

Frontiers
in
Artificial
Intelligence
and
Applications

PROCEEDINGS OF CECNET 2021

The 11th International Conference on Electronics, Communications and Networks (CECNet), November 18-21, 2021

Edited by
Antonio J. Tallón-Ballesteros



IOS Press

PROCEEDINGS OF CECNET 2021

It is almost impossible to imagine life today without the electronics, communications and networks we have all come to take for granted. The 6G network is currently under development and some chips able to operate at the Terahertz (THz) scale have already been introduced, so the next decade will probably see the consolidation of 6G-based technology, as well as many compliant devices.

This book presents the proceedings of the 11th International Conference on Electronics, Communications and Networks (CECNet 2021), initially planned to be held from 18-21 November 2021 in Beijing, China, but ultimately held as an online event due to ongoing COVID-19 restrictions. The CECNet series is now an established annual event attracting participants in the interrelated fields of electronics, computers, communications and wireless communications engineering and technology from around the world. Careful review by program committee members, who took into consideration the breadth and depth of those research topics that fall within the scope of CECNet, resulted in the selection of the 88 papers presented here from the 325 submissions received. This represents an acceptance rate of around 27%.

Providing an overview of current research and developments in these rapidly evolving fields, the book will be of interest to all those working with digital communications networks.



ISBN 978-1-64368-240-2 (print)

ISBN 978-1-64368-241-9 (online)

ISSN 0922-6389 (print)

ISSN 1879-8314 (online)

PROCEEDINGS OF CECNet 2021

Frontiers in Artificial Intelligence and Applications

The book series Frontiers in Artificial Intelligence and Applications (FAIA) covers all aspects of theoretical and applied Artificial Intelligence research in the form of monographs, selected doctoral dissertations, handbooks and proceedings volumes. The FAIA series contains several sub-series, including 'Information Modelling and Knowledge Bases' and 'Knowledge-Based Intelligent Engineering Systems'. It also includes the biennial European Conference on Artificial Intelligence (ECAI) proceedings volumes, and other EurAI (European Association for Artificial Intelligence, formerly ECCAI) sponsored publications. The series has become a highly visible platform for the publication and dissemination of original research in this field. Volumes are selected for inclusion by an international editorial board of well-known scholars in the field of AI. All contributions to the volumes in the series have been peer reviewed.

The FAIA series is indexed in ACM Digital Library; DBLP; EI Compendex; Google Scholar; Scopus; Web of Science: Conference Proceedings Citation Index – Science (CPCI-S) and Book Citation Index – Science (BKCI-S); Zentralblatt MATH.

Series Editors:

Joost Breuker, Nicola Guarino, Pascal Hitzler, Joost N. Kok, Jiming Liu,
Ramon López de Mántaras, Riichiro Mizoguchi, Mark Musen, Sankar K. Pal,
Ning Zhong

Volume 345

Recently published in this series

- Vol. 344. B. Brodaric and F. Neuhaus (Eds.), *Formal Ontology in Information Systems – Proceedings of the Twelfth International Conference (FOIS 2021)*
- Vol. 343. M. Tropmann-Frick, H. Jaakkola, B. Thalheim, Y. Kiyoki and N. Yoshida (Eds.), *Information Modelling and Knowledge Bases XXXIII*
- Vol. 342. P. Hitzler and M.K. Sarker (Eds.), *Neuro-Symbolic Artificial Intelligence: The State of the Art*
- Vol. 341. A.J. Tallón-Ballesteros (Ed.), *Modern Management based on Big Data II and Machine Learning and Intelligent Systems III – Proceedings of MMBD 2021 and MLIS 2021*
- Vol. 340. A.J. Tallón-Ballesteros (Ed.), *Fuzzy Systems and Data Mining VII – Proceedings of FSDM 2021*
- Vol. 339. M. Villaret, T. Alsinet, C. Fernández and A. Valls (Eds.), *Artificial Intelligence Research and Development – Proceedings of the 23rd International Conference of the Catalan Association for Artificial Intelligence*
- Vol. 338. C. Frasson, K. Kabassi and A. Voulodimos (Eds.), *Novelties in Intelligent Digital Systems – Proceedings of the 1st International Conference (NIDS 2021), Athens, Greece, September 30 – October 1, 2021*

ISSN 0922-6389 (print)
ISSN 1879-8314 (online)

Proceedings of CECNet 2021

The 11th International Conference on Electronics,
Communications and Networks (CECNet),
November 18–21, 2021

Edited by

Antonio J. Tallón-Ballesteros

University of Huelva, Spain



IOS Press

Amsterdam • Berlin • Washington, DC

© 2022 The authors and IOS Press.

This book is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

ISBN 978-1-64368-240-2 (print)

ISBN 978-1-64368-241-9 (online)

doi: 10.3233/FAIA345

Publisher

IOS Press BV

Nieuwe Hemweg 6B

1013 BG Amsterdam

Netherlands

fax: +31 20 687 0019

e-mail: order@iospress.nl

For book sales in the USA and Canada:

IOS Press, Inc.

6751 Tepper Drive

Clifton, VA 20124

USA

Tel.: +1 703 830 6300

Fax: +1 703 830 2300

sales@iospress.com

LEGAL NOTICE

The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

Preface

Electronics, Communication and Networks coexist and it is not possible to conceive the current society without any of the previous terms. 6G network is currently under development and more researchers are joining the research on 6G. Additionally, some chips able to operate at the Terahertz (THz) scale have been already introduced. Probably, next decade would be the scenario to observe the consolidation of 6G-based technology as well as lots of compliant devices.

The Conference on Electronics, Communications and Networks (CECNet) series has been established as a mature event after ten previous years of existence. CECNet is held annually covering many interrelated groups of topics such as:

- Electronics technology.
- Communication engineering and technology.
- Wireless communications engineering and technology.
- Computer engineering and technology.

The 11th International Conference on Electronics, Communications and Networks (CECNet 2021) was primarily scheduled to be held in Beijing, China on November 18–21, but was finally transformed into an online conference for the same reason as last year (COVID-19).

This book contains the papers accepted and presented at the 11th International Conference on Electronics, Communications and Networks (CECNet 2021), held on 18–21 November 2021 in a virtual way instead of onsite participation in Beijing (China). All papers were carefully reviewed by program committee members and took into consideration the breadth and depth of the research topics that fall into CECNet scope.

CECNet 2021 received 325 submissions and after a vivid discussion stage, the committee decided to accept 88 papers, which represents an acceptance rate of 27.07%.

I would like to thank all the keynote speakers and authors for their effort in preparing a contribution for this leading international conference. Moreover, I am very grateful to the people, especially the program committee members and reviewers, who devoted time to evaluate the papers. It is a great honour to continue with the publication of these proceedings in the prestigious series Frontiers in Artificial Intelligence and Applications (FAIA) by IOS Press. Our particular thanks also go to J. Breuker, N. Guarino, J.N. Kok, R. López de Mántaras, J. Liu, R. Mizoguchi, M. Musen, S.K. Pal and N. Zhong, who are the FAIA series editors, for supporting again this conference.

Finally, I hope you enjoy, so far, your virtual visit to Beijing, although with hopes of attending the conference in the future, face-to face.

Antonio J. Tallón-Ballesteros
Huelva city (Spain)
University of Huelva (Spain)

This page intentionally left blank

Contents

Preface	v
<i>Antonio J. Tallón-Ballesteros</i>	
Internet User Stickiness in the Social Q & A Community from the Perspective of Interactivity – A Case Study on Zhihu Website	1
<i>Sheng-jun Yuan and Jia-xin Chen</i>	
Webliometric Indicators as Elements of the AI Technique of Estimation of the Language Teacher's Net Proficiency	7
<i>Matukhin Pavel, Provotorova Elena, Petrova Marina, Gracheva Olga, Rybakova Irina and Ismailova Kholisakhon</i>	
Sequence-in-Sequence Learning for SOH Estimation of Lithium-Ion Battery	14
<i>Thien Pham, Loi Truong, Mao Nguyen, Akhil Garg, Liang Gao and Tho Quan</i>	
A Pointwise Evaluation Metric to Visualize Errors in Machine Learning Surrogate Models	26
<i>Seyed Shayan Sajjadinia, Bruno Carpentieri and Gerhard A. Holzapfel</i>	
Multi-Stage Prediction of Feed System Time Series	35
<i>Lihua Shen, Biling Wang and Hongjun Liu</i>	
Multi-Class Imbalanced Corporate Bond Default Risk Prediction Based on the OVO-SMOTE-Adaboost Ensemble Model	42
<i>Jie Sun and Jingmei Zhu</i>	
An Effective Deep Neural Network Architecture for Cross-Subject Epileptic Seizure Detection in EEG Data	54
<i>Imene Jemal, Amar Mitiche, Lina Abou-Abbas, Khadidja Henni and Neila Mezghani</i>	
Prediction of Alzheimer's Disease Based on Coordinate-Dense Attention Network	63
<i>Yongmei Tang, Xiangyun Liao, Weixin Si and Zhigang Ning</i>	
Study on Dune Morphology in Computer Mapping Based on 3S Technology	71
<i>Zhidan Ba</i>	
A Study of Forest Swamp Mapping in Hani Wetland Integrating Sentinel-1 and Sentinel-2 Satellite Images	77
<i>Jingfa Wang</i>	
The RBF Hyperparameter in Metamodel-Assisted Evolutionary Search	84
<i>Yoel Tenne</i>	
Online Life Forecasting Method and State Assessment Architecture of Power Cable Based on Recurrent Neural Network and Internet of Things	90
<i>Changcheng Song, Binglei Xue, Fumu Lu, Feng Lan, Zheng Yang, Chunlei Zhang and Peiran Yuan</i>	

Virtual Power Plant Platform for Demand Response Based on Microservice Architecture	102
<i>Guodong Jiang, Zuixi Xie, Tao Han, Hongwei Du, Rongqiang Feng, Shouquan Luo, Xueqiong Wu and Jiaqi Zhang</i>	
Distributed Autonomous Economic Control Strategy for Microgrid Considering Event Triggering Mechanism	118
<i>Yulong Xiong, Shihong Miao, Weichen Yang and Zhiwei Liu</i>	
Analysis of Current Production Status and Key Existing Problems of Tower-Shaped Solar Molten Salt Storage Tank	132
<i>Hang Xi, Qiong Wu, Xiaojun Xie, Ruigang Zhang, Bo Yang and Qian Liu</i>	
Application Status and Research of Instrumentation Device in Solar Thermal Power Generation Industry	141
<i>Xiaojun Xie, Hang Xi, Qiong Wu, Ruigang Zhang, Bo Yang and Zengbo Liu</i>	
Thermal System Simulation Study of Wide-Load Out-Of-Stock Technical Transformation to Reduce SCR Inlet Flue Gas Temperature of 300MW Subcritical Boiler	148
<i>Tao Qin, Jun Rong, Guang Yang, Yankai Wang, Yi Han, Bin Cai and Yingli Yu</i>	
Research on Identification of Power Grid Weakness Based on Bayesian Inference	159
<i>Jiang Hu, Wei Li, Wenxia Liu, Xianggang He and Yu Zhang</i>	
Discussion on Integrated Modeling Technology of New Generation Centralized Control Station Equipment Monitoring System	171
<i>Yi Zhai, Tianpin Xu, Qijing Yang, Weihua Cao, Zhichao Qi and Chunyao Liu</i>	
Development and Application of Virtual Simulation Teaching Platform for the Transcritical CO ₂ Two-Stage Compression Refrigeration System Based on LabVIEW	178
<i>Yuyao Sun, Jinfeng Wang, Jing Xie and Chenlong Wu</i>	
Study on Optimization Path of Energy Consumption Structure of Container Terminal	188
<i>Ruiyu Li, Haifeng Fang, Guodong Wang and Qingfang Ma</i>	
Development and Implementation of High Precision Positioning Technology for Electric Vehicle Wireless Charging Coils	199
<i>Jie Li, Kejun Qian, Yafei Li, Yi Liu, Zifeng Liu and Linlin Tan</i>	
Study on Type Selection of Relays in Outdoor Terminal Box of Converter Station	211
<i>Jianying Wu, Zhen Wang and Xilin Yan</i>	
Research on Distribution Network Reliability Investment Estimation Method Based on Sequence Linearization Correlation Analysis	220
<i>Huazhen Cao, Chong Gao, Yaxiong Wu, Hao Li, Zijun Wang and Chengmin Wang</i>	
Research on the Construction of Online Culture Trading Platform Based on Netizens' Wishes	233
<i>Ren Song, Weiwei Wang and Yang Liu</i>	

Parametric Optimization of Integrated Circuit Assembly Process: An Evolutionary Computing-Based Approach <i>Tatjana Sibalija</i>	239
Smart Manufacturing and Jobs <i>Juvenal Mendoza-Valencia</i>	247
Uncertain Machine Load Forecasting Based on Least Squares Support Vector Machine <i>Qiaofeng Meng</i>	255
Edge Loss for Remote Sensing Image Super-Resolution <i>Jiaoyue Li, Weifeng Liu, Kai Zhang and Baodi Liu</i>	262
Phase Tracking Sequences for 5G NR in 52.6-71 GHz Band: Design and Analysis <i>Alexander Maltsev, Andrey Pudeev, Seonwook Kim, Suckchel Yang, Seunghwan Choi and Sechang Myung</i>	268
Investigating Automated Hyper-Parameter Optimization for a Generalized Path Loss Model <i>Usman Sammani Sani, Daphne Teck Ching Lai and Owais Ahmed Malik</i>	283
Calculation of RMS Values for Variable Frequency Sinusoidal Signals, Using Phasors and Digital SAL and CAL Filters <i>Yonatan Aguirre, Flabio Gutierrez, Richard Abramonte, Nilthon Arce and Antenor Aliaga</i>	292
Optical Vortices Sharp Focusing by Silicon Ring Gratings Using High-Performance Computer Systems <i>Dmitry Savelyev</i>	300
An Explainable Convolutional Neural Networks for Automatic Segmentation of the Left Ventricle in Cardiac MRI <i>Jun Liu, Feng Deng, Geng Yuan, Xue Lin, Houbing Song and Yanzhi Wang</i>	306
Total Variance Approach on Commercial Vehicle Handling Dynamics <i>Yaohua Guo and Jinquan Ding</i>	315
Compressive Video Sensing Based on Intra-Inter-Frame Constraints and Genetic Algorithm <i>Yuchen Yue, Hua Li and Jianhua Luo</i>	321
Koch Snowflake Fractal Antenna Design in the Deep Space Bands for a Constellation of Cubesat Explorers <i>Orlando Francois Gonzales Palacios, Ricardo Erick Diaz Vargas, Patrick H. Stakem and Carlos Enrique Arellano Ramirez</i>	335
Searching and Rescuing Victims in Emergency: A Comprehensive Survey <i>Huibo Bi and Erol Gelenbe</i>	343
Design of Bluetooth Automatic Check in System Without Operation Based on Cross Mobile Platform <i>Yiqin Bao and Zhengtang Sun</i>	349

