. 3, the information will be extracted to form the complete IfcDoor information set.

The output of the BIM model is an IFC file. Here, IfcDoor is used as an example. The correspondence between the IFC file and the IFC standard is as follows. It mainly contains four parts: the Globally Unique Identifier GlobalID in the IfcRoot, OwnHistory, Name, and Description; IfcObject Object type ObjectType; IfcProduct object position information ObjectPlacement, representing information Representation and tag information Tag.

4.2 O&M System Data Transfer

The O&M platform uses the IFC file as a medium for model and information transmission.

Therefore, the IFC file parsing module and the BIM database need to be added as data management layer which used for information integration and the conversion between IFC file and model. The system's DFD (data flow), as shown in Fig. 4, is divided into four parts. The model layer is responsible for information generation and processing, and is the initial source of data. It can form IFC files containing basic information such as geometric information and attribute information; The data management layer has the function of data management. The O&M management platform is the most central part, with information flow, interactive feedback, model

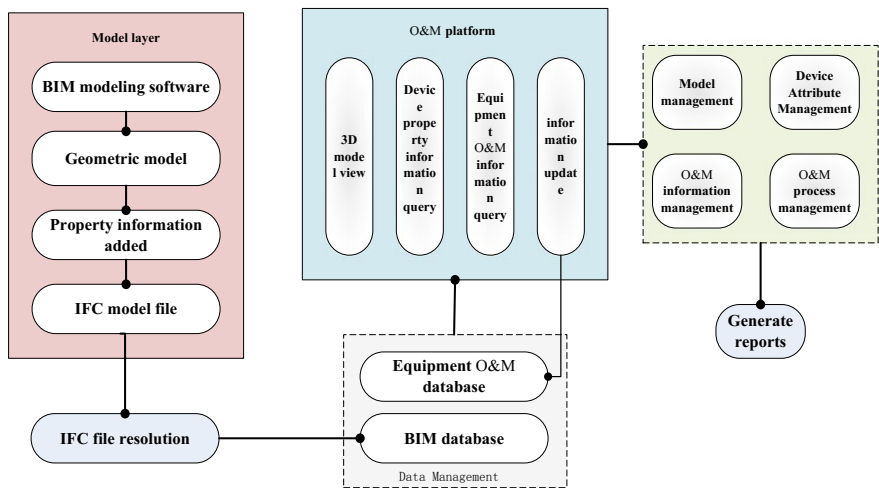


Fig. 4 Information flow chart

visualization, and the overall control of the system are only recorded in the device operation and maintenance database, but also displayed in the platform in real time; Report management can extract related information from the platform, implement information query and generate report display results.

In summary, after the IFC model file information is imported into the system, a closed data cycle will be formed between the model layer, data management layer, O&M management platform, and report management module to ensure the authenticity and reliability of the BIM model data.

5 Conclusions

In this paper, it combines the basic concepts of BIM technology and O&M management system, establishes the design and development framework of underground engineering O&M management system based on Java and BIMserver, and provides key technical points and technical routes for the development of visualized O&M platform. In the next step of development, the prototype system is intended to be applied to the actual project. It also needs to combine the device location technology and develop the system's information transmission protocol to form an interactive O&M platform with the integration of information and physics.

Acknowledgements The project was partially funded by the National Key Research and Development Project, China, with No: 2017YFC0704100 “New Building Intelligent System Platform Technology”.

References

1. Vanlande, R., Nicolle, C., Cruz, C.: IFC and building lifecycle management. *Autom. Constr.* **18**(1), 70–78 (2009)
2. IFC 2x Edition 3
3. Kensek, K.: BIM guidelines inform facilities management databases: a case study over time. *Buildings* **5**(3), 899–916 (2015)
4. Codinhoto, R.: Challenges in the implementation of BIM for FM—case manchester town hall complex. In: *Proceedings of Computing in Civil and Building Engineering 2014*, pp. 572–577 (2014)
5. <https://github.com/opensourceBIM/BIMserver/releases>
6. Pärn, E.A., Edwards, D.J., Sing, M.C.P.: The building information modelling trajectory in facilities management: a review. *Autom. Constr.* **75**, 45–55 (2017)

A Fully Distributed Genetic Algorithm for Global Optimization of HVAC Systems



Shiqiang Wang, Jianchun Xing and Juelong Li

Abstract To solve the high labor and maintenance cost problems in actual engineering, a decentralized heating, ventilation, and air-conditioning (HVAC) system is configured following its physical layout. In a decentralized HVAC control system, each of the updated smart equipment can communicate with the adjacent nodes collaboratively to fulfill the load requirement. Furthermore, to achieve the global optimal operation of an HVAC system, a fully distributed constrained optimization is formulated. In this paper, a fully distributed genetic algorithm (GA) is developed to solve the proposed constrained optimization. The proposed method is confirmed to be effective to realize the global optimization of HVAC system through simulation study.

Keywords HVAC system · Global optimization · Self-Organizing · Decentralized genetic algorithm

1 Introduction

A typical heating, ventilating, and air-conditioning (HVAC) system is a multi-variable, nonlinear, and large-scale system. Each system comprises a set of interconnected subsystems, including the heating and cooling plants, the ventilation system, and one or more zones served by the terminal units of the ventilation system. Each subsystem consists of several components, such as sensors, mechanical and electrical actuators, and controllers [1]. The energy consumption of air-conditioning systems accounts for a large proportion of the total energy consumption in buildings, thus it is important to improve the energy efficiency of such systems [2]. A small increase in system operating efficiency can result in substantial energy savings. Much research has been done on HVAC optimal control, including simulation and experiment.

S. Wang · J. Xing (✉) · J. Li

School of Defense Engineering, Army Engineering University of PLA, Nanjing, Jiangsu, China
e-mail: xjc@893.com.cn

© Springer Nature Singapore Pte Ltd. 2019

Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_57

However, the deployment of advanced controls in buildings has been progressing very slowly [3].

Currently, in the practical HVAC system operation engineering, the typical variables (such as supply chilled water temperature, return cooling temperature, and supply air temperature) are always set as constant. But the automatic control provisions and even the computerized building management systems have been mainly provided for full load settings, and not necessarily part load situations [1]. However, the cooling load is always changing in the building, therefore, HVAC system cannot operate in the optimal condition such that extra energy is consumed. There have been many research works reported for improving either individual component or partial system efficiencies. But the above practical and research methods either fix some control variables or overlook some control devices to reduce the complexity of the HVAC optimization problem, not considering the global optimization of total HVAC system [4–6]. Obviously, the simplified methods cannot achieve the global optimal operation of HVAC system.

Several researchers have devoted to a global HVAC system optimization for minimizing the energy consumption. Lu et al. proposed a global optimization for overall HVAC systems in [7, 8]. Besides, a number of studies focused on the real-time optimization algorithm of HVAC systems such as classic linear/quadratic programming [9], or gradient-based iterative methods [10], mixed-integer linear programming (MILP) [11]. Recently, to address the complexity of large-scale HVAC systems, swarm intelligence and evolutionary optimization algorithms have been developed, such as genetic algorithms [12], particle swarm optimization (PSO) algorithm [13, 14], radial basis function neural network (RBFNN) [15], expert control [16], and other computational intelligence techniques [17]. Besides, a modified genetic algorithm is devised for a mixed-integer nonlinear constraint optimization problem to find optimal set points to minimize the overall system energy consumption [18].