Exploration and Research on Integrating Programming Education into Junior Middle School Mathematics Classroom <i>Yiqin Bao, Gui Wang and Zhengtang Sun</i>	356
RISC-V Based Safety System-on-Chip with Hardware Comparator <i>Eike Hahn, Dominik Kalinowski, Waldemar Mueller, Mohamed Abdelawwad and Josef Boercsoek</i>	362
Classification of Knee Osteoarthritis Severity Using Modified Masks to Preprocess X-ray Images in a Deep Learning Model <i>Ching-Chung Yang</i>	369
Securing an IoT Medical System Using AI and a Unidirectional Network Device: Application to a Driver <i>Georges El Hajal, Roy Abi Zeid Daou, Yves Ducq and Josef Boercsoek</i>	379
A Master Station and Terminal Data Exchange Method Based on Symmetric and Asymmetric Algorithms <i>Fan He, Zhengquan Ang, Guanglun Yang, Qingqin Fu, Ling Yi, Pingjiang Xu, Jia Liu, Zhaoqing Liang, Changsheng Niu and Jianying Chen</i>	387
Approach to Define the Reliability of Safety-Related Machine Learning Based Functions in Highly Automated Driving <i>Pengyu Si, Ossmane Krini, Nadine Müller and Aymen Ouertani</i>	394
Analysis of Signal Processing Methods for Formation of Radar Range Profiles <i>Nguyen Van Loi, Le Thanh Son, Tran Trung Kien, Tran Van Truong and Tran Vu Hop</i>	402
Three Term Weighted Type Fractional Fourier Transform Based Generalized Hybrid Carrier and Its Application into PAPR Suppression <i>Yong Li, Zhiqun Song, Teng Sun and Bin Wang</i>	414
A Novel Dual Mode Decision Directed Multimodulus Algorithm (DM-DD-MMA) for Blind Adaptive Equalization <i>Sagi Tadmor, Sapir Carmi and Monika Pinchas</i>	420
Efficiency Measurement of Compressed Air Compressors Using High Availability SoC with 1oo2 Redundancy Architecture <i>Mohamed Abdelawwad, Eike Hahn, Josef Boercsoek, Julian Fairbrother, Tarek Al Shahadat and Peter Otto</i>	431
A New Identity-Based Encryption Scheme with Accountable Authority Based on SM9 <i>Ke Wang, Yuan Zhao, Song Luo and Zhi Guan</i>	440
Incremental 2D Grid Map Generation from RGB-D Images <i>Tingfeng Ye, Juzhong Zhang, Yingcai Wan, Ze Cui and Hongbo Yang</i>	450
Flat OFC Generation Based on DPMZM Cascaded Dual-Parallel PolM with Frequency Multiplication Circuit <i>Xin Wang, Sijia Liu, Ximing Wang, Yingxi Miao, Caili Gong and Yongfeng Wei</i>	459

Algebra Based Human Skeleton Sequence Reduction and Action Recognition <i>Shibin Xuan, Kuan Wang, Lixia Liu, Chang Liu and Jiaxiang Li</i>	467
An Empirical Study on the Fossilization of English Language Learning in the Context of Multimedia Network Teaching <i>Lihua Huang</i>	477
A Study on the Efficiency and Countermeasures of Network-Based Autonomous Learning Platform for Vocational College English Learners <i>Liping Jiang</i>	484
Detecting of All Zero Blocks in HEVC for RDOQ <i>Nana Shan, Henglu Wei, Wei Zhou and Zhemin Duan</i>	493
Visible Light Wireless Data Center Links with Distinct Beam Configurations <i>Jupeng Ding, Chih-Lin I and Lili Wang</i>	500
A Study on Fault Tolerance Technology of Flight Control Computer for Unmanned Aerial Vehicle <i>Chen Zhang and Jihui Pan</i>	507
Ultra-HD Video Streaming in 5G Fixed Wireless Access Bottlenecks <i>Koffka Khan and Wayne Goodridge</i>	515
Design and Implementation of Highly Integrated Electronic System for Small-Size Spacecraft <i>Xinfeng Yan, Wei Liu, Gaosi Li and Jian Geng</i>	525
SDN-Enabled 3C Resource Integration in Green Internet of Electrical Vehicles <i>Handi Chen, Xiaojie Wang, Zhaolong Ning and Lei Guo</i>	536
Video Pre-Caching Joint Hand-Off and Content Delivery in Multi-Access Edge Computing Based EONs <i>Yutong Chai, Shan Yin, Lihao Liu, Liyou Jiang and Shanguo Huang</i>	544
Research and Implementation of High-Speed Data Streams Symbol Synchronization Algorithm Using Training Sequence in IMDD-OOFDM System <i>Dongsheng Zhang, Gang Zhang, Jiawei Wu, Yunjie Xiao, Liang Liang, Chao Lei and Huibin Zhang</i>	550
Spectrum Availability Aware Routing and Resource Allocation for Point-to-Multipoint Services in Mixed-Grid Optical Networks <i>Feng Wang, Qingcheng Zhu, Xiaolong Li, Jiaming Gu, Zhenhua Yan, Xiaosong Yu and Yongli Zhao</i>	557
Low-Cost Deployment Scheme of VNF and PNF in Optical Datacenter Networks <i>Jianghua Wei, Xin Li, Jingjie Xin, Ying Tang, Lu Zhang and Shanguo Huang</i>	563
Complex-Coefficient Microwave Photonic Filter Based on Orthogonally Polarized Optical Single-Sideband Modulation <i>Jinwang Qian, Junling Sun, Pengge Ma, Xinlu Gao and Shanguo Huang</i>	570

Node Importance Based Protection in Power-Grid Optical Backbone Communication Networks	575
<i>Xiaobo Li, Guoli Feng, Run Ma, Lu Lu, Kaili Zhang, Xinnan Ha, Wenbin Wei, Ning Wang, Xiaosong Yu and Yongli Zhao</i>	
A High-Performance Infrared Imaging System with Adaptive Contrast Enhancement	581
<i>Feng Deng, Zhong Su, Rui Wang, Jun Liu and Yanzhi Wang</i>	
ReDCN: A Dynamic Bandwidth Enabled Optical Reconfigurable Data Center Network	591
<i>Xinwei Zhang, Zuoqing Zhao, Yisong Zhao, Yuanzhi Guo, Xuwei Xue, Bingli Guo and Shanguo Huang</i>	
Routing Optimization Based on OSPF in Multi-Layer Satellite Network	597
<i>Mingjiang Fu, Bingli Guo, Hai Yang, Chengguang Pang and Shanguo Huang</i>	
Data Augmentation Algorithm Based on Generative Antagonism Networks (GAN) Model for Optical Transmission Networks (OTN)	603
<i>Liang Chen, Kunpeng Zheng, Yang Li, Xuelian Yang, Han Zhang, Yangyang Liang, Junhua Huang, Yuan Zhang, Pengfei Xie and Yongli Zhao</i>	
OSNR Prediction Based on Federal Learning in Multi-Domain Optical Networks	612
<i>Junhua Huang, Bohan Zhu, Hongxi Zhou, Qiwei Zheng, Zhuo Chen, Tong Lin, Yuanyuan Zhao, Ziqi Dong, Xiaofang Gao and Yongli Zhao</i>	
Research on 3D Fingerprint Positioning Based on MIMO Base Station	621
<i>Chang Xia, Yijie Ren, Xiaojun Wang, Weiguang Sun, Fei Tang and Xiaoshu Chen</i>	
High-Performance Fingerprint Localization in Massive MIMO System	631
<i>Yijie Ren, Zhixing Xiao, Yuan Tang, Fei Tang, Xiaojun Wang and Xiaoshu Chen</i>	
Scattering Clustering Method for Terminal Fingerprint Positioning Based on MIMO Base Station	640
<i>Fei Tang, Yijie Ren, Xiaojun Wang, Weiguang Sun and Xiaoshu Chen</i>	
A Novel OLTC Fault Diagnosis Method Based on Optimized Long Short-Term Memory Parameters	650
<i>Yan Yan and Hongzhong Ma</i>	
Automated Contextual Anomaly Detection for Network Interface Bandwidth Utilisation: A Case Study in Network Capacity Management	659
<i>Esmaeil Zadeh, Stephen Amstutz, James Collins, Craig Ingham, Marian Gheorghe and Savas Konur</i>	
An Algorithm of Vocabulary Enhanced Intelligent Question Answering Based on FLAT	667
<i>Jing Sheng Lei, Shi Chao Ye, Sheng Ying Yang, Wei Song and Guan Mian Liang</i>	
Detection Method of Computer Room Personnel Based on Improved DERT	677
<i>Dan Su, Qiong-lan Na, Hui-min He and Yi-xi Yang</i>	

Design of Speech Denoising Algorithm Based on Wavelet Threshold Function and PSO <i>Lanyong Zhang, Ruixuan Zhang and Papavassiliou Christos</i>	684
Application of Text Mining Method for Classification of Work Order in Power Grid Production <i>Yadi Zhao, Zhifeng Wei, Bingqiang Gao and Shuo Zhang</i>	693
Modeling Ebola Virus Dynamics by Colored Petri Nets <i>Dmitry A. Zaitsev, Peyman Ghaffari and Virginia Sanz Sanchez</i>	707
Image Segmentation by Pairwise Nearest Neighbor Using Mumford-Shah Model <i>Nilima Shah, Dhanesh Patel and Pasi Fräntti</i>	716
I Want to Control Your Move: Human-Human Interface (HHI) Neuromuscular Electrical Stimulator (NMES) <i>Ching Yee Yong and Terence Tien Lok Sia</i>	724
Analysis of Heart Rate Asymmetry During Sleep Stages <i>Chang Yan, Peng Li, Yang Li, Jianqing Li and Chengyu Liu</i>	731
Hierarchical Optimization Model Based on V2G Technology <i>Wei Luo, Xiaohu Zhu and Liansong Yu</i>	737
Research of Power Distribution System Dynamic Simulation Based on Grid's Harmonic Current <i>Wei Qin, Tiansong Gu and Hongliang Li</i>	745
Study on Inter-Turn Short Circuit Test for Distribution Reactor <i>Wei Qin, Tiansong Gu and Hongliang Li</i>	753
Subject Index	763
Author Index	769

This page intentionally left blank

Internet User Stickiness in the Social Q & A Community from the Perspective of Interactivity - A Case Study on Zhihu Website

Sheng-jun YUAN and Jia-xin CHEN¹

School of Business, Guilin University of Electronic Technology, Guilin 541004, China

Abstract. This paper mainly explores how the user's perceptual interaction towards a specific website affects the user stickiness and the interaction mechanism of various influencing factors of user stickiness, which may encourage socialized Q & A community operators to adopt more appropriate ways to enhance user's stickiness. We comprehensively use the planning behavior theory and technology acceptance model with the perceptual perspective of interaction. Based on the framework of "belief → attitude → intention", an attempt is made to construct a conceptual model of influencing factors of user stickiness in socialized Q & A community by collecting the data through questionnaire survey, and using spss23.0 and amos23.0 to analyze the data and validate the model. The empirical research results prove that the factors such as perceived usefulness, perceived ease of use, information exchange perception, social interaction perception satisfaction, and heart flow experience have positive effects on social users. We conclude that the operators can build and consolidate user engagement, attract and retain the users by enhancing interactive perception, optimizing user experience, and improving their satisfaction.

Keywords. Interactivity, social Q & A community, user stickiness, Zhihu website

1. Introduction

The social Q & A community is a system, in which users can ask questions, answer and evaluate in the process of interaction[1-3]. At present, the representative of domestic socialized Q & A community development is Zhihu website, which was established in 2011. The Q & A community in China, taking Zhihu as an example, is developing extensively with the spread of Internet. Though the survey shows that by 2019, the number of registered Zhihu has reached 200 million, but the data reveals that most of the registered users are silent and one-time consumers. On an average, less than 40% of active users visit Zhihu more than 7 times a week, with low user stickiness and low continuous participation. The core purpose of socialized Q & A community is to generate knowledge through continuous interaction to meet users' access to information and create value [4] (Wang Wei 2018).

¹ Corresponding Author, Jia-xin CHEN, Guilin University of Electronic Technology, Guilin 541004, China;
Email: 954743805@qq.com

The fundamental way to establish and consolidate user stickiness is to improve users' interactive perception. Only some scholars built the comprehensive planning behavior theory based on the technology acceptance model to establish the integrated use model (UTAUT) [5] and other researchers added the Expectation Confirmation Theory on the basis of the integrated model to put forward the information system continuous use model (ecm-it) [6]. But the obvious problem of these two models is that they ignore the impact of interpersonal interaction on stickiness in interactive perception and the role of factors such as flow experience. This paper aims ① to integrate the factors affecting user stickiness from the perspective of interactivity, ② to test how the influencing factors interact and, ③ which factors play a leading role? This study will more comprehensively explain the influencing factors of user stickiness, and help the operators to improve their existing strategies to enhance user stickiness.

2. Research design and hypothesis

2.1 The establishment of research model

Review of relevant literature reveals that the perceptual interaction is a multi-dimensional construct, and its dimensions are divided into different standards. In fact, the process of users participating in the social Q & A community is a combination of "human interaction with the system" and "interaction between people". Therefore, we divide the interaction into two forms: human-machine interaction and interpersonal interaction, from the perspective of user participation by are using the division method of Hoffman [7]. According to the technology acceptance model, we divide human-machine interaction into perceived ease of use and perceived usefulness [8].

In 2000, Kurniawan [9] proposed a model of influencing factors of customer stickiness of the network retailer. This model takes community involvement, website attractiveness and convenience as prepositional variables whereas, pleasure and customer satisfaction are the intermediate variables and customer stickiness is a result variable. This model verified that five predictive variables had a significant impact on customer stickiness. Based on this, an attempt is made here to formulate a conceptual model of the influence factors of the community user stickiness by taking the socialized Q & A community as a research object, combining its characteristics and comprehensively using the interactive theory, the technical acceptance model, and the heart flow experience theory.

2.2 Research hypotheses

(1) Interactive perception and satisfaction

For the social Q & A community, the community will achieve high satisfaction when it perceives that the community is easy to use and can get quick help [10]. Accordingly, in the social Q & A community, we put forward the following hypothesis:

H1a: Perceived ease of use has a positive impact on satisfaction.

H1b: Perceived usefulness has a positive impact on satisfaction.

A successful social Q & A community not only promotes the user's continuous information interaction in the community, but also helps users to establish social relations network in the process of interaction, using the platform to constantly carry out effective

information transmission and emotional interflow thereby, promoting the user satisfaction. Therefore, in the social Q & A community, we put forward the following hypothesis:

H2a: Information exchange perception has a positive impact on satisfaction.

H2b: Social interaction perception has a positive impact on satisfaction.

(2) Interactive perception and the heart flow experience

The experience of heart flow refers to the overall feeling of an individual fully invested in a certain activity.. Whether individuals can enter the heart flow experience or not depends on skills and challenges. Skill is a kind of ability for individuals to successfully complete given tasks whereas, the challenge means all the efforts required to accomplish the task. Accordingly, in the social Q&A community, we put forward the hypothesis that:

H3a: Perceived ease of use has a positive impact on the flow of heart experience

H3b: Perceived usefulness has a positive impact on the flow of heart experience

In the background of social Q&A community, there are few studies on the effects of interpersonal interaction on cardiac experience. However, the positive relationship between interpersonal interaction and heart flow experience has been verified many times in other contexts, such as tourism community, online learning platform, and online game. This paper extends the conclusion to the social Q&A community. Accordingly, we put forward the hypothesis:

H4a: Information exchange perception has a positive impact on the flow of heart experience;

H4b: Social interaction perception has a positive impact on the flow of heart experience.

The heart flow experience can attract individuals and significantly affect their subsequent attitudes and behaviors. Users may expect to experience positive experiences in social Q&A communities, and they will be satisfied when expectations are confirmed, which is confirmed in the background of online banking, sports website, and mobile shopping, and so on. Accordingly, in the social Q & A community, we put forward the hypothesis that:

H5: The heart flow experience positively affects satisfaction

When individuals enter the heart flow experience and reach the optimal state, they may immerse themselves in it and spend a lot of time digging deep into the community. When expecting this experience again, the frequency of return visits will increase. One of the characteristics of user stickiness is user's continuous use behavior. Accordingly, in the social Q&A community, we put forward the hypothesis that:

H6: The heart flow experience positively affects user stickiness

When users have positive attitude and overall attachment to the contents, functions, products and services in the system, stickiness is bound to produce. Satisfaction is a powerful determinant of sustained behavior, and is the best valuation of relationship strength in long-term relationship marketing. Many studies in the internet background have shown that user satisfaction is an important driving force for stickiness. Accordingly, in the social Q & A community, we put forward the hypothesis that:

H7: Satisfaction positively affects user stickiness.

3. Research Methods

3.1 Questionnaire Design and Collection

The measurements of the related variables in this essay are mainly cited from maturity scales at home and abroad, and are adjusted according to the research contents and China's national conditions. Seven potential variables are considered. The final questionnaire comprises of two parts.

The questionnaire was spread in some places including Guangxi province, Guangdong Province, Beijing, Shanghai, and Sichuan province. Totally, 430 questionnaires were collected including the invalid ones comprising of incomplete filling, non-subscribers from Zhihu community and complete in all respect. 372 effective questionnaires were collected with an efficiency rate is 86.5%.

It is observed that Cronbach's α and Combinatorial Reliability (CR values) of every dimension in this scale is close to or higher than 0.8, which satisfy the demand of reliability. The standard loading of every latent variable are all higher than 0.7, CR values are higher than 0.8 and the average variance extracted (AVE) are higher than 0.8, confirming that the scales have a good convergent validity. From the result, the square roots (diagonal) of the AVE of every latent variable are higher than the correlation coefficients among the latent variables and others, which shows that the distinctiveness among every latent variable is pretty strong, and the discriminant validity of the model is good.

3.2 Reliability & Validity of data

The analysis was mainly by SPSS23.0 and AMOS23.0. We used Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis for the reliability and validity verification.

Firstly, we did KMO test and Bartlett Test of Sphericity to the valid samples to confirm that the questionnaire have construct validity. The result showed that the KMO value is 0.847, higher than 0.5, the chi-square value of Bartlett Test of Sphericity is 4459.009, reaching the significant level ($p=0.000 < 0.001$), indicating that the samples are suitable for factor analysis.

Secondly, we carried out the reliability and validity verification from the index including standard loading of every dimension, combinatorial validity, combinatorial reliability, average variances extracted, etc. It is observed that Cronbach's α and Combinatorial Reliability (CR values) of every dimension in this scale is close to or higher than 0.8, which satisfy the demand of reliability. The standard loading of every latent variable are all higher than 0.7, CR values are higher than 0.8 and the average variance extracted (AVE) are higher than 0.8, confirming that the scales have a good convergent validity. The result shows that the square roots (diagonal) of the AVE of every latent variable are higher than the correlation coefficients among the latent variables and others, which shows that the distinctiveness among every latent variable is pretty strong, and the discriminant validity of the model is good.

3.3 Model Amendment

After reliability and validity verification of the questionnaire, we used AMOS23.0 to verify the hypothetical model. Putting the data of every latent variable into the model

to have the first fitting, we found that the regression coefficient of social interaction perception and flow experience is 0.08, the t value is 0.145, and the P value is higher than 0.1 revealing that the social interaction perception has no direct influence on the flow experience. Then we deleted path H4b, and carried out a small adjustment to the residual relationship.

We did the goodness-of-fit analysis of the amended model by fitting every index of confirmatory factor. The results reveal that CMIN/DF is less than 3, PGFI is more than 0.5, index like GFI, AGFI, IFI, TLI, CFI are all higher than 0.8, and RMSEA is smaller than 0.08, showing that all meet the relevant standard. In general, the confirmatory factor analysis model fits well. The model and data also fit well. The subsequent analysis was also tested on the basis of the revised model.

3.4 Hypothetical Test

The results shows that the hypotheses H1a, H1b, H2a, H2b are tested. Also from the perspective of the influencing factors of flow experience, the effects range from high to low for perceived usefulness, perceived ease of use, and information exchange perception ($0.317 > 0.226 > 0.135$), respectively. Thus, hypothesis H3a, H3b and H4a are tested. Flow experience positively affects the satisfaction (0.151) hence, hypothesis H5 is also tested. Similarly, satisfaction influences more to user stickiness than the flow experience ($0.390 > 0.261$), so hypothesis H6 and H7 are tested.

4. Research conclusions and Counter measures

4.1 Research conclusions

The results show that the four dimensions of interactive perception have a significant positive impact on satisfaction, with the greatest effect on perceived usefulness. This means that in the social Q & A community, the user's biggest motivation is to use his own instrumental demand, which is consistent with the product positioning of the Q & A community. Secondly, the impact of perceived ease of use on satisfaction is higher than that of information interaction perception and social interaction perception. The influence of social interaction perception on satisfaction is lower than that of the information exchange, indicating that users are more inclined to exchange information in the community. Information search, browsing and effective response enhance satisfaction. Social interaction perception has a significant positive impact on satisfaction, indicating that interpersonal and emotional communication between like-minded users in the community has a positive impact on satisfaction.

The study of the heart flow experience found that the perceived usefulness has a greater impact over perceived ease of use, which corroborates with the literature. As one of the dimensions of interpersonal interaction perception, information exchange perception has a positive impact on the experience of heart flow. In the context of the social Q & A community, the frequent exchange of information between users, especially the exchange of hidden information, strengthens the depth and breadth of information cognition, resulting in immersion and the state of experience. In addition, research shows that user stickiness is the result variable of satisfaction and heart flow experience, and heart flow experience positively affects user satisfaction.

4.2 Counter measures

At present, the socialized Q & A community is still in the development stage. The empirical results of this paper have a certain guiding role for the marketing practice activities of community operators. We strongly trust that through these results community operators should enhance users' continuous access intention and increase users' stickiness by adopting the following strategies:

(1) Based on the data, we believe that the first thing operators should do is to build the usefulness of the community and focusing on improving the information quality. To solve these problems operators can establish an external control mechanism to control the response quality, integrate internal and external resources, build an expert team, improve professionalism, review existing information through various means and eliminate the useless information.

(2) Secondly, we believe that operators should improve the quality of information system, optimize interface design through easy construction, simplify user operation and enhance user feedback mechanism according to the users' differentiated needs. In addition, when new functions appear in the product, they should be quickly offered to the user to ensure rapid mastery.

(3) Finally, we believe that operators should pay attention to develop users' interpersonal interaction and improve users' perception of information exchange. To improve users' perception of social interaction, operators should, on the one hand, establish cooperative relations with other platforms and expand users' original relationship network; on the other hand, on the basis of protecting users' information security, encourage users' information disclosure and promote the interaction of users in the same industry and across industries.

References

- [1] Morris M R, Teevan J, Panovich K. What do people ask their social networks, and why?: a survey study of status message q & a behavior [C]. International Conference on Human Factors in Computing Systems. 2010 Jan;62(1):1739-1748.
- [2] Kang M, Kim B, Gloor P, et al. Understanding the effect of social networks on user behaviors in community-driven knowledge services [J]. Journal of the Association for Information Science & Technology. 2011 Jan;62(6): 1066-1074.
- [3] Gazan R. Social Q & A [J]. Journal of the Association for Information Science & Technology. 2011 Mar;62(12): 2301-2312.
- [4] Shah C, Oh S, Oh J S. Research agenda for social Q & A ☆ [J]. Library & Information Science Research.,2009 May;31(4): 205-209.
- [5] Shah C, Kitzie V. Social Q & A and virtual reference—comparing apples and oranges with the help of experts and users [J]. Journal of the Association for Information Science & Technology. 2012 Jul; 63(10):2020-2036.
- [6] Chenxiaoyu, Deng shengli, Sun yameng. Progress and Prospects of User Information Behavior Research in Social Question and Answering Communities [J]. Documentation, Information & Knowledge. 2015 Jan;2(4):71-81.
- [7] Hoffman D L, Novak T P. Marketing in hypermedia computer-mediated environments: Conceptual foundations [J]. Journal of Marketing. 1996 Oct; 60(3): 50-68.
- [8] Davis, Fred D, Bagozzi, et al. User acceptance of computer technology: a comparison of two theoretical models [J]. Management Science.1989 May;35(8): 982-1003.
- [9] Kurniawan S H. Modeling Online Retailer Customer Preference and Stickiness: A Mediated Structural Equation Model [C]. Pacific Asia Conference on Information It & society. 2000 Apr;20(8): 238-252.
- [10] Bhattacherjee A. Understanding Information Systems Continuance: An Expectation-Confirmation Model [J]. Mis Quarterly. 2001 Sept;25(3): 351-370.

Webliometric Indicators as Elements of the AI Technique of Estimation of the Language Teacher's Net Proficiency

Matukhin PAVEL^{a,1}, Provotorova ELENA^a, Petrova MARINA^a, Gracheva OLGA^b, Rybakova IRINA^a and Ismailova KHOLISAKHON^c

^aPeoples' Friendship University of Russia (RUDN University), Moscow, Russia

^bInstitute of International Educational Programs of the Moscow State Linguistic University (IMOP MGLU), Moscow, Russia

^cTchaikovsky Moscow State Conservatory, Moscow, Russia

Abstract. The article deals with the issues related to the scientometric indexing of the net resources of the group of language teachers and scientists. A detailed accounting of all types of publications allows us to obtain initial information about the level of their Internet engagement. The analysis was carried out including a wide range of genres of publications, the structure of publications, the format and language of publications, the corpus of academic subjects, the language aspects of the publications under study, the composition of the authors of publications, based on their position. The authors' use of wide opportunities to present their developments on various Internet resources provides them with the opportunity to be detected by search systems based on artificial intelligence (AI) as well as Big Data techniques and to be most fully characterized by existing and prospective scientometric systems.

Keywords. monitoring of publication activity, scientometric indexing, Big Data, artificial intelligence systems.