However, an HVAC system is spatially and temporally distributed, with a large-scale, complex, and nonlinear control system. The current centralized control architecture for HVAC systems has several deficiencies during construction and operation. First, network construction is a complicated and time-consuming job because a considerable amount of secondary development, such as configuration and commissioning, is necessary on-site. In addition, the control model and algorithm are written into the centralized controller, which needs to be reprogrammed from case to case because the system configuration or device type changes. Thus, the design and deployment of a centralized control system requires tedious and time-consuming manual calibration procedure.

Besides, during operation process, a centralized method needs information transmission to the supervising computer for control and alarm processes, which can cause grievous link congestion and operational lag. To date, most HVAC control systems are not fully operational and only decorative.

In some cases, the conventional centralized structure has become a barrier for the control of HVAC systems; hence, the desired HVAC industry requires a new platform. This new platform should be open, flexible, and plug and play to enable easy installation with low configuration cost.

Motivated by this observation, a novel decentralized HVAC system control structure and optimization method are proposed to meet theoretical challenges and practical needs. The rest of this paper is organized as follows: in Sect. 2, the HVAC system model is presented and a decentralized optimization is formulated. Then a decentralized genetic algorithm is designed to solve the novel optimization problem in Sect. 3. Section 4 focuses on performance assessment through simulations and experiments. Conclusions are drawn in Sect. 5.

2 System Model and Problem Definition

For the global optimization of HVAC systems, the objective is to provide all the conditioned spaces with the required cooling loads at minimal energy consumption. To investigate the performance of the proposed optimization strategy, a simplified but representative HVAC system was constructed. The system consists of one chiller, one chilled pump, one cooling pump, one cooling tower, one air-handling unit (AHUs), and one conditioned zone as shown in Fig. 1a. The cooling load Q_{cool} can be measured during one sample period and treated as constants for the optimization process.

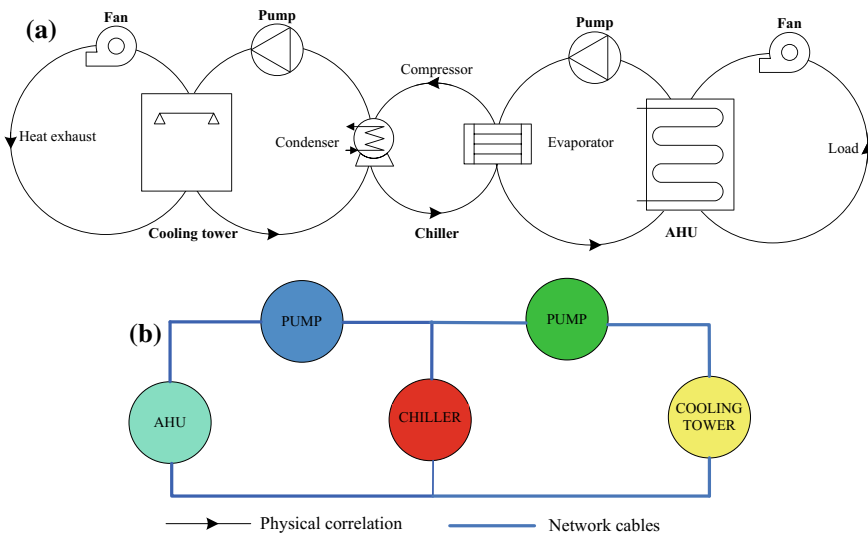


Fig. 1 HVAC system structure and its decentralized control system. **a** Typical HVAC system. **b** Decentralized HVAC control system

In the decentralized control structure, an HVAC system is divided into cooling tower, chiller, pump, and AHU based on physical distance and coupling relationship. Then, a microprocessor unit is embedded in each device such that it is updated as a smart agent. And, each smart agent is connected using network cables if they are involved in the same physical laws equation. Accordingly, the communication topology can be obtained as Fig. 1b based on physical correlation such as heat exchange equation. These agents can work independently and cooperate together to fulfill the global optimizing task on the HVAC system. It can be seen that the decentralized control system topology is closely mapped from the actual system layout and it makes the controller setup process more straightforward.

Then, the energy consumption P_{total} of the HVAC system can be got as

$$\min P_{total} = P_{chiller} + P_{CHWpump} + P_{CWpump} + P_{tfan} + P_{cfan}, \quad (1)$$

where $P_{chiller}$, $P_{CHWpump}$, P_{CWpump} , P_{tfan} , and P_{cfan} denotes the energy consumption of chillers, chilled pump, cooling pump, cooling tower, and AHU, respectively. And, the specific cost functions of each device are described as follows.

2.1 Cost Function

(1) Chiller

Chiller is the key device which connects cooling and chilled circulating water systems. And, its energy consumption is affected by evaporating temperature, condensing temperature, and cooling load. Thus, its energy consumption is a function of supply chilled water temperature, return cooling water temperature, and cooling load as

$$P_{chiller} = a_0 + a_1(T_{CWS} - T_{CHWS}) + a_2(T_{CWS} - T_{CHWS})^2 + a_3(T_{CWS} - T_{CHWS})Q_{cool} + a_4Q_{cool} + a_5Q_{cool}^2, \quad (2)$$

where T_{CWS} is the entering cooling water temperature, °C; T_{CHWS} is the supply chilled water temperature, °C; a_0 , a_1 , a_2 , a_3 , a_4 , and a_5 is the model parameters. And, the parameters are acquired from the manufacture catalog data or by on-site testing data.

(2) Chilled pump

And, the power consumptions of variable-speed drive (VSD)-driven pumps are influenced by two parameters: water flow rates and pressure difference. To reduce the energy consumption on pipe, the valve could keep fully open condition on the AHU cooling coils of the simplified HVAC system in Fig. 1a. Thus, the pressure difference on pump only involves water flow rate variable because the resistance on pipe network is determined. The power can be modeled as a cubic polynomial of the flow

rate whose coefficients were obtained from regression of performance data. Thus, the energy consumption of VSD chilled pumps can be formulated as [3]

$$P_{CHWpump} = b_0 + b_1 m_{CHW} + b_2 m_{CHW}^2, \quad (3)$$

where m_{CHW} is the supply chilled water flow rate, m^3/h ; b_0 , b_1 , and b_2 are the parameters.

(3) Cooling pump

A variable-speed cooling water pump is used to deliver a variable-rate condensing water flow in the experimental platform. As the above analysis, the cost function of the cooling pump is

$$P_{CWpump} = c_0 + c_1 m_{CW} + c_2 m_{CW}^2, \quad (4)$$

where m_{CW} is the supply cooling water flow rate, m^3/h ; c_0 , c_1 , and c_2 are the parameters.

(4) AHU

VSD fan is the main energy consumption component in an AHU. And, the power consumptions of VSD fans are influenced by two parameters: air flow rates and pressure difference between the duct inlets and outlets. To reduce the energy consumption on duct, the valve could keep fully open condition in the terminal in Fig. 1a. Thus, the pressure difference on fan only involves air flow rate variable because the resistance on duct network is determined. And its cost function can also be described as

$$P_{cfan} = d_0 + d_1 m_{SA} + d_2 m_{SA}^2, \quad (5)$$

where m_{SA} is the AHU air flow rate, m^3/h ; d_0 , d_1 , and d_2 are the parameters.

(5) Cooling tower

Cooling tower is responsible for transferring the cooling load from cooling water to atmosphere. The cooling water exchanges heat with outdoor air by fans such that the cooling water temperature is decreased. Thus, fan is also the major energy consumption component in the cooling tower. And, its energy consumption of cooling tower fan can be formulated as the function of air flow rate as

$$P_{tfan} = e_0 + e_1 m_a + e_2 m_a^2, \quad (6)$$

where m_a is the cooling tower air flow rate, m^3/h ; e_0 , e_1 , and e_2 are the parameters.

2.2 Constraints Establishment

The energy consumption function on an HVAC system involves with multiple variables. Any change of these independent variables will influence the efficiency of the whole HVAC system, and they are constrained by the physical relationships.

- (1) The interaction between AHU and chilled pump can be described by the heat exchange model on the cooling coils. And, the heat transfer model of AHU cooling coils is

$$Q = \frac{p_0 m_{SA}^q}{1 + p_1 \left(\frac{m_{SA}}{m_{CHW}} \right)^q} (T_{MA} - T_{CHWS}), \quad (7)$$

where T_{MA} is the mixed air temperature. This model does not need iterative computations, which makes it suitable for real-time implementation. In this model, no geometric data of coils are required and only three empirical parameters (p_0 , p_1 , and q) need to be identified from manufacture catalog data or experimental data.

- (2) The interaction between cooling tower and cooling pump can be described by the heat exchange model as

$$Q = \frac{r_0 m_{CW}^t}{1 + r_1 \left(\frac{m_{CW}}{m_a} \right)^t} (T_{CWR} - T_{wb}), \quad (8)$$

where T_{wb} is the wet bulb temperature outdoor. The parameters (r_0 , r_1 , and t) are identified from manufacture catalog data or experimental data.

The above performance function including the specific performance parameters in cost function and the related constraints can be written into the corresponding decentralized controller chip by the manufacturer who knows the equipment's performance as it leaves the factory.

For purpose of making the parallel locally coupled devices work under this structure, a feasible algorithm is necessary. Thus, the following paragraph mainly focuses on the study and analysis of a decentralized algorithm.

3 Decentralized Optimization Algorithm Based on Log-Linear Model

For the decentralized optimization problem described in (2)–(6), a penalty function method is considered. Generally, the cost function of each smart device is augmented with a penalty term associated with its constraints can be denoted as (9) for device i ($i = 1, 2, \dots, N$).

$$\min J_i(x_i) = f_i(x_i) + P_i(x_i), \tag{9}$$

where $f_i(x_i)$ denotes the energy consumption of each device, variable x_i denotes the decision variable on device i , and the local penalty function, P_i , penalizes the violation of the nonlinear constraints that x_i is associated with. Then, (9) means that each smart device i should select a certain strategy from its strategy space to satisfy the corresponding constraints with the least energy consumption.

This feature makes genetic algorithms (GAs), as a typical swarm intelligence algorithm is able to solve the complicated problem, global optimization for the overall HVAC systems.

3.1 Algorithm Design

However, the parallelism of the current GA and other swarm intelligence algorithms lie in its large number of populations, that is, all the populations, ants, and fishes in these algorithms involve identical global chromosomes. Thus, learning from the nature phenomenon, a decentralized GA (DGA) is proposed for the decentralized HVAC optimization system in this paper. In this so-called decentralized algorithm, each smart device is viewed as an individual can cooperate with adjacent individuals to realize the global optimal operation of HVAC system. And, the actions of mutation, crossover, and selection operations are addressed on each agent to explore the optimal solution independently. Different from the traditional GA, a local exchange operator is designed in the DGA. And the local exchange operator is described as: each individual can intersect with adjacent nodes on coupling variables in its local chromosomes, and product new chromosomes as shown in Fig. 2.

The detailed algorithm description is shown in Fig. 3. As seen from Fig. 3, DGA is different from traditional evolutionary or swarm intelligent algorithms, where the populations of each smart agent only cover its own related variables and cooperate with its coupled agent. And this process is being done simultaneously at each device agent. Further, the optimality of DGA needs to be explained.

Fig. 2 Exchange operator between adjacent agents

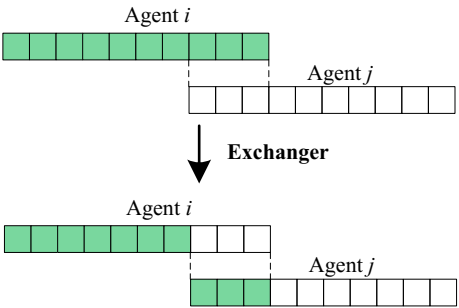
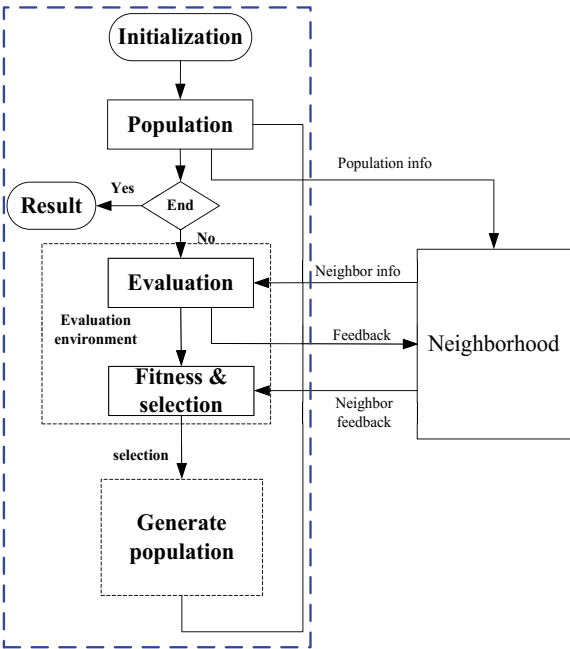


Fig. 3 Flowchart of DGA



4 Application to HVAC Systems

4.1 Experimental Setup

To investigate the performance of the proposed optimization strategy, the simplified but typical HVAC system is shown in Fig. 1 was applied in this section. Some simulations are used to show the advantages of the proposed decentralized method. And the device parameters are listed in Table 1.

The DGA method is coded in MATLAB (R2016b). For comparison, the current centralized GA is studied and illustrated here. In addition, several traditional methods, such as fixed T_{CHWS} control and T_{CWS} setpoint control, are also presented. And the supply chilled water temperature was 7 °C and the return cooling water temperature was 25 °C in the simulation.