1. Introduction

Currently, we can observe an intensive growth in the number of theoretical and applied research in the field of development and application of Big Data and Artificial Intelligence (AI) systems for evaluating the performance indicators of scientific and pedagogical staff of higher education, particularly in the field of foreign language teaching (FLT). An integral part of the work of teachers are development of teaching resources, participation in scientific conferences and seminars, regular publication of scientific papers. At the same time, there is publication methods' redistribution in favor of the network placement of the results of the intellectual activity of teachers. One of the tasks of Big Data as well as AI systems is to monitor publication activity. To solve the problems of indexing the plenty of works of authors and their teams, it is necessary to develop a set of indicators based on which the selection of publications for inclusion in a particular database is carried out, and the corresponding indexes are calculated. Also,

¹ Corresponding Author, Peoples' Friendship University of Russia (RUDN University), Moscow, Russia; E-mail: m-pg@mail.ru.

it is necessary to develop the information and technological competence of the authors themselves who create electronic educational resources. In this case, there is a need to create tools that allow authors to search for information posted on the global network and carry out statistical analysis of the obtained data sets [1].

The interest in monitoring and analyzing the publication activity of various groups of authors engaged in the study of specific scientific issues or who are members of the same teaching staff is essential for evaluating the organization's effectiveness and management of the work of the department, faculty.

One example of indexing systems is the Scopus system of ELSEVIER B. V. Corporation. [<https://www.scopus.com>], which is currently the primary source of quantitative information about the publication activity of employees of scientific and educational institutions of the Russian Federation. The main quantitative characteristic of this scientometric system is the Hirsch index which depends on the technical means of distributing scientific publications. It can undergo significant changes over time as a particular work becomes available to different groups of researchers. Such considerations allow us to suggest that the AI systems being developed today will soon form and calculate more complex indicators.

Another example of accounting and monitoring of publication activity is the Digital Object Identifier (DOI) system by the CrossRef. A big plus is a possibility of as-signing such a number not only to the entire scientific work but also to each part. DOI eliminates an error in identifying an information product and facilitates its search in the network, facilitating access to it.

Russia has created its electronic scientometric database – the National Electronic Library [elibrary.ru], combined with the RSCI (Russian Science Citation Index) service, which repeats the principles of the Scopus system in several parameters.

Considering the central systems of scientometric indexing, we can conclude that there is an active development of such services. In addition, the strategic direction of their development is the search for forms of application of AI elements and systems that allow monitoring and evaluating the publication activity of individual authors and author groups to determine their scientific productivity and level.

2. Methods

This study aimed to develop a set of characteristics that describe the level of information and technical competence and Internet engagement of the authors of electronic language educational resources. It is based on a statistical analysis of those information products that they created for use at different levels of language education in the humanities, natural science, and technical direction. At the same time, only published works in electronic format and placed in open access in the databases of electronic publications were studied. It should be clarified that speaking about the Internet engagement of the authors; we mean their focus on the use of tools and resources of the Internet in their work. It also includes the authors' competence in network communications with the editorial offices of electronic publications of various profiles, knowledge of the principles of the device, and operation of the main scientometric systems, which allows authors to develop and implement a targeted policy of promoting their publications. In our article, we rely on the results of a study conducted using the collector_stat v. 1. 0 programs package, which was developed by the staff of the Joint Electronic Resources Fund "Science and Education" (OFERNiO) [2]. This instrument allows selecting from

the database of the fund containing huge electronic resources. Unfortunately, this tool does not allow making complete statistical analysis of data, which reduces the possible potential of the program. Because of this, it was decided to try to combine the features offered by collector_stat v. 1.0 and the MS EXCEL table processor. An array of 561 papers was taken as a base for the study. The team of authors consisted of 90 authors who created these works for 11 years – from 2005 to 2020.

The set of initial research data is further called in work as the webliography of a group of authors. This term is introduced by analogy with the traditional concept of bibliography, which can be considered a subset of webliography, which includes printed works that are not presented in electronic versions on the Internet resources. It should be noted that the webliography contains the full texts of manuscripts and basic information about them, given according to the MARK IV standard. They can be obtained using information search engines when the full texts of publications are not available.

Similarly, we use the term "webliometry" to describe a set of techniques and quantitative characteristics of a webliographic massive obtained by mathematical and statistical methods.

During the research, special attention was paid to the manuscripts of FLT in the field of natural and computer science. The studied webliography corpus of teachers of general and language disciplines includes information products on Latin, Russian as a foreign language, anatomy, and physics for medical students and future information technology specialists. The works included in the webliography of the study were created for use at preparatory faculties for students from Asia, Africa, and Latin America of different departments including medical and agricultural institutes.

Almost all project participants have prepared and applied many new forms of presentation of educational content (electronic notes, presentations, videos, tests, etc.). To provide the complete presentation of the works of each participant, not only traditional types of publications, such as monographs, articles, teaching aids, were included. We use the more general concept of "Information product" to refer to all types of materials created by teachers. By accumulating electronic methodological materials on teaching language and general education disciplines, the teacher can create his educational blogs. Unfortunately, access to such information products is sometimes difficult, and artificial intelligence systems should play a saving role in solving this problem.

It was necessary to form a set of appropriate webliometric parameters based on the obtained statistical data. It aimed to get a general landscape characterizing the level of Internet engagement and the richness of the tools' range used by the authors of electronic teaching and scientific publications. We have selected such indicators that allow us to get the most general idea of various forms of publication activity of project participants. Here we present a few of the main ones from among these characteristics.

3. Results and discussions

We took a corpus of 561 publications for scientific, methodological, and educational purposes to conduct a statistical study. The authors of these works are 90 employees of educational institutions. First, the structure of publications was considered, describing the number of publications with each participant's participation in the project. We have analyzed separately Russian-language and English-language works. At the same time,

all the works of the study participants that are contained in the OFERNiO database were considered [3].

These collections have a structure resembling a cone. The analyzed array of webliography is based on the works of 5-6 authors, which are characterized by the largest number of publications included in the study sample. This group of authors makes up no more than 10 percent of members of the sample, but the number of publications of these authors per person is the largest. It can be concluded that these authors show the greatest creative activity and can be considered the initiators and organizers of creative and educational projects. Above them are the names of authors who participated in three and two creative projects. The number of these authors is also quite small. The most significant group consists of authors whose works are presented in the database only in the singular. It does not mean that these authors have no other manuscripts at all. It indicates that other works of these authors were not included in the studied corpus of publications (which again indicates the need to create comprehensive AI-based search algorithms that will allow, in the absence of a single depository, to build the complete array of webliography as a set of publications from different sources containing the necessary information about the publications of a particular author).

Table 1. Genres of IP.

Genre	Number	Percentage
article	107	19,1%
talk abstracts	88	15,7%
annotation	45	8,0%
IC (information card)	45	8,0%
report	45	8,0%
ATD USPD)	45	8,0%
ATD copy	45	8,0%
ATD text	45	8,0%
ERR certificate)	45	8,0%
TA (teaching aid)	28	5,0%
EMM	18	3,2%
contributions	3	0,5%
guide	2	0,4%

The next stage of the analysis was structuring the webliography and getting an opportunity to determine which type of publication is the most common. Analyzing the sample available to us, we determined that the project participants used a wide range of genres of electronic scientific and methodological publications presented by them on the website of the OFERNiO resource. The list of genres is given in table 1. ATD USPD means advertising and technical description, a unified system of program documentation, ERR - electronic resource registration, EMM - educational and methodical manual.

The analysis of academic subjects presented in the studied corpus of scientific and methodological IP is of interest. The list of 22 items includes Language disciplines (English, Latin, Russian, and Russian as a foreign language), Information disciplines (programming, computer science), Natural science disciplines (biology, physiology).

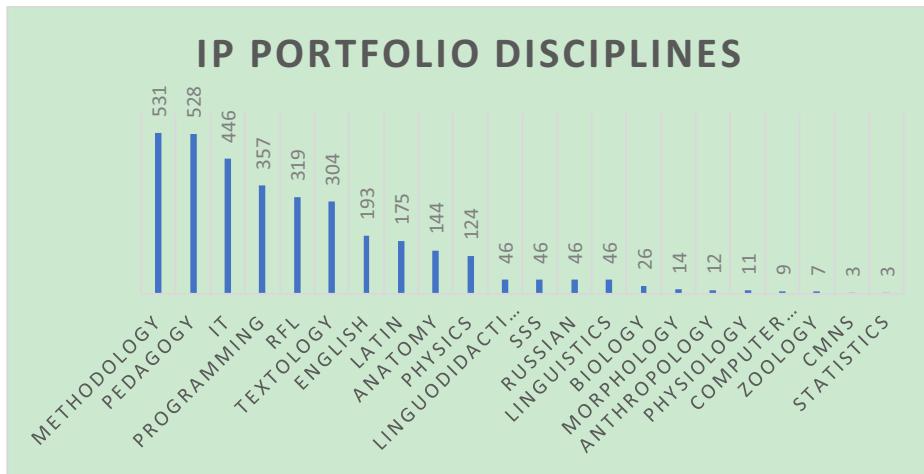


Figure 1. The structure of the IP corpus academic disciplines.

Most of the scientific and methodological papers related to general pedagogy and teaching methods are presented. These works contain various examples of the author's techniques. Works on information technologies and programming represent the second largest group. A significant part of the studied works belongs to the field of language disciplines. These are both languages themselves (Russian and foreign), as well as linguistics and linguodidactics. Here SSS means the scientific style of speech, RFL - Russian as a foreign language, CMNS - concepts of modern natural science).

The diagram shown in Figure 1 gives a visual representation of the structure of the webliographic array. There are several main groups in the spectrum. These general methods and pedagogics form the basis of the pyramid. It characterizes the group of developers as specialists focused not only on creating manuals but also on publicly presenting their methods and pedagogical technologies for the study, analysis, and application by their colleagues. Information technologies and Programming form the next tier of the structure. It reflects the orientation of IP developers to use modern tools in their manuals. This group also includes Testology, the achievements of which are considered in the preparation of electronic forms of training and control-measuring components of educational and methodological complexes. A particular group is formed by language disciplines (RFL, Russian and Latin) [4]. The scientific publications on the Russian language are included in the same group in a separate line. At the same time, these publications can also be attributed to the group of disciplines of the scientific and methodological profile, which includes Linguodidactics, Linguistics. Publications on these disciplines, in turn, form the basis for academic subjects that are taught at the preparatory faculty for foreign citizens and the initial bachelor's degree courses. The latter also includes such disciplines as Morphology and Anthropology.

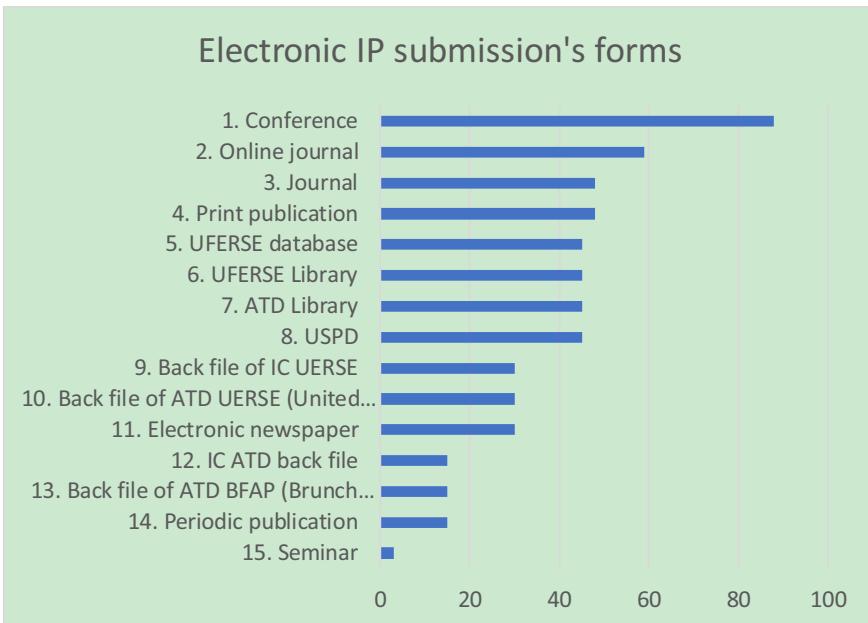


Figure 2. Forms of submission of electronic IP.

The most popular format of publications was conference materials in abstracts and articles, which occupy 18 % of publications. The second most popular is articles in scientific journals, which make up 11 %. The next ones are publications directly in OFERNiO. In addition, the materials of scientific seminars are the least represented. An analysis of the types of educational products created by the project participants was carried out. It should be noted that in most cases, the publication could be assigned to different categories. All of them were identified as informational, most of them as language and profile-specific. A significant part can be considered as soft-ware and algorithmic materials and electronic means of educational and methodological support of the learning process. Forms of the submission of the electronic IP are listed in diagram shown in Figure 2.

The works are mainly presented in two languages – Russian and English. Out of the total number of authors (90), 50 people took part in writing joint works in English, that is, more than half, while the number of English-language works is only 9 percent. Many joint works explain this discrepancy in indicators. In addition, these works are intended mainly for students of Russian universities who are studying in Russian. Accordingly, the demand for Russian-language works is much higher than for English-language ones. The presence of English-language publications is probably due to the need to be published in foreign journals, which can also confirm a sufficiently high level of authors of publications.

4. Conclusion

Summing up the above analysis of the overall picture and details of the study of several indicators of Internet engagement by a group of authors-developers of the considered corpus of webliography, we can draw a crucial key conclusion.

A detailed account of publications allows us to identify the main nuances and shades of the spectrum of special and general information technology tools that were used by the developers of electronic IP, which allows us to obtain initial information about the level of their Internet engagement. The authors' use of wide opportunities to present their developments on various Internet resources provides them with a chance to be detected by AI-based search systems and to be most fully characterized by new scientometric systems based on an integrated approach to evaluating the webliography of individual authors and author groups. The systematic registration of the results of developments in structures like OFERNIO not only allows each of the authors to present their results in a wide range of genres of scientific and pedagogical publications. When forming indicators of the publication activity of teachers, presented in the form of electronic IP, it is advisable to combine the data array of collector_stat-type systems based on a sample from the OFERNIO database with the works of authors posted on other platforms. The use of complex characteristics and statistical methods of their analysis makes it possible to establish the presence and determine the relationship between different types of information products and different groups of authors [5, 6].