Table 1 Device parameters

Devices	Parameters						Control variable
Chiller	a_0	a_1	a_2	a_3	a_4	a_5	$7 \leq T_{chws} \leq 12$
	183.48	-0.29	-0.12	0.016	0.11	1.95e-5	$25 \leq T_{cws} \leq 32$
Chilled pump	b_1	b_2		b_3			$10 \leq m_{CHW} \leq$
	8.80	0.02		0.02			40
Cooling pump	c_0	c_1		c_2			$30 \leq m_{CW} \leq 60$
	15.56	0.12		0.01			
AHU	d_0	d_1		d_2			$40 \leq m_{SA} \leq 80$
	1.51e-3	-7.94e-2		0.79			
Cooling tower	e_0	e_1		e_2			$60 \leq m_a \leq 150$
	9.41e-3	0.14		2.10			
Cooling coils (AHU)	p_0	p_1		q			
	2.65	0.52		0.77			
Cooling coils (cooling tower)	r_0	r_1		t			
	2.63	1.12		1.11			

4.2 Results and Analysis

Assume the cooling load is 400 kW one sample period. Then, DGA can be downloaded to decentralized microprocessor chip in each smart device, and then each smart device can be active. And every smart device can communicate and cooperate with other smart device nodes to find an optimal operating point at a given cooling load. Then, the optimization results are also illustrated in Fig. 4.

Further, the energy consumption in different methods is shown in Fig. 5. As seen from Fig. 5, the power cost can be reduced by approximately 9% comparing to the traditional method. And the proposed decentralized method can achieve the same performance as in a centralized manner. However, the decentralized method is more practical from the engineering perspective.

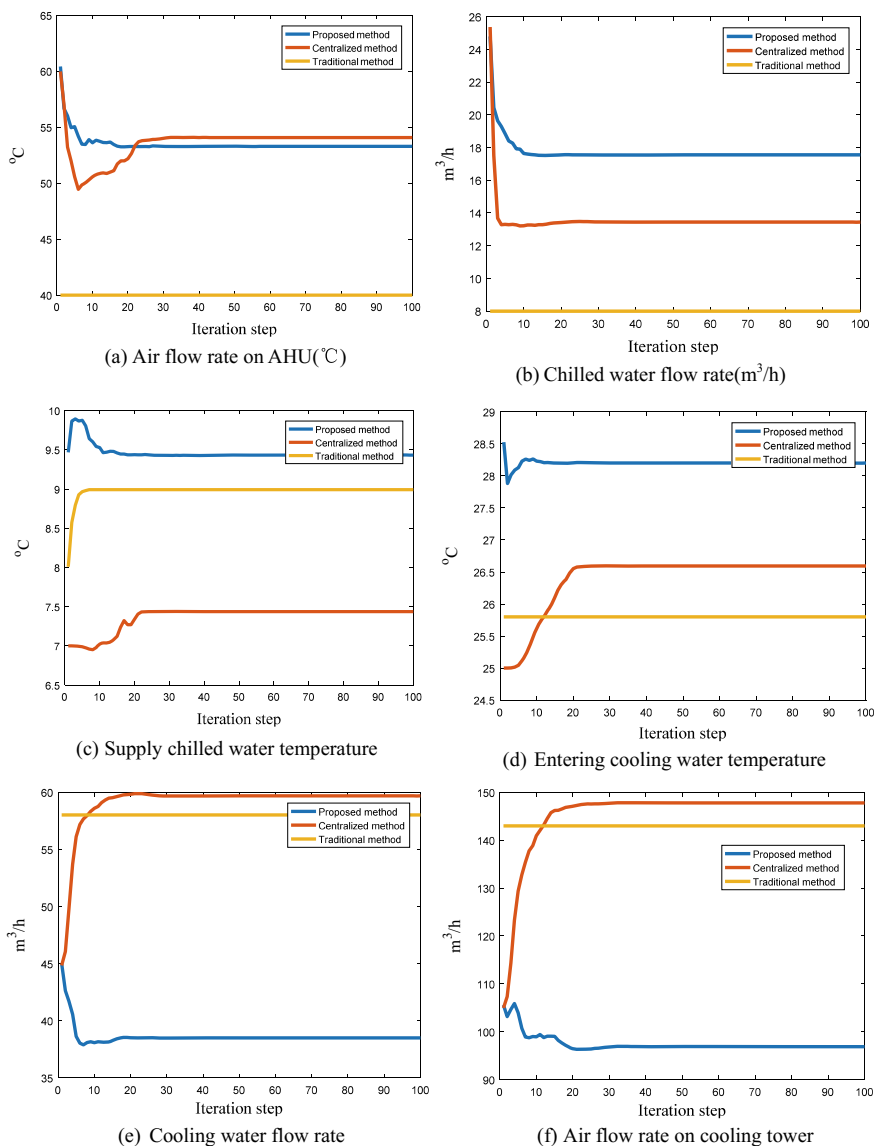
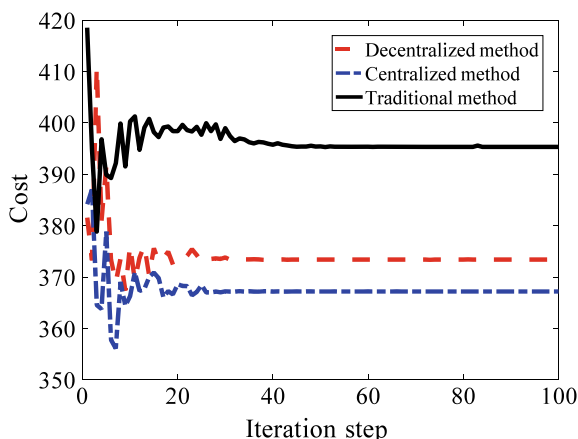


Fig. 4 Iterative process of decision variables on each device. **a** Air flow rate on AHU (°C). **b** Chilled water flow rate (m³/h). **c** Supply chilled water temperature. **d** Entering cooling water temperature. **e** Cooling water flow rate. **f** Air flow rate on cooling tower

Fig. 5 Iterative process of energy in decentralized method



5 Conclusions

In this paper, a general decentralized methodology is proposed for optimizing the HVAC system control to deal with the actual engineering issues. A novel optimization model for the HVAC system is established, a fully distributed genetic algorithm is explored to achieve the global optimal operation of HVAC system, the convergence of the novel method is theoretically analyzed, and the proposed method is confirmed effective for realizing the global optimal operation of HVAC system through simulation study.

Acknowledgements This work is supported by the National Key Research and Development Project of China No. 2017YFC0704100 (entitled New Generation Intelligent Building Platform Techniques).

References

1. Fong, K.F., Hanby, V.I., Chow, T.T.: HVAC system optimization for energy management by evolutionary programming. *Energy Build.* **38**(3), 220–231 (2006)
2. Pérez-Lombard, L., Ortiz, J., Pout, C.: A review on buildings energy consumption information. *Energy Build.* **40**(3), 394–398 (2014)
3. Cai, J., Kim, D., Jaramillo, R., et al.: A general multi-agent control approach for building energy system optimization. *Energy Build.* **127**, 337–351 (2016)
4. Marini, D.: Optimization of HVAC systems for distributed generation as a function of different types of heat sources and climatic conditions. *Appl. Energy* **102**(2), 813–826 (2013)
5. Wang, S., Burnett, J.: Online adaptive control for optimizing variable-speed pumps of indirect water-cooled chilling systems. *Appl. Therm. Eng.* **21**(11), 1083–1103 (2001)
6. Fan, B., Jin, X., Du, Z.: Optimal control strategies for multi-chiller system based on probability density distribution of cooling load ratio. *Energy Build.* **43**(10), 2813–2821 (2011)
7. Lu, L., Cai, W., Chai, Y.S., et al.: Global optimization for overall HVAC systems—part I problem formulation and analysis. *Energy Convers. Manag.* **46**(7), 999–1014 (2005)

8. Lu, L., Cai, W., Soh, Y.C., et al.: Global optimization for overall HVAC systems—part II problem solution and simulations. *Energy Convers. Manag.* **46**(7–8), 1015–1028 (2005)
9. Sun, J., Reddy, A.: Optimal control of building HVAC&R systems using complete simulation-based sequential quadratic programming (CSB-SQP). *Build. Environ.* **40**(5), 657–669 (2005)
10. Ahn, B.C., Mitchell, J.W.: Optimal control development for chilled water plants using a quadratic representation. *Energy Build.* **33**(4), 371–378 (2001)
11. Dullinger, C., Struckl, W., Kozek, M., et al.: A general approach for mixed-integer predictive control of HVAC systems using MILP. *Appl. Therm. Eng.* (2017)
12. Seo, J., Ooka, R., Kim, J.T., et al.: Optimization of the HVAC system design to minimize primary energy demand. *Energy Build.* **76**(2), 102–108 (2014)
13. Kusiak, A., Li, M., Tang, F.: Modeling and optimization of HVAC energy consumption. *Appl. Energy* **87**(10), 3092–3102 (2010)
14. He, X., Zhang, Z., Kusiak, A.: Performance optimization of HVAC systems with computational intelligence algorithms. *Energy Build.* **81**, 371–380 (2014)
15. Attaran, S.M., Yusof, R., Selamat, H.: A novel optimization algorithm based on epsilon constraint-RBF neural network for tuning PID controller in decoupled HVAC system. *Appl. Therm. Eng.* **99**, 613–624 (2016)
16. Ling, K.V., Dexter, A.L.: Expert control of air-conditioning plant. *Autom.* **30**, 761–773 (1994)
17. Ahmad, M.W., Mourshed, M., Yuce, B., et al.: Computational intelligence techniques for HVAC systems: a review. *Build. Simul.* **9**(4), 359–398 (2016)
18. Lu, L., Wenjian, C., Lihua, X., Shujiang, L., Yeng, C.S.: HVAC system optimization—in-building section. *Energy Build.* **37**(1), 11–22 (2005)

Open-Neutral Fault Detection in Underground Space Based on Genetic Support Vector Machine



Xuechen Zhao, Ping Wang and Jianchun Xing

Abstract When an open-neutral fault occurs in an underground space low-voltage distribution system, the neutral-point voltage will offset, which will make a lot of equipment unable to work normally. In order to solve such problem, the changes of third harmonic current and neutral-point offset voltage were studied and analyzed, the parameters of Support Vector Machine (SVM) was optimized by using a genetic algorithm. The output of SVM whether normal or fault state can be distinguished by taking the variation of harmonic current and neutral offset voltage as the input of SVM. The experimental result shows that the method of optimizing the parameters of SVM based on the genetic algorithm can effectively detect the fault.

Keywords Open-neutral fault · Fault detect · Harmonic · Neutral-point offset voltage · SVM

1 Introduction

In low-voltage distribution system, three-phase four-wire is widely used, and neutral line holds neutral-point potential at zero. When open-neutral fault occurs, the neutral-point potential will offset, and the load will have varying degrees of over-voltage, damaging electrical equipments, and even causing building fire. Therefore, it is necessary to accurately judge whether the neutral line breaks down.

X. Zhao (✉) · P. Wang · J. Xing
College of Defense Engineering, Army Engineering University of PLA,
Nanjing, Jiangsu 210007, China
e-mail: 13815442225@163.com

P. Wang
e-mail: wp@893.com.cn

J. Xing
e-mail: xjc@893.com.cn

© Springer Nature Singapore Pte Ltd. 2019
Q. Fang et al. (eds.), *Advancements in Smart City and Intelligent Building*,
Advances in Intelligent Systems and Computing 890,
https://doi.org/10.1007/978-981-13-6733-5_58

The existing open-neutral fault detection methods are divided into two categories: parameters monitoring and external devices, the type of parameters monitoring is divided into harmonic monitoring and voltage monitoring. In the literature [1], a detection method of monitoring phase voltage is proposed, a detection method of monitoring phase voltage is proposed, which can detect the open-neutral fault by determining whether the phase voltage is overvoltage, but this method is only applicable to single-phase load. Literature [2] and [3] use the third harmonic current change to judge whether the open-neutral fault occurs, by analyzing fluctuation law of third harmonic current and fundamental current can obtain criterion to distinguish between normal and fault state, and through the simulation test to verify the effectiveness of the proposed criterion, but when the fault occurs at the end of the line, the proposed criterion cannot guarantee high accuracy rate. In the literature [4], a detection method with the neutral line current criterion is proposed, which can be misjudged when the three-phase load imbalance is high, and the accuracy rate needs to be further improved. Through introducing the adaptive coefficient, literature [5] improves the protection criterion of the neutral-point offset voltage and enhances the sensitivity of the traditional protection methods. The main idea of the external devices detection method is to add other signals to diagnose the open-neutral fault. In the literature [6], the constant current source is added, the open-neutral fault can be detected by monitoring load voltage.

1.1 Harmonic Characteristics

During normal operation of low-voltage distribution system, line impedance is far less than the load impedance, harmonic by via neutral line flow system side, but when an open-neutral fault occurs, and harmonic current can only flow to the load side, the harmonic current at the beginning of the line will only be the load harmonic current before the break position.

When an open-neutral fault occurs, the neutral-point potential is not offset if the three-phase load is in a state of balance, but the third harmonic current in three-phase line will mutate, as shown in Fig. 1, when the time is 0.5 s the open-neutral fault occurs. Most low-voltage distribution systems are in the state of unbalance now harmonic current will also mutate when an open-neutral fault occurs, as shown in Fig. 2. At the same time, the neutral-point potential will offset, causing light load overvoltage, heavy load under voltage and making equipment fail to work.

When an open-neutral fault occurs, whether the load is in a state of balance or not, the third harmonic current in three-phase line will mutate. Therefore, by detecting the change of the third harmonic current, the open-neutral fault can be detected.