References

- [1] Galkina AI, Matukhin PG, Matyash GA, Provotorova EA. State and cooperative resources as a means of access to information and methodological support for educational communication. Proceedings of the 13th International Conference-Exhibition Information Technologies in Education; Moscow: LLC NPP "BIT pro"; 2013. p.30-31.
- [2] Galkina AI, Grishan IA. The program of Collector_stat v. 1.0 - No. AAAA-A17-117041810029-2 dated 23.10.2017; Moscow: CITIS; 2017.
- [3] Portal of the United Fund for Electronic Resources of Science and Education, URL <http://ofernio.ru/portal/newspaper.php>, last accessed 2021/06/25.
- [4] Matukhin PG, Galkina AI, Gracheva OA, Komissarova NV, Zaporozhets VV, Saltykova OV. Reflexive portfolio of a team of authors - developers of information products of the FTP "Russian language 2016-2020". Educational Technologies and Society. 2018; 4:360-372.
- [5] Galkina A and Grishan I. The electronic portfolio of a research university in the mirror of webliometrics. Bulletin of Science and Practice. 2021; 7(5):441-450.
- [6] Galkina AI, Burnasheva EA, Grishan IA, Kadyrova EA. Electronic reflective portfolio of a university teacher in diagrams, tables, graphs. Educational Technology and Society. 2018; 21(2):500-514.

Sequence-in-Sequence Learning for SOH Estimation of Lithium-Ion Battery

Thien PHAM^a, Loi TRUONG^a, Mao NGUYEN^b, Akhil GARG^c, Liang GAO^c, and Tho QUAN^{a1}

^a*Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Vietnam National University Ho Chi Minh City, Vietnam*

^b*Faculty of Computer Science, Ho Chi Minh City University of Computer Science, Vietnam National University Ho Chi Minh City, Vietnam*

^c*School of Mechanical Science and Engineering, Huazhong University of Science and Technology Wuhan, China*

Abstract. State-of-Health (SOH) prediction of a Lithium-ion battery is essential for preventing malfunction and maintaining efficient working behaviors for the battery. In practice, this task is difficult due to the high level of noise and complexity. There are many machine learning methods, especially deep learning approaches, that have been proposed to address this problem recently. However, there is much room for improvement because the nature of the battery data is highly non-linear and exhibits higher dependence on multidisciplinary parameters such as resistance, voltage and external conditions the battery is subjected to. In this paper, we propose an approach known as bidirectional sequence-in-sequence, which exploits the dependency of nested cycle-wise and channel-wise battery data. Experimented with real dataset acquired from NASA, our method results in significant reduction of error of approximately up to 32.5%.

Keywords. Lithium-ion Batteries, BiLSTM, Auto Regression.

1. Introduction

Nowadays, *Lithium-ion Batteries* (LIB) have been used in many electronic devices as the main power supply to avoid power interruption. Especially, the recent growing quantity of moving objects explodes the demand for LIB. In particular, LIB has been used enormously due to its advantages of durability, stability, high-capacity, low-cost, light-weight and small-scale. In another approach, users need to have recent batteries monitoring data during using them. This demand comes from the need to lookahead malfunction, foresee the risks, avoid explosion or optimize the battery life. Therefore, it is essential to predict LIB capacity accurately and promptly in order to provide the exact input for above demands, making the issue of *State-of-Health* (SOH) estimation for LIB highly essential to maintain proper functioning [1].

¹Corresponding Author: Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Vietnam National University Ho Chi Minh City, Vietnam; Email: qtTho@hcmut.edu.vn.

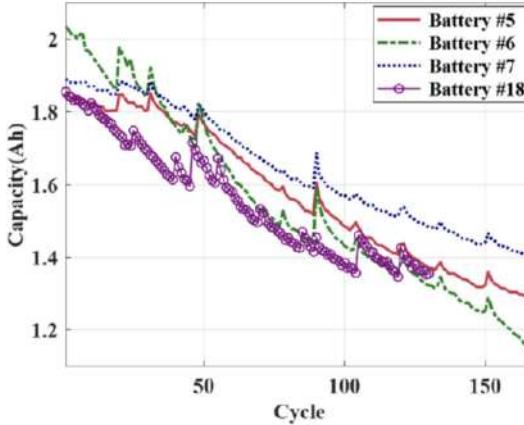


Figure 1. Capacities of real LIBs, in Ampere Hour (Ah), recorded during charge cycles increases [3].

Basically, SOH is a metric to reflect the condition of a battery, compared to its ideal conditions. In the actual lifetime, the capacity of a battery is normally reduced during the discharge cycle and needs a charge cycle afterward. These charge and discharge cycles, empirically, influence the battery SOH over a long time. There are a number of metrics used for SOH estimation of a battery. Perhaps the most popular one is introduced in [2], where SOH is calculated as $SOH = \frac{C_{full}}{C_{nom}} \times 100\%$, where C_{nom} (*Nominal Capacity*) is the original capacity given when the battery was manufactured and C_{full} (*Full Capacity*) is the full capacity practically charged at the current cycle. Theoretically, C_{full} is gradually degraded after every charge and discharge cycles and the SOH estimation task is thus equivalent to that of C_{full} prediction, given the current cycle of the battery. However, this estimation is a non-trivial exercise due to the nonlinear nature of battery capacity change one practically used. Moreover, different batteries introduce different rates of SOH decay. Figure 1 shows an example of SOHs measured from four batteries [3]. It can be observed that the SOHs of the cycles create a time-series trend and there are clear degradations with various rates over around 150 cycles of charge and discharge of all the four modules. Moreover, the batteries C_{full} have not always strictly decreased but occasionally increased in a periodical manner, making *spikes* over the trends.

Research in this field shows that even though each battery possesses a different decay rate, historical data of the energy status of the battery, can be used to predict the energy status in the next few cycles [4]. With the recent advancement of AI methods, much work considers leveraging sequence-based *deep learning* models to handle this issue. In [5], a model using importance sampling and neural networks is introduced. Perhaps the most widely used sequence-based deep learning model for this problem is based on the well-known deep learning architecture of *Long Short-term Memory* (LSTM) [6], [7], [2], which was applied in various works with remarkable results [5], [6].

Traditionally, in LSTM-based approaches, the previous information of the battery capacity in some last cycles will be used to predict the corresponding capacity statuses in the next cycles. However, some additional information from other *channels*, such as voltage, current and temperature, can be complementarily used for insightful analysis of the battery status. Thus, the collected channel-wise data of a battery become those of

multivariate time series [2]. In [8], a multi-channel charging model is proposed for the prediction task. Other partial data also used in [9] with similar purpose.

Notably, perhaps the most remarkable achievement so far is reported in the work of [2]. In this work, the authors proposed an importance sampling strategy of 11 points in a charge cycle. To construct a predictive model, other channel-wise information, as above discussed, has been incorporated to introduce a multivariate time series data model, which is then eventually handled by an LSTM architecture. This work is considered state-of-the-art now in this area.

Thus, in this study, we exploit an observation that not only the cycle-wise relationship during a lifetime of a battery, but also its internal channel-wise information. To realize this idea, we suggest a novel architecture, known as *sequence-in-sequence* (seq-in-seq) model, where the battery energy is analyzed during two nested sequences, including local sequence and global sequence, which reflect the channel-wise and cycle-wise information respectively. Moreover, we also present a combined architecture of Bi-LSTM [10] and *Autoregression* [11], applied in each local and global sequences, to better represent their sequential relationships in different views, resulting in our ultimate model, socalled *BiAR-SeqInSeq*.

In the rest of paper, we summarize preliminaries in Section 2. The overall architecture and details of BiAR-SeqInSeq model are given in Section 3. Section 4 discusses our experimental results when verifying our model with real NASA battery dataset. Finally, Section 5 concludes the paper.

2. Preliminaries

2.1. Charge Cycle of a Lithium-ion Battery

In common working conditions, a LIB lifetime spans up to 2-3 years or equivalently, 300-500 charge and discharge cycles. In the prior cycle, LIB internal channels (current, voltage, temperature) vary and LIB capacity increases gradually to reach a fully-charged state. This charge cycle can be roughly divided into 2 separate periods: *Constant Current* (CC) one with gradually increased voltage and *Constant Voltage* (CV) one with gradually decreased current, as depicted in Figure 2. A charge cycle starting with CC period as the current is kept unchanged while voltage is gradually increased. By the end of CC period, the charge cycle is continuously controlled under CV period as the current is gradually decreased and the voltage remains stable. This CC-CV process is completed when the current reaches a threshold constant value [12]. The battery then goes for a later discharge phase to accumulate SOH.

In [4], there are 5 crucial features that can be extracted during this CC-CV process to infer battery lifetime as follows (Figure 2).

- x_1 : Initial charge voltage
- x_2 : CC charge capacity, calculated by accumulation the rectangle $ICCT$, where I_{CC} is the constant current in the charge, and t the duration of CC period
- x_3 : CV charge capacity, accumulated by the integral of the area of decreasing current during CV period
- x_4 : final charge voltage (voltage measured at battery terminal)
- x_5 : final charge current (current measured at battery terminal)

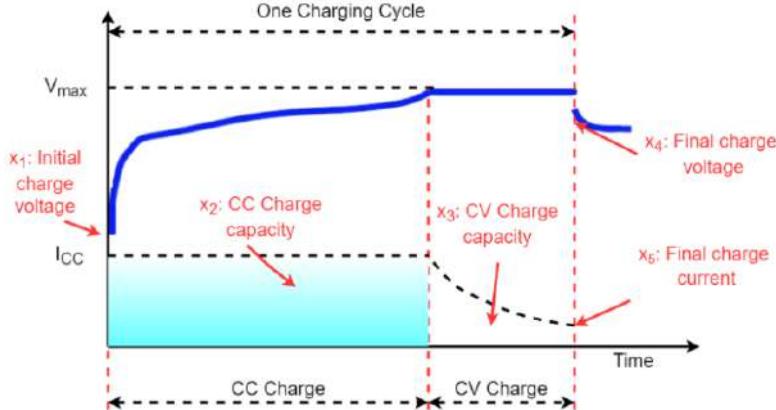


Figure 2. Charge cycle and handcrafted features in that cycle [4].

Obviously, those x_1 values are usually unstable considering many different cycles. Moreover, x_1 is taken in the non-sampling area [5] and its value practically plays a very small impact on the output. Therefore, in this study, we will ignore x_1 contribution in the calculation. The other 4 values x_2, x_3, x_4, x_5 are encoded as *cycle-wise* data to predict the battery SOH, as subsequently discussed.

Additionally, we also adopt the information of *peak* status, which represents that the SOH value of the current cycle is staying at the peak comparing to its neighbour using z-scores or not [13] as a cycle-wise information, alongside with the C_{full} information of the considered cycle.

2.2. A Baseline Sampling-based Model

Obviously, current and voltage are essential in the learning model to predict battery lifetime or SOH. However, these values are continuous and one would naturally consider sampling them. In [2], temperature values are additionally sampled at these timepoints to create multi-channel input for their prediction. Those attributes of current, voltage and temperature introduce channel-wise information of a LIB.

Figure 3 depicts the channel-wise raw values of a battery. Moreover, in [5], the authors proposed 11 important timepoints in a charge cycle to extract the sampling current, voltage and temperature as depicted. The authors explained the need to maintain important data points which carry the shape of the derivative ΔV at 11 positions in the charge cycle. Those 11 points of corresponding to the 35th, 52th, 70th, 88th, 90th, 92th, 94th, 96th, 98th, 99th, 100th produce the corresponding sampled channel-wise data.

Figure 4 illustrates how channel-wise data are encoded in [2]. The sampled values of voltage V_j , current I_j and temperature T_j at the above mentioned 11 data points are formed together as 11×3 matrix $\xi_n = [[V_1, I_1, T_1][V_2, I_2, T_2], \dots, [V_S, I_S, T_S]], S = 11$. The authors also use the C_j , which is the SOH value at the j^{th} cycle, as other additional information in their final encoded feature vectors of a cycle. The features vectors of the past cycles are then fed into a LSTM model to make the prediction for the next cycle, as subsequently discussed. In our study, we consider this model as the baseline model in our subsequent experiments.

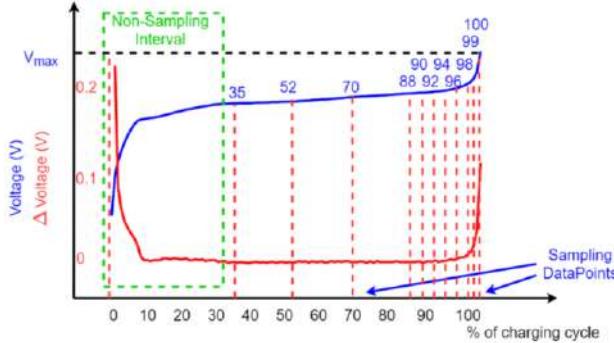


Figure 3. Sampling important values of current and voltage during a charge cycle [5].

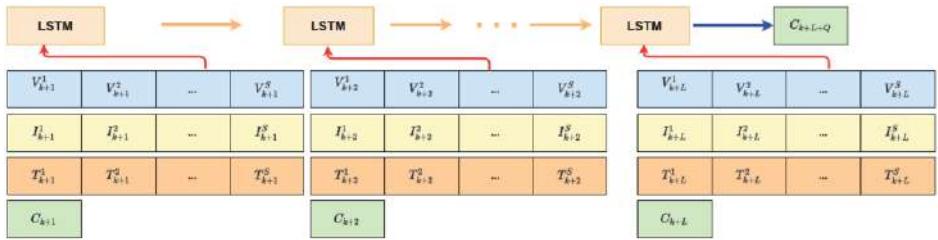


Figure 4. Sampling-based features for SOH estimation [2].