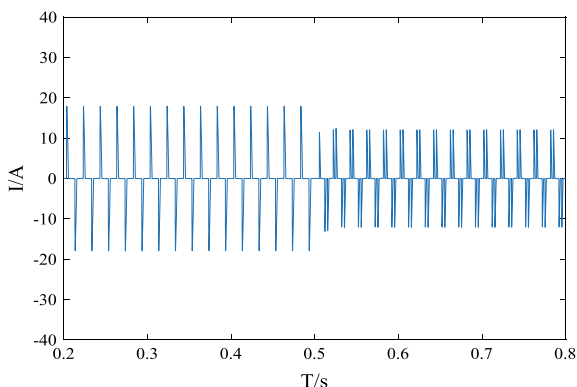


Fig. 1 Harmonic current in three-phase balanced load

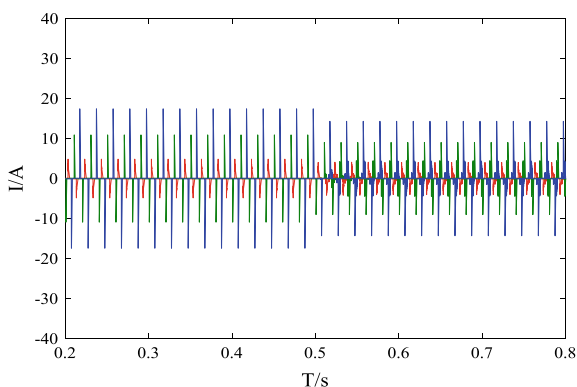


Fig. 2 Harmonic current in three-phase unbalanced load

1.2 Neutral Voltage Offset Analysis

When an open-neutral fault occurs, the neutral-point potential will mutate, and the fault can be detected by using this feature. Using Simulink to establish the simulation model, three-phase load balance and unbalance were simulated after a fault occurs at a certain point on the neutral line, the neutral-point potential offset was obtained, as shown in Figs. 3 and 4. It can be seen from the figures that the addition of non-linear load will distort the neutral-point offset voltage. When the load is in a state of balance, the neutral offset voltage is small and cannot affect the normal operation of equipment, as shown in Fig. 3. However, when the load is in a state of unbalance, the neutral-point offset voltage increases significantly, as shown in Fig. 4, which may lead to a series of adverse consequences. In conclusion, when an open-neutral fault occurs, the neutral-point voltage will offset, and its offset size is largely related to the load imbalance.

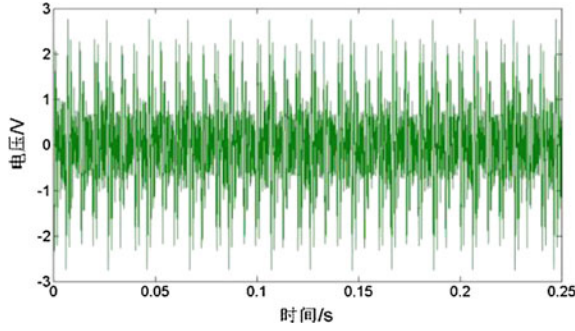


Fig. 3 Neutral-point excursion with open-neutral fault (with balanced load)

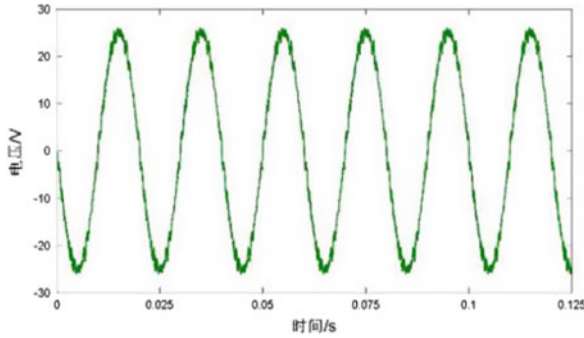


Fig. 4 Neutral-point excursion with open-neutral fault (with unbalanced load)

2 Open-Neutral Fault Detection Based on Genetic Support Vector Machine

SVM is a machine learning method developed on the basis of statistical learning theory. Its basic idea is to find the optimal hyperplane of two kinds of samples in the sample space under the linear condition. In nonlinear cases, first choose a nonlinear transformation $\varphi(x)$, the n -input, one-dimensional output vector (x_i, y_i) from the original sample space is mapped to high-dimensional feature space, and then establish optimal hyperplane in high-dimensional feature space:

$$w\varphi(x) + b = 0 \quad (1)$$

where w is the weight vector and b is the deviation.

Therefore, the two classification problems of the original sample space can be expressed as

$$y_i(w\varphi(x) + b) \geq 1, i = 1, 2, \dots, l \quad (2)$$

Considering that some samples cannot be separated by (2), nonnegative relaxation variable ξ_i is introduced in order to consider maximum classification interval and minimum error sample when determining the optimal hyperplane, the constant C control punishment degree of wrongly classified sample, so constructing optimal hyperplane is converted to the following optimization problem:

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} & y_i(w\varphi(x) + b) \geq 1 - \xi_i \end{aligned} \quad (3)$$

Introducing Lagrange multiplier α_i , converting the original problem into dual problem for solving:

$$\begin{aligned} \max & \left(\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \varphi(x_i) \varphi(x_j) \right) \\ \text{s.t.} & 0 \leq \alpha_i \leq C, \quad \sum_{i=1}^l y_i \alpha_i = 0 \end{aligned} \quad (4)$$

The optimal classification function is obtained by solving the above problems:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^m y_i \alpha_i K(x, x_i) + b \right\} \quad (5)$$

where $K(x, x_i) = \varphi^T(x) \varphi(x_i)$ is the kernel function, and m is the number of support vectors.

2.1 Kernel Function Selection

SVM is a kind of learning algorithms based on statistical learning theory, the selection of kernel function, and kernel parameters have a decisive effect on the performance of SVM, choosing different kernel functions can produce different SVM algorithms, and the commonly used kernel functions are polynomial kernel function, quadric kernel function, sigmoid kernel function, and RBF kernel function. Since RBF kernel function can effectively deal with nonlinear classification problems, this paper selects RBF kernel function expressed is as follows:

$$K(x_i \cdot x_j) = \exp \left(-\frac{\|x_j - x_i\|^2}{2\sigma^2} \right) = \exp \left(-g \|x_j - x_i\|^2 \right) \quad (6)$$

After selecting RBF as the kernel function, the parameters that the SVM model needs to optimize are punishment factor C and kernel parameter g , punishment factor C determines the size of training error and the strength of generalization ability. The kernel parameter g affects the distribution complexity of sample data in the feature space and the correlation degree of support vectors. Only the right parameters are selected can make the SVM have good classification effect, SVM parameters generally are determined by grid searching, but the method needs to determine the range of parameters and step length by the human, and the recognition effect is not ideal. In order to overcome the above defects, avoiding the blindness of parameters selection, this article chooses the genetic algorithm for the SVM parameters optimization.

2.2 SVM Parameters Selection Based on Genetic Algorithm Optimization

The kernel function is a key to influence the performance of the SVM classification, the choice of kernel function and to find the optimal combination of parameters (C , g) method is particularly important, deciding the complexity of the linear classification in high-dimensional space, reducing the feature vectors dimension, improving real-time of the classification. Genetic algorithm is an optimal search algorithm, which simulates the mechanism of natural selection and genetic evolution. Because of optimizing the two parameters C and g , set up a primitive population of 40 individuals, each individual has 2 random variables, a fitness function is the accuracy of SVM classification.

2.2.1 Selection Operator

The selection operator mainly simulates the natural selection mechanism of biological evolution, in order to keep excellent individuals of every generation into the next generation without suffering the influence of crossover and mutation, the method of roulette wheel selection is adopted to select individuals so that individuals with high fitness are more likely to enter the next generation.

2.2.2 Crossover and Mutation Operators

The selection of crossover probability P_C and variation probability P_M plays a key role in algorithm convergence speed, but the standard genetic algorithm's P_C and P_M are fixed so that the algorithm is prone to "premature". Therefore, an automatic change of P_C and P_M values is adopted according to fitness value, when the individual fitness is low, increasing the value of crossover and mutation probability. The specific operation is as follows:

$$p_c = \begin{cases} p_{c1} - (p_{c1} - p_{c2}) \frac{(f' - f_{\arg})}{f_{\max} - f_{\arg}} & f' \geq f_{\arg} \\ p_{c1} & f' < f_{\arg} \end{cases} \quad (7)$$

$$p_m = \begin{cases} p_{m1} - (p_{m1} - p_{m2}) \frac{(f'' - f_{\arg})}{f_{\max} - f_{\arg}} & f'' \geq f_{\arg} \\ p_{m1} & f'' < f_{\arg} \end{cases} \quad (8)$$

where p_{c1} and p_{c2} represent the largest and smallest crossover probability, p_{m1} and p_{m2} represent the maximum and minimum mutation probability, $f_{\arg}$ represents the average fitness value of each generation group, $f_{\max}$ represents the group's largest fitness value in each generation, f' stands for individuals with greater fitness in the crossover, and f'' represents individuals with greater fitness in the mutate.

In this paper, genetic algorithm is adopted to optimize the parameters of SVM model and to find the optimal parameter combination (C, g). The specific optimization steps are as follows:

- (1) System initialization, including parameters and initial population.
- (2) Calculate the objective function value.
- (3) Judge whether the termination condition has been reached. If so, output the result; otherwise, turn to step (4).
- (4) Calculate fitness values, simultaneously carry on the selection, the crossover, and the variation.
- (5) Calculate the objective function value of the subgeneration and replug the sub-generation into the population. Return to step (3).

3 Simulation Studies

When open-neutral fault occurs in underground space, the harmonic current of the three-phase line will change and the neutral-point potential will offset. The traditional detection method is to use the harmonic current or neutral offset voltage alone as a criterion of diagnosis, but it is easy to have a phenomenon of misjudgment. In order to solve the above problem, this article selected the three-phase line harmonic current I_A, I_B, I_C , and the neutral-point offset voltage U_0 as fault characteristics. In order to reduce all kinds of sample data due to the quantitative difference, the normalization procedure was applied to the original data obtained from the simulation, the specific equation is

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (9)$$

where x_i is the original sample data, $x_{\max}$ is the maximum value in the same sample data, $x_{\min}$ is the minimum value in the same sample data, and x'_i is the normalization data.

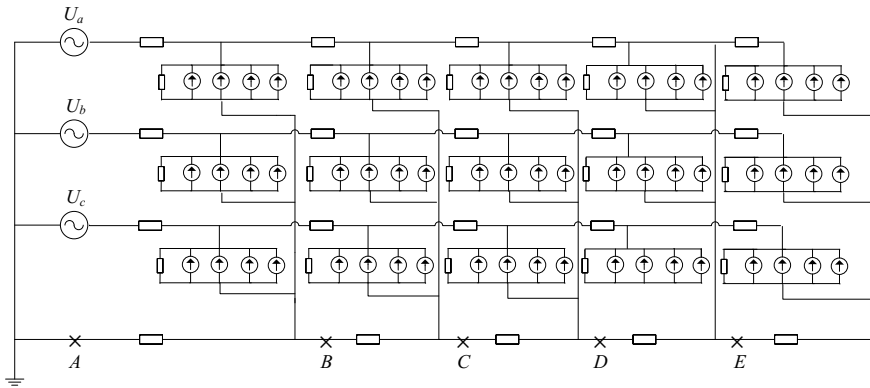


Fig. 5 The configuration of simulated low-voltage feeder

Based on GA-SVM, the fault diagnosis process of open-neutral fault in N underground space low-voltage distribution system is as follows:

- (1) Normalized the training samples and test samples;
- (2) Optimize the parameters of SVM by genetic algorithm;
- (3) The optimized parameters are substituted into the classification algorithm of SVM and establish corresponding classification model; and
- (4) Diagnose test samples by the established classification model.

Build a low-voltage distribution system as shown in Fig. 5, different loads are added to the three-phase line, the data acquisition module is added to the first end of the line and the neutral line. In Fig. 5, U_a , U_b , and U_c represent the three-phase power supply, A, B, C, D, and E represent different broken line positions. The above simulation diagram is used for simulation in Simulink, adding different load combinations and setting different break positions can obtain sample data used in the open-neutral fault.

In the simulation experiment, the SVM penalty factor C parameters change in the range of [0.1, 100], kernel function parameter g change in the range of [0.1, 1000], the biggest evolutionary iteration number is 200, and the population is 40. In the whole process of SVM parameter optimization, the accurate classification rate of fitness function is adopted as the evaluation standard. The optimal fitness value remained stable throughout the optimization process, is 98.8764%, finally determining the optimal parameter combination $C = 3.9824$, $g = 2.9691$, set the test data into the optimized SVM for diagnosing (Fig. 6).

Combined with the above optimization model, 100 groups of sample data are selected as test sets which are used for classification. The test results are shown in Table 1. It can be seen from the table that open-neutral fault detection based on GA-SVM better than other methods in terms of diagnosis speed and accuracy.

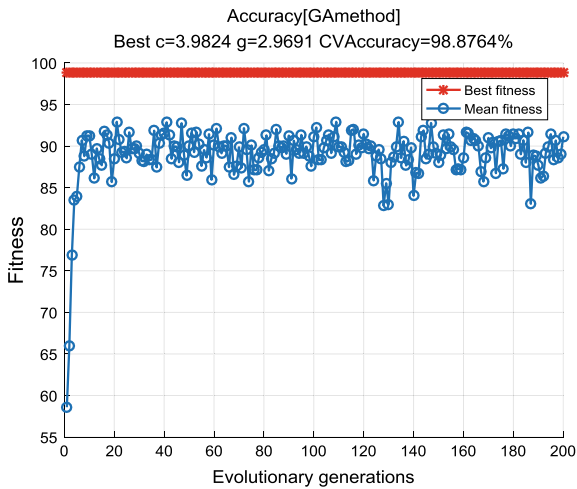


Fig. 6 GA find the fitness curve of optimal parameters

Table 1 Relationship between processing times and reworking rate

Test model	Sample numbers	Accuracy (%)	Time (s)
GA-SVM	100	98.8764	1.89
SVM	100	89.37	2.47
Grid-search SVM	100	92.54	2.18
PSO-SVM	100	96.83	1.96

4 Conclusions

This paper dealt with the problem of fault detection when an open-neutral fault occurred in an underground space low-voltage distribution system. The variation of harmonic current and neutral offset voltage were taken as the input of SVM, and the genetic algorithm was used to optimize SVM parameters, the normal or fault state can be distinguished from the output of SVM. The experimental result shows that the method of optimizing the parameters of SVM based on genetic algorithm can detect the fault well comparing to other methods. But in the process of simulation, the normal fluctuation of a load was ignored. When the load fluctuates normally, the fault detection method may be misjudged, so the influence of load fluctuation on the detection method needs to be considered in the further research.