2.3. Long Short-Term Memory

As discussed, these sequences of feature vectors over cycles will be used to train LSTM [14], which is virtually a sequence of multiple logical cells, unfolded from a physical cell. Figure 5 presents an architecture of a basic LSTM cell, which was introduced in [14]. LSTM is in fact an upgraded version of Recurrent Neural Network (RNN) [15]. At the time t , this network will produce output $y(t)$ from the input $x(t)$. However, the output of this network at the previous iteration will also be used as part of the input of the next step, or recurrent input, together with new actual input. In the standard LSTM model, processing information is more complicated when modules containing computational blocks are repeated over many timesteps to selectively interact with each other in order to determine what information will be added or removed. This process is controlled by three gates namely *input gate*, *output gate*, and *forget gate*. Controlling the flow of information inside an LSTM model can be described as the following mathematical equations:

$$i = \sigma(W_i x_t + W_{hi} h_{t-1} + b_i) \quad (1)$$

$$f = \sigma(W_f x_t + W_{hf} h_{t-1} + b_f) \quad (2)$$

$$o = \sigma(W_o x_t + W_{ho} h_{t-1} + b_o) \quad (3)$$

$$(C)_t = (C)_{t-1} \otimes f + i \otimes \tanh(W_C x_t + W_{hC} h_{t-1} + b_C) \quad (4)$$

$$h_t = o \otimes \tanh((C)_{t-1}) \quad (5)$$

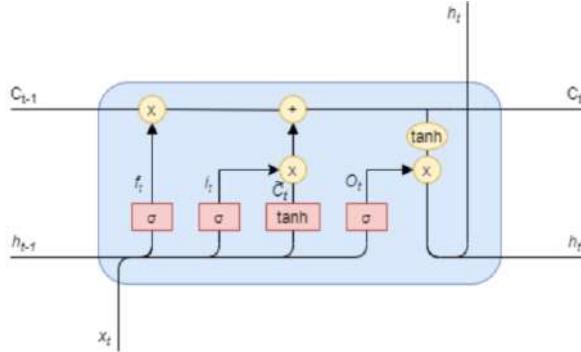


Figure 5. A basic LSTM cell [14].

In Eq. (1) - Eq. (5), $i, f, o, (C), h$ denote input gate, forget gate, output gate, internal state, and hidden layer, respectively. Here, W_i, W_f, W_o, W_C , and b_i, b_f, b_o , and b_C represent the weights and bias of three gates and a memory cell, in the order given. Concretely, the activation function *sigmoid* helps an LSTM model to control the flow of information because the range of this activation function varies from 0 to 1 so that if the value is 0 then all of the information will be cut off, otherwise the entire flow of information will pass through. Similarly, the output gate will allow information to be revealed appropriately due to the sigmoidal activation function then the weights will be updated by the element-wise multiplication of output gate and internal state activated by non-linearity *tanh* function.

In the LSTM neural network architecture proposed by [2], input vector x_i at each timestep is feature vector of each cycle as described above. This LSTM architecture uses L previous cycle to predict C_{full} of next cycle. As discussed, this work has enjoyed the best performance so far in terms of SOH prediction.

3. The BiAR-SeqInSeq Model

In this section, we present our proposed BiAR-SeqInSeq model. For the sake of convenience, technical details of our model is presented in separate subsections. Firstly, we present the *Autoregression* technique. Then, we present our feature encoding strategy for each charge cycle as the premise for our next discussion about the sequence-in-sequence bidirectional architecture. Finally, we present how to combine these techniques into our ultimate model.

3.1. The Autoregression Technique

In statistics, econometrics and signal processing, an *Autoregression (AR)* model is a *random process*, which specifies that the output variable depends linearly on its own previous values and on a stochastic term. Thus, the model can be considered as a form of a stochastic difference equation evaluated on the values from the past.

In general, given the value $C_{k+1}, C_{k+2}, \dots, C_{k+L}$ where C_j is the value the j^{th} time-point and C_{k+L} the current time, the value of the next value C_{k+L+1} can be evaluated as Eq. (6)

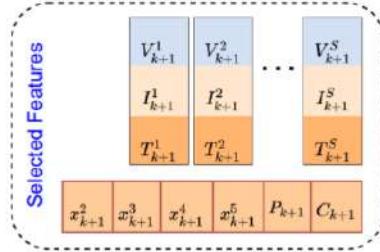


Figure 6. The features selected inside a cycle (channel-wise features).

$$C_{k+L+1} = c + \sum_{j=1}^L \varphi_j C_{k+j} + \varepsilon_{k+L+1} \quad (6)$$

where c is a initial constant (or bias), φ_j is the corresponding coefficient of C_{k+j} and ε_{k+L+1} is a random noise. There are different methods to evaluate φ_j , one of which is to approximate them by means of a neural network, as we further discuss and apply in our works.

3.2. Cycle Encoding

Extended from [2], we rearrange 3 channel information inside each cycle into 11 vectors $\xi_j = (V_j, C_j, T_j); j = 1..11$, corresponding to 11 sampling time steps as illustrated in Figure 6. Additionally, in each cycle, we also additionally extract an encoded vector $E_V = (x_2, x_3, x_4, x_5, peak, C)$, where the meaning of $x_2, x_3, x_4, x_5, peak, C$ are already discussed in Section 2.1. ξ_j and E_V will be used in our sequence-in-sequence training mechanism in later discussion.

3.3. Bidirectional Seq-in-seq Training Strategy

With feature vectors described above for each cycle, the features are encoded as the input using the sequence-in-sequence mechanism to train our predictive model as illustrated in Figure 7. We call this mechanism sequence-in-sequence because the gradients flow through two nested sequences during the training phase: local sequence and global sequence, or *channel-wise* and *cycle-wise* sequences.

3.3.1. Bidirectional LSTM

In both channel-wise and cycle-wise sequences, data-points are trained in forward and backward directions between time steps, hence the name bidirectional. Recall that in [2], the author used LSTM architecture to learn from sequential data. This LSTM is unidirectional, which reads input data from left to right order. In this research, we use another variation of LSTM, known as Bidirectional LSTM, or Bi-LSTM [10] to better capture sequential relationships in each sequence. As illustrated in Figure 7, Bi-LSTM neural network composed of two independent LSTM networks in opposite directions, w.r.t the order of the input data-points. The output hidden vectors to those LSTM manner are then concatenated in a cell-wise manner, producing the final output of the Bi-LSTM network.

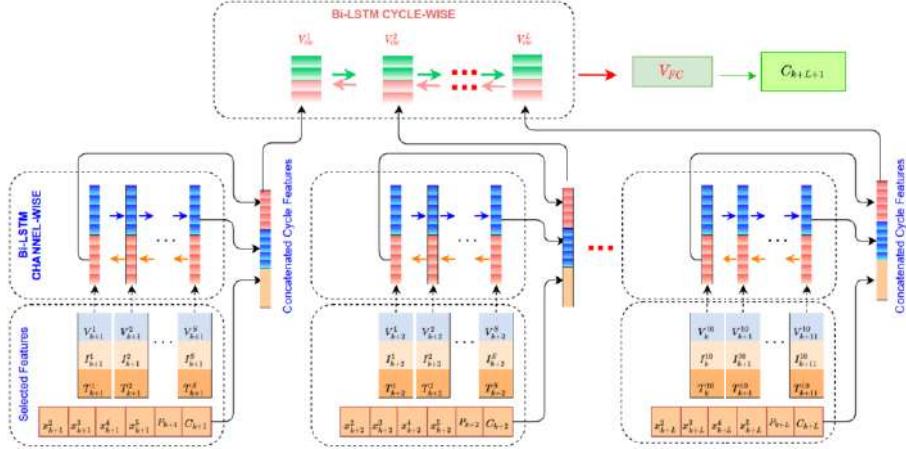


Figure 7. The Bidirectional seq-in-seq training strategy.

3.3.2. Channel-wise sequence

Channel-wise sequence is equivalent to a Bi-LSTM neural network whose input are 11 vectors $\xi_j = (V_j, C_j, T_j)$; $j = 1..11$ above discussed. The final output of this Bi-LSTM is then concatenated with the vector $E_V = (x_2, x_3, x_4, x_5, peak, C)$ as the final output of this processing step. We denote this result vector as a channel-wise feature vector V_{cw}^j .

3.3.3. Cycle-wise Sequence

Cycle-wise sequence is also a Bi-LSTM neural network, whose inputs are channel-wise feature vectors V_{cw}^j , $(j = 1..11)$, encoded from $L = 11$ consecutive cycles in the past, after processed in the channel-wise sequence step. The final output of this global BiLSTM, denoted as V_{FC} , will flow through a fully connected layer to predict SOH of the next cycles.

3.4. The ultimate BiAR-SeqInSeq Model Architecture

Figure 8 captures our ultimate model, in which bidirectional seq-in-seq architecture is combined with Autoregression technique to improve prediction results. We enhance our prediction by combining the two results from two computational branches, known as Non-linear Branch and Linear Branch, whereas Non-linear Branch output is the above mentioned V_{FC} , now is referred as C_{non_linear} . Meanwhile, in Linear-Branch, we apply Autoregression on the SOH values of the L processed cycles ($C_{k+1}, C_{k+2}, \dots, C_{k+L}$) to obtain the $C_{regression}$ result (Eq. (7)). Finally, the C_{non_linear} will be added up $C_{regression}$ based on the trainable combined parameter as the below Eq. (8).

$$C_{regression} = \sum_{j=1}^L \varphi_j C_{k+j} \quad (7)$$

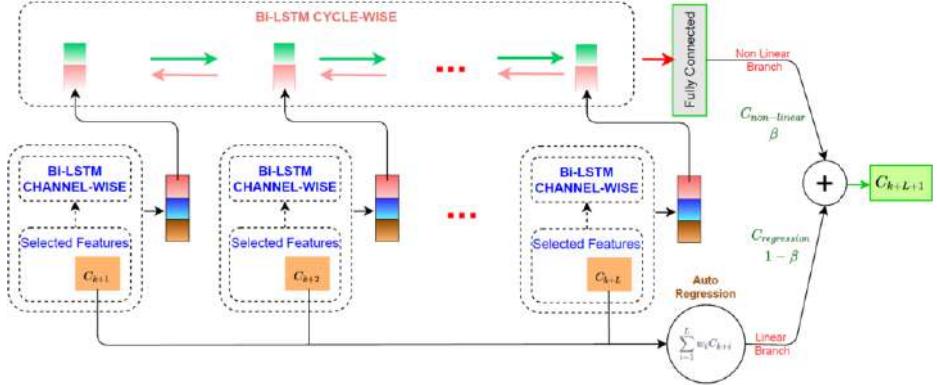


Figure 8. The ultimate BiAR-SeqInSeq mode.

$$C_{k+L+Q} = \beta C_{non_linear} + (1 - \beta) C_{regression} \quad (8)$$

As discussed, the weights φ_j and β are trainable parameters optimized in the training of the whole model to achieve the best prediction performance.

4. Experiments

In this section, we describe the experiments to confirm the performance of our proposed BiAR-SeqInSeq model. The datasets of 4 NASA batteries and the cross-validation techniques used in [2] are inherited. The quantitative comparisons will be given as below.

4.1. Dataset

Dataset used in this research are provided by NASA [16], including 4 batteries named B0005, B0006, B0007, B0018. The first three battery modules provide full 168 cycles (charge and discharge) and the last B0018 123 cycles. These dataset specification is presented in Table 1. Test conditions of these batteries are presented in Table 2.

Table 1. Summary of 4 NASA battery modules over the cycles.

Battery No.	Battery Structure	Length of Capacity Data
B0005	1×616	168
B0006	1×616	168
B0007	1×616	168
B0018	1×319	123

Table 2. Test condition of NASA LIB dataset.

Battery No.	Constant Charge Current(A)	Charge Cut-off Voltage(V)	Discharge Current(A)	Discharge Cut-off Voltage(V)
B0005	1.5	4.2	2.0	2.7
B0006	1.5	4.2	2.0	2.5
B0007	1.5	4.2	2.0	2.2
B0018	1.5	4.2	2.0	2.5

We used all 4 batteries historical data to develop the models, in which 80% of the dataset was used for training and the remaining 20% was used for testing.

4.2. Evaluation results

In our experiments, we develop the following models.

- LSTM-based Multi-Channel Prediction [2]: As discussed, this is the current SOTA performance of SOH estimation problem.
- LSTM-based Multi-Channel Prediction: based specifically on [2], apply basic fine-tuning and noise-removal techniques to record the output. We have no seq-in-seq and cycle-wise features in this experiment.
- Con2DLSTM: Another additional model we develop for experiment. This architecture uses 2D CNN to extract cycle features and use LSTM to predict [17].
- Bi-SeqInSeq NoAR: similar to above BiLSTM-seq-in-seq-Autoregression but without Autoregression is applied.
- BiLSTM-seq-in-seq-Autoregression (BiAR-SeqInSeq): This is the model we presented.

Experiments results are summarized in Tab. 3. We can see that the baseline LSTM-based Multi-Channel [2] method is still outperforming deep-learning based time series analysis methods such as its baseline. However, models Conv2D_LSTM and Bi-SeqInSeq results even better than the baseline due to their capability to capture both local features and global features. Besides that, Autoregression technique helps to coordinate channel-wise and cycle-wise information most reasonably by trainable parameters learnt by back propagation mechanism when training the end-to-end model, which also significantly contributes to our best performance enjoyed.

Table 3. Summary of performance outcome.

Index	Methods	RMSE	Reduction outcome
(1)	LSTM-based Multi-Channel [2]	0.0268 - 0.030	
(2)	LSTM-based Multi-Channel (No multi-task, no Conv2D)	0.040	
(3)	Conv2D_LSTM [17]	0.033 - 0.034	Reduce 17.5% compare to (2)
(4)	Bi-SeqInSeq NoAR (Without Autoregression)	0.036	Reduce 10.0% compare to (2)
(5)	BiAR-SeqInSeq	0.027	Reduce 10% compare to (1) and 32.5% compare to (2)

The experiment result on Tab. 3 shows that the overall loss is reduced approximately 10% from 0.030 (reported in [2]) to 0.027 ($(0.030-0.027)/0.030 = 10\%$) compared between LSTM-based Multi-Channel to BiAR-SeqInSeq. Moreover, if we compare with the result of the same method re-implemented and worked on the same data set, the loss is reduced from 0.040 (LSTM-based Multi-Channel) to 0.036 (Bi-SeqInSeq NoAR, reduce $((0.04-0.036)/0.040 = 10\%$) and to 0.027 (BiAR-SeqInSeq, reduce $((0.040-0.027)/0.040 = 32.5\%$), which are significant.