References

1. Li, Z., Liu, W., Wang, C.: The utility model relates to a new type of neutral line breaking protection module for low voltage circuit breakers. CN201020553735[P].2010-09-10
2. Yong, J., Zhou, H.: A harmonic-based approach for open-neutral fault detection in low voltage systems. *Power Syst. Protect. Control* **19**, 49–57 (2014)
3. Yong, J., Zhou, C., Yang, B.: A harmonic-based approach for open-neutral fault detection in low voltage systems. In: IEEE, International Conference on Harmonics and Quality of Power. IEEE, pp. 53–58 (2012)
4. Liu, C., He, Z., Chen, Z.: A new method for zero-line disconnection detection with zero sequence current criterion. *Energy* **42**(4), 40–45 (2015)
5. Chen, Q., Liu, C., Liao, W.: Principle and implementation of adaptive neutral line breaking protection. *Electric Power* **50**(5), 46–51 (2017)
6. Wang, P., Liu, B.: Detection method of neutral line breaking in three-phase four-wire low-voltage power supply system. *Electrotech. Appl.* **16**, 48–50 (2015)

Author Index

B

Bai, Yan, 241
Bai, Zhu, 293, 301
Bao, Pengxiang, 101
Bian, Shuai, 301
Bi, Kaiyuan, 75

C

Cao, Huaqian, 485, 497
Chen, Bing, 85
Cheng, Lei, 101
Chen, Li, 419, 433, 557
Chen, Qing, 557
Chen, Si, 485, 497
Chen, Wei, 573
Chen, Xi, 197, 407
Chong, Faguang, 371

D

Dai, Yunchuang, 385
Dang, Chuangyin, 363
Ding, Mengli, 601
Dong, Xiaoting, 75
Duan, Yuzhou, 353

F

Fang, Qiansheng, 11, 23
Fan, Shanshan, 333
Feng, Zengxi, 249
Fu, Baochuan, 363
Fu, Hongrui, 3
Fu, Jundong, 419, 433, 557

G

Guan, Feng, 343
Guan, Xiaohong, 197
Guan, Xingquan, 333
Guo, Lingzhong, 321
Guo, Xiujuan, 477

H

Han, Bingcheng, 515
Han, Guanghong, 407
Han, Zhonghua, 75
Hong, Dejian, 23
Huang, Bingqiang, 309
Huang, Kuan, 3
Huang, Luming, 419
Huang, Mingxia, 293
Huang, Pei, 309
Hu, Junning, 231

J

Jia, Haiqi, 231
Jiang, Hongmei, 261
Jiang, Shiqi, 525
Jiang, Ziyang, 385, 457
Jia, Zhiyang, 281
Jin, Hang, 231

L

Leng, Tianhang, 433
Li, Chuan, 363
Li, Congxin, 231
Li, Ding, 125

Li, Guoping, 583
 Li, Huai, 183
 Li, Hui, 353
 Li, Huifei, 343
 Li, Jiake, 485, 497
 Li, Juelong, 609
 Li, Meng, 95
 Li, Qi, 515, 525
 Liu, Mingsheng, 537
 Liu, Qin, 101
 Liu, Qite, 219
 Liu, Wei, 183
 Liu, Weidong, 333, 371
 Liu, Xiaodong, 163
 Liu, Yang, 125
 Liu, Zhengmin, 241
 Li, Wei, 271
 Li, Ye, 231

M

Ma, Liangdong, 59
 Ma, Xiaoteng, 153
 Ma, Zhixian, 537

P

Peng, Tongrui, 115
 Piao, Wenjie, 173

Q

Qian, Xuegen, 219
 Qiao, Feng, 321
 Qin, Yugen, 59
 Qiu, Yunxia, 419, 557

R

Rao, Zhongtian, 249
 Ren, Qingchang, 241, 261
 Ren, Yanhuan, 207

S

Shang, Jianglan, 515
 Shao, Qianqian, 343
 Song, Haolin, 3, 321
 Song, Yabo, 353
 Sun, Fukang, 547
 Sun, Haochen, 321
 Sun, Kailai, 153
 Sun, Liangliang, 231
 Sun, Lishuang, 371

Sun, Wei, 137
 Sun, Xixian, 75

T

Tokhi, M. Osman, 115
 Tu, Chang, 485, 497

W

Wang, Honglin, 95
 Wang, Jiaming, 271
 Wang, Jinling, 371
 Wang, Mingyue, 11
 Wang, Ping, 485, 497, 621
 Wang, Shiqiang, 219, 583, 593, 609
 Wang, Xiaohui, 505
 Wang, Xin, 101
 Wang, Xuetao, 385
 Wang, Yahui, 469
 Wang, Ye, 547
 Wang, Yifan, 385, 397
 Wang, Yuhao, 85
 Wang, Zhaoyan, 321
 Wei, Liming, 477
 Wu, Wenmiao, 11, 23
 Wu, Xudan, 525
 Wu, Yangyun, 477
 Wu, Yixiu, 333
 Wu, Zhengtian, 363

X

Xiang, Xinjian, 309
 Xie, Chenlei, 137
 Xie, Liqiang, 601
 Xing, Jianchun, 447, 573, 601, 609, 621
 Xing, Yan, 371
 Xue, Guangtong, 447
 Xu, Yiyang, 59

Y

Yang, Guoqing, 85
 Yang, Qiliang, 447, 601
 Yang, Xiong, 45, 207
 Yang, Xu, 75
 Yang, Yalong, 11, 23
 Yan, Zhe, 173
 Yao, Yufeng, 115
 Ye, Qinghong, 469
 Yin, Xiaoli, 537
 Yong, Jing, 35

Yue, Jinglin, [433](#)
Yu, Junqi, [45](#), [207](#), [219](#), [249](#)
Yu, Mingwei, [515](#)
Yun, An, [469](#)
Yu, Qiuyue, [101](#)
Yu, Zhen, [183](#)

Z

Zhang, Haiqiang, [293](#)
Zhang, Hui, [163](#)
Zhang, Jili, [59](#), [537](#)
Zhang, Junlin, [45](#)
Zhang, Ping, [163](#)
Zhang, Rui, [11](#), [23](#)
Zhang, Xuwei, [573](#)
Zhang, Yi, [137](#)
Zhang, Yongming, [173](#)
Zhang, Yuhan, [593](#)
Zhang, Yunfeng, [343](#)

Zhang, Zhenya, [485](#), [497](#)
Zhao, Anjun, [45](#), [207](#), [219](#), [249](#)
Zhao, Jin, [35](#)
Zhao, Qianchuan, [153](#), [385](#), [397](#), [407](#), [457](#)
Zhao, Shuo, [447](#)
Zhao, Tianyi, [271](#), [281](#)
Zhao, Xuechen, [621](#)
Zhao, Yunfeng, [469](#)
Zheng, Jinjin, [363](#)
Zhou, Meng, [45](#), [207](#)
Zhou, Qizhen, [573](#)
Zhuang, Luping, [197](#)
Zhu, Quanmin, [115](#)
Zhu, Shaofeng, [547](#)
Zhu, Xueli, [173](#)
Zhu, Xulai, [11](#), [23](#)
Zong, Shiqi, [505](#)
Zou, Jianhong, [153](#)