5. Conclusion

We have proposed some set of improvements on a SOTA method in Lithium-ion Battery SOH prediction using LSTM-based deep learning model. Based on the deep study of LIB data, we have developed a sequence-in-sequence advancement. Moreover, thanks to the dependency on cycle-wise SOH time-series, we keep the Autoregression portion to play its role in final prediction.

The improvements originate from the deep understanding of the nature seq-in-seq in LIB data. Before our study, the existing LIB researches have not fully investigated the dependency. Therefore, we create deep improvement on the outcome. The experiment result shows that the proposed method creates improvement in almost all cases and the ultimate model BiAR-SeqInSeq definitely gave significant improvement.

From the promising result presented, we will continue data-driven investigation, applying more natural enhancements into the model. The current method, still, shows some shortcomings at the discrete 11 datapoints sampling approach. Moreover, the extracted features have room to apply an attention mechanism. As presented in [18], the attention model can extract the sequence contribution and prioritize the features contribution into the final output. We will apply the above ideas into our future research.

Acknowledgement

This research is funded by Vietnam National University HoChiMinh City (VNU-HCM) under grant number C2021-20-13.

References

- [1] Xiaodong X, Chuanqiang Y, Shengjin T, Xiaoyan S, Xiaosheng S, Lifeng W. Remaining useful life prediction of Lithium-Ion batteries based on wiener processes with considering the relaxation effect. *Energies*, MDPI, Open Access Journal, vol. 12(9), pages 1-17; 2019.
- [2] Kyungnam P, Yohwan C, Won JC, Hee-Yeon R, Hongseok K. LSTM-Based battery remaining useful life prediction with Multi-Channel Charging Profiles. *IEEE Access*, vol. 8, pages 20786-20798; 2020.
- [3] Saha B, Goebel K. Battery data set, NASA Ames Prognostics data repository. NASA Ames, Moffett Field, Calif, USA; 2007.
- [4] Chao H, Gaurav J, Schmidt CL, Carrie S, Sullivan M. Online estimation of lithium-ion battery capacity using sparse Bayesian learning. *Journal of Power Sources*, vol 289, pages 105-113; 2015.
- [5] Ji W, Chenbin Z, Zonghai C. An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks. *Journal of Applied Energy*, vol 173, pages 134-140; 2016.
- [6] Yuefeng L, Guangquan Z, Xiyuan P, Cong H. Lithium-ion battery remaining useful life prediction with long short-term memory recurrent neural network. *Proceedings of the Annual Conference of the PHM Society*; 2017.
- [7] Funai Z, Po H, Xiaohui Y, RUL prognostics method based on real-time updating of LSTM parameters. *Chinese Control And Decision Conference (CCDC)*; 2018.
- [8] Yohwan C, Seunghyung R, Kyungnam P, Hongseok K. Machine Learning-Based Lithium-Ion battery capacity estimation exploiting Multi-Channel Charging Profiles. *IEEE Access on Advanced Energy Storage Technologies and Their Applications*, vol 7, pages: 75143 - 75152; 2019.
- [9] Benvolence C, Chunhui C, Lien-Kai C, Wei-Chih S, Mi-Ching T. Long Short-Term Memory approach to estimate battery remaining useful life using partial data. *IEEE Open Access*, vol 8, pages 165419-165431; 2020.

- [10] Alex G, Santiago F, Jürgen S. Bidirectional LSTM networks for improved phoneme classification and recognition. International Conference on Artificial Neural Networks, page 799-804; 2005.
- [11] Guokun L, Wei-Cheng C, Yiming Y, Hanxiao L. Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks Attention for multivariate time series. International ACM SIGIR Conference on Research Development in Information Retrieval; 2018.
- [12] Peter K, Simon FS, Christian von L, Holger H, Raghavendra A, Andreas J. Lifetime analyses of Lithium-Ion EV batteries. 3rd Electromobility Challenging Issues conference (ECI), Singapore; 2015.
- [13] Van Brakel. Z-score based streaming peak detection, 2014 (accessed October 5, 2020). URL <https://stackoverflow.com/questions/22583391>.
- [14] Sepp H, Jürgen S. Long Short-Term Memory. Neural Computation, vol 9, pages 1735–1780; 1997.
- [15] David E.R, Geoffrey E.H, Ronald JW. Learning representations by back-propagating errors. Nature, vol 323, pages 533–536; 1986.
- [16] Nasa dataset online at: Available online: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>. Access Aug, 2021.
- [17] Xingjian S, Zhourong C, Hao W. Convolutional lstm network: A machine learning approach for precipitation nowcasting. Computer Vision and Pattern Recognition; 2015.
- [18] Ashish V, Noam S, Niki P, Jakob U, Llion J, Aidan N.G, Lukasz K, Illia P. Attention is All you Need. Computer Science; 2017.

A Pointwise Evaluation Metric to Visualize Errors in Machine Learning Surrogate Models

Seyed Shayan SAJJADINIA ^{a,1}, Bruno CARPENTIERI ^a and
Gerhard A. HOLZAPFEL ^{b,c}

^a*Department of Computer Science, Free University of Bozen-Bolzano, 39100 Bozen-Bolzano, Italy*

^b*Institute of Biomechanics, Graz University of Technology, 8010 Graz, Austria*

^c*Department of Structural Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

Abstract. Numerical simulation is widely used to study physical systems, although it can be computationally too expensive. To counter this limitation, a surrogate may be used, which is a high-performance model that replaces the main numerical model by using, e.g., a machine learning (ML) regressor that is trained on a previously generated subset of possible inputs and outputs of the numerical model. In this context, inspired by the definition of the mean squared error (MSE) metric, we introduce the pointwise MSE (PMSE) metric, which can give a better insight into the performance of such ML models over the test set, by focusing on every point that forms the physical system. To show the merits of the metric, we will create a dataset of a physics problem that will be used to train an ML surrogate, which will then be evaluated by the metrics. In our experiment, the PMSE contour demonstrates how the model learns the physics in different model regions and, in particular, the correlation between the characteristics of the numerical model and the learning progress can be observed. We therefore conclude that this simple and efficient metric can provide complementary and potentially interpretable information regarding the performance and functionality of the surrogate.

Keywords. Visualization, surrogate modeling, numerical analysis, machine learning metric, PMSE

1. Introduction

Numerical methods such as finite element (FE) methods [1] are powerful tools to simulate various physics behavior [2–6], but for some practical applications, they may be computationally prohibitive due to the high nonlinearity and large number of equations involved. This motivates the use of a surrogate model, which is an efficient model that can replace the main numerical model using machine learning (ML) or any other efficient data-driven method [7]. In this regard, a simple idea, among many others, is to first

¹Corresponding author at: Department of Computer Science, Free University of Bozen-Bolzano, 39100 Bozen-Bolzano, Italy; e-mail: ssajadinia@unibz.it.

generate a subset of possible inputs and outputs of the numerical model in order to train the ML surrogate model [8], which now enables real-time simulation for several scientific disciplines, e.g., in medicine [9–13] and in mechanical engineering [14–18]. Although this technique still requires implementation of an expensive numerical solver to create the dataset, over time, the highly efficient surrogate can substitute it for the next simulations.

The mean squared error (MSE) metric is often used to evaluate the performance of such ML methods. Metrics of this type find the average of errors, by some comparison between the predicted and target values, over all the samples' geometrical points that define the physical system [19, 20]. Despite their advantages, especially in ML training, they may not show the progress in learning at each point, or in other words, they are not pointwise in a sense that they cannot distinguish the importance of each point. If the average pointwise errors are not considered, the overall accuracy in different regions of the physical system cannot be shown, and then, we may lose some information about the importance of each point in ML evaluation. To the best of the authors' knowledge, the advantages and feasibility of a pointwise metric in this application area of ML have not yet been explored, and assessing them is our main contribution.

This study aims to present the pointwise MSE (PMSE) metric that efficiently evaluates the surrogates and may elucidate the role of numerical model definition. While related work is reviewed in Section 2, this metric is mathematically defined in Section 3. Next, we try to show its importance with a simple experiment worked out in Section 4, and then some final remarks are provided in Section 5. We shared our research data and code at github.com/shayansss/pmse.

2. Related Work

Global metrics, such as the mean absolute error [11], MSE [20], coefficient of determination [21], are widely used for ML training or model evaluation. Despite their benefits, especially in simple quantitative comparison of the ML and numerical models, they might be barely interpretable [22]. Therefore, it is a common practice to visualize the prediction, comparing to the corresponding testing target, typically by mapping the results onto the numerical model [23–25]. However, they may not replace the local metrics that can enable pointwise visualization of the errors associated with all the samples.

Several studies, e.g., [26–29], used local surrogates or pointwise metrics, e.g., using the local maximum absolute error [27]. Although, some of these local errors, such as the pointwise cross validation error [28], may not be necessarily reliable surrogates for the point errors of test samples, but they still can increase the overall performance, while with possibly higher computational costs [29]. Few studies [30–32] defined pointwise metrics to measure and precisely visualize errors of samples, but they have not been used to study the correlation between the numerical model and these errors. In this paper, we propose a similar pointwise metric, but it is used only for testing to keep the training fast while we use it to especially interpret the relationship between the numerical and ML models.

3. Evaluation Metrics of Geometrical Points

In a typical supervised surrogate model, the ML regressor can predict the outputs of the numerical model extracted at N assigned points evaluated on M test samples. Denoting by $\bar{y}_{m,n}$ and $y_{m,n}$ the prediction and target values of the n^{th} point of the m^{th} sample, respectively, the performance of the model can then be assessed as follows

$$\text{SE}_m = \frac{1}{N} \sum_{n=1}^N (\bar{y}_{m,n} - y_{m,n})^2. \quad (1)$$

Here, SE_m is the squared error of sample m , which is averaged over the whole points. This can again be averaged over all the test samples in order to find the MSE by

$$\text{MSE} = \frac{1}{M} \sum_{m=1}^M \text{SE}_m. \quad (2)$$

To propose the PMSE, we used an equation similar to Eq. (2), but this time, the average is only calculated over the constituent points using

$$\text{PMSE}_n = \frac{1}{M} \sum_{m=1}^M (\bar{y}_{m,n} - y_{m,n})^2, \quad (3)$$

where PMSE_n indicates the PMSE value of point n , which, together with the other errors of the other points, allow us to visualize a contour plot of them on the physical system. This can reveal some new interpretable information about the training performance, considering that we can then see the accuracy of all the geometrical points. Note that having two metrics simultaneously may not affect the computational cost of the surrogate significantly, since, here, we use the PMSE metric only for evaluation of the trained model, rather than training.

This metric may be applicable when the points are attached to the simulated materials in the Euclidean space that should have similar locations at some stage, regularly at the beginning of the numerical simulation, to which the contour can be mapped. In particular, this condition can be realized when the metric is applied to the mesh nodes of a typical FE simulation, where the physical variables are measured.

4. Experiment

4.1. Setup

Physics problem. We are seeking to simulate a group of contact problems on a rectangular 2D sample (with a dimension of $20 \times 50 \text{ mm}^2$) with a circular hole within 2.5 mm in radius (see Figure 1). We assign an incompressible neo-Hookean model to the sample because it is highly used for this type of simulations. Next, we assume that the model, fixed to the bottom, has a hard contact through a semi-circular rigid indenter (with a radius of 10 mm) moving in both x_1 and x_2 directions. See this review study [33] for more about the governing equations of these models.

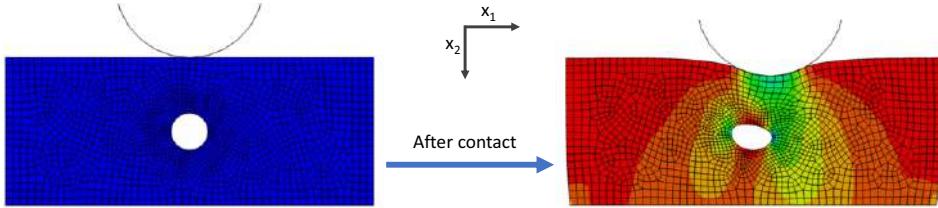


Figure 1. An example of the initial and final configurations of the numerical model. The maximum and minimum calculated values are indicated by a color range from red (165 MPa) to blue (-2750 MPa).

Numerical simulation. The physics problem has complex nonlinear properties, making its numerical simulation time-consuming enough to justify surrogate modeling. By FE numerical modeling, the complex physical system is discretized into a numerical model formed by a mesh of elements with simpler equations that subsequently approximates the physics involved for each point in each element. Since, here, a high-level FE solver is employed, we just need to assign different conditions of the physics problem, as stated in the former paragraph, and then, we run the solver to obtain the simulation data on each point, yielding contour plots of the Cauchy stress in the x_2 direction, see, e.g., Figure 1 (right). These values can roughly be interpreted as the distribution of the contact effect within the sample in the given direction.

Dataset. The input features include the movements of the indenter in the x_1 direction (l_1) and in the x_2 direction (l_2) as well as the material parameter (C). The outputs of the ML model are collected by rerunning the numerical simulation for a wide range of input values sampled by the uniform distribution (U), including $l_1 \sim U$ (-3 mm, 3 mm), $l_2 \sim U$ (0.1 mm, 3 mm), and $C \sim U$ (0.001 MPa, 1000 MPa). Once the data are normalized, 75 and 25 samples are randomly selected for the training and validation sets (used for the ML training), respectively, and 100 samples are inserted into the test set.

Training. A feed-forward neural network with the ReLU activation function is selected, as it is highly utilized for surrogate modeling of FE analysis [20]. While the learning rate is sufficiently small, with a value of 10^{-4} , to record the gradual progress of learning, in order to keep the training still fast, it is implemented by the Adam optimizer [34]. To speed up training even more, we reduce the number of parameters by applying only a single hidden layer that contains one-hundredth less than the number of neurons in the last layer. During training, the model is saved 8 times every 40 epochs for the subsequent evaluation.

Implementation. The numerical simulation is carried out by Abaqus [35], an FE solver, with its hybrid linear plane strain element formulation. This software is controlled by a Python script to generate the data set by assigning the input values to the numerical model and then iterating the simulation. As soon as the data set is created, the ML model is trained by the Keras [36] library with the TensorFlow [37] backend. This is implemented by Jupyter Notebook [38] on an Intel Core i5 CPU and 8 GB RAM. The evaluation metrics are calculated on the test set using the vectorized functions in NumPy [39], and ultimately, the visualization code is implemented using the Matplotlib [40] library along with another Python script to visualize the PMSE in Abaqus. More implementation details are available in the shared GitHub repository.

4.2. Results

The model was trained successfully, and the surrogate worked efficiently, considering that the numerical data generation took 55 min, whereas the training and evaluation of the ML surrogate took only about 2 min. Therefore, if the future repetition of the simulation is desired, the surrogate can guarantee real-time simulation.

The overall performance on the test set is shown in Figure 2 using the SE_m and MSE values during training. While the SE_m box plot provides some additional information about sample outliers, both estimate the overall performance without interpreting the role of the numerical model definition in training.

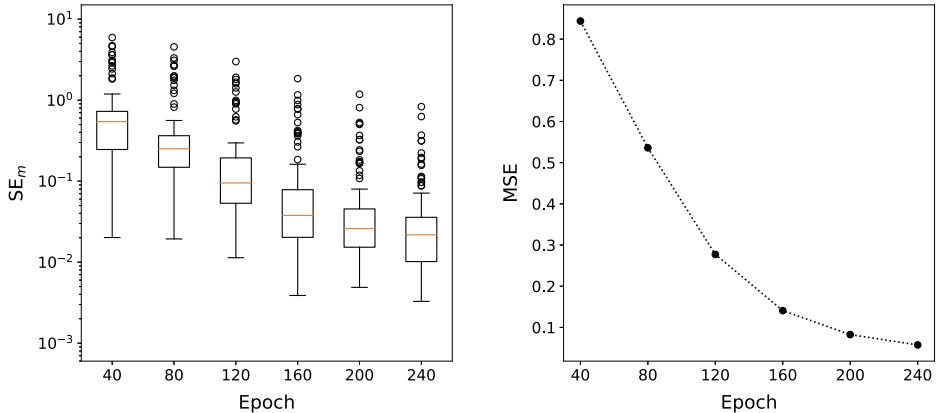


Figure 2. Squared errors of each test sample (SE_m in the left plot) and their mean values (MSE in the right plot) in relation to the number of epochs.

Figure 3 shows the corresponding PMSE contours, mapped to the initial configuration. In the first 40 epochs, the contour is mostly green, which means that most regions are still partially imprecise. These snapshots could indicate the learning progress across the points, which could be of high importance, e.g., while we needed more and more epochs to reach high accuracy (shown in blue), the area around the circular hole was well-trained after 160th epoch. This could be an important basis for the analysis, because in a typical numerical simulation we can prioritize the accuracy in a certain region in order to reduce the computational time.

Other relevant results include: (i) by halving the model geometry into the right and left parts, due to the overall symmetry of the model, we expect similar complexity for each part, which can gradually lead to grossly symmetrical contour errors; (ii) some small zones near the top remained red or green in all snapshots, which could be explained by the fact that they underwent more deformation, caused by different movements of the contacting indenter, and consequently, the output data was more complex; (iii) looking at the last two snapshots, some areas where the elements change from irregular to regular rectangular shapes are not completely blue, suggesting that a more regular element shape can improve training. All of these results demonstrate the potential functionality of the pointwise metric to further interpret the association of the ML performance and the numerical model.

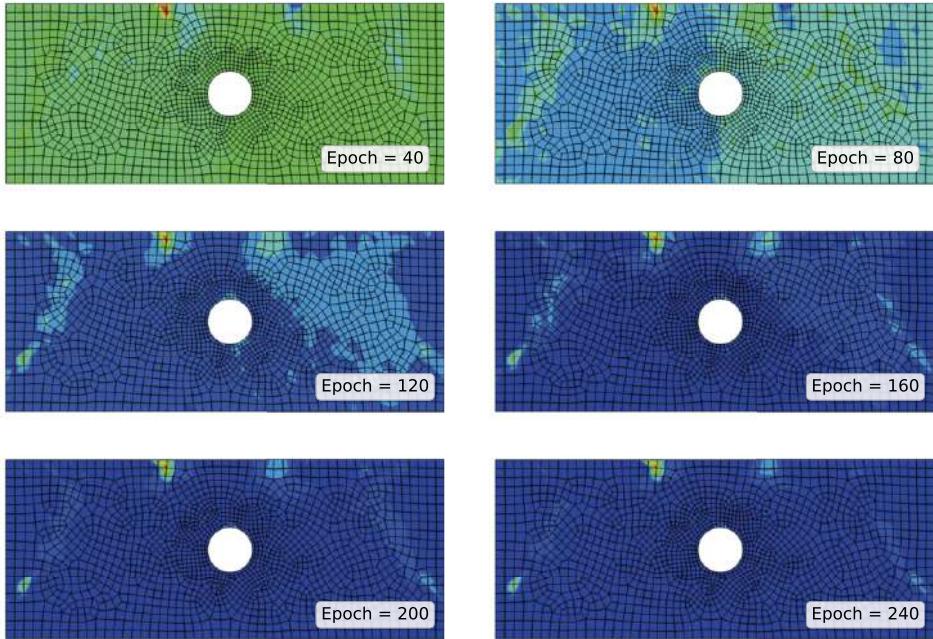


Figure 3. Contour plots of pointwise mean squared errors on the test set with respect to the trained model at different epochs. The contours are mapped onto the initial configuration of the model under contact before deformation. The color range from red to blue shows the values from 2 to 0, respectively.

For the sake of comparison, in Figure 4, we display the absolute error snapshots of a testing sample with the highest error in the first 40 epochs. We see that these errors are unsurprisingly larger than the PMSE values shown in Figure 3, especially on the right side of the plots. The contours do not show the small red zone that is observable in all PMSE snapshots, and they also show significantly less accuracy than the PMSE contours in the elements around the circular hole, particularly at late epochs, probably because of the large displacement boundary conditions contained in the input values. Therefore, the proposed metric proves to be clearly more generalizable.

5. Discussion and Conclusions

This research presented the PMSE, a simple and efficient metric for evaluating ML surrogate modeling by focusing on each point of the numerical model separately. Our experimental results demonstrate their important role by indicating more complex points for training that, along with our understanding of numerical simulation, could be used to interpret the possible correlation between the numerical model and learning progress. From this, we conclude that this metric can provide useful and complementary information about the performance of the model.

To thoroughly examine the benefits of the proposed metric, ideally it should have been tried on a variety of different numerical problems. However, we only experimented it on a simplified numerical problem, which is not a major limitation for this study, as the

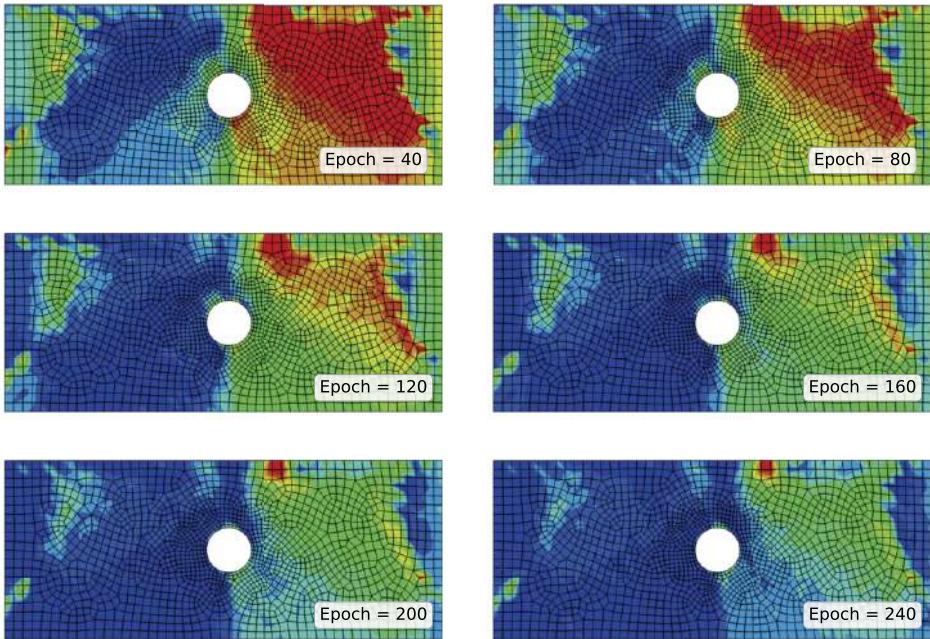


Figure 4. Contour plots of absolute errors corresponding to the sample with the highest squared error at the beginning, illustrated at different epochs. The contours are mapped onto the initial configuration of the model under contact before deformation. The color range from red to blue shows the values from above 4 to 0, respectively.

results obtained are sufficient to achieve our main objective, i.e., presenting preliminary evidence about the usefulness and feasibility of the pointwise metric.

Although the new metric is derived from the definition of MSE, similar pointwise metrics could be introduced using, e.g., the root mean squared error. We hope that such a pointwise metric will be applied to surrogate models in different (but applicable) domains in the future.

References

- [1] Belytschko T, Liu WK, Moran B. Nonlinear finite elements for continua and structures. Chichester: John Wiley & Sons; 2000.
- [2] Chen Q, Chen W, Wang G. Fully-coupled electro-magneto-elastic behavior of unidirectional multi-phased composites via finite-volume homogenization. Mechanics of Materials. 2021;154:103553.
- [3] Pang L, Liu W, Zheng Q, Du Y, Meng X, Li X. Evaluation and analysis of metal mine filling based on numerical simulation and actual measurement. Environmental Earth Sciences. 2021;80(16):505.
- [4] Torabi MR, Hosseini M, Akbari OA, Afrouzi HH, Toghraie D, Kashani A, et al. Investigation the performance of solar chimney power plant for improving the efficiency and increasing the outlet power of turbines using computational fluid dynamics. Energy Reports. 2021;7:4555-65.
- [5] Sajjadnia SS, Haghpanahi M, Razi M. Computational simulation of the multiphasic degeneration of the bone-cartilage unit during osteoarthritis via indentation and unconfined compression tests. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2019;233(9):871-82.

- [6] Sajjadinia SS, Carpentieri B, Holzapfel GA. A backward pre-stressing algorithm for efficient finite element implementation of in vivo material and geometrical parameters into fibril-reinforced mixture models of articular cartilage. *Journal of the Mechanical Behavior of Biomedical Materials*. 2021;114:104203.
- [7] Queipo NV, Haftka RT, Shyy W, Goel T, Vaidyanathan R, Kevin Tucker P. Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*. 2005 Jan;41(1):1–28.
- [8] Liang L, Liu M, Martin C, Sun W. A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. *Journal of The Royal Society Interface*. 2018;15(138):20170844.
- [9] Holzapfel GA, Linka K, Sherifova S, Cyron CJ. Predictive constitutive modelling of arteries by deep learning. *Journal of The Royal Society Interface*. 2021;18(182):20210411.
- [10] Shim VB, Holdsworth S, Champagne AA, Coverdale NS, Cook DJ, Lee TR, et al. Rapid Prediction of Brain Injury Pattern in mTBI by Combining FE Analysis With a Machine-Learning Based Approach. *IEEE Access*. 2020;8:179457–65.
- [11] Liang L, Liu M, Martin C, Sun W. A machine learning approach as a surrogate of finite element analysis-based inverse method to estimate the zero-pressure geometry of human thoracic aorta. *International Journal for Numerical Methods in Biomedical Engineering*. 2018 Aug;34(8):e3103.
- [12] Martin-Guerrero JD, Ruperez-Moreno MJ, Martinez-Martinez F, Lorente-Garrido D, Serrano-Lopez AJ, Monserrat C, et al. Machine learning for modeling the biomechanical behavior of human soft tissue. *IEEE International Conference on Data Mining Workshops, ICDMW*. 2016;0:247–253.
- [13] Martínez-Martínez F, Rupérez-Moreno MJ, Martínez-Sober M, Solves-Llorens JA, Lorente D, Serrano-López AJ, et al. A finite element-based machine learning approach for modeling the mechanical behavior of the breast tissues under compression in real-time. *Computers in Biology and Medicine*. 2017 Nov;90:116–124.
- [14] Gao W, Lu X, Peng Y, Wu L. A deep learning approach replacing the finite difference method for in situ stress prediction. *IEEE Access*. 2020;8:44063–44074.
- [15] Kalyuzhnyuk AV, Lapin RL, Murachev AS, Osokina AE, Sevostianov AI, Tsvetkov DV. Neural networks and data-driven surrogate models for simulation of steady-state fracture growth. *Materials Physics and Mechanics*. 2019;42(3):351–358.
- [16] Pfaff T, Fortunato M, Sanchez-Gonzalez A, Battaglia PW. Learning mesh-based simulation with graph networks. *International Conference on Learning Representations*. 2021:1–18.
- [17] Xiao S, Hu R, Li Z, Attarian S, Björk KM, Lendasse A. A machine-learning-enhanced hierarchical multiscale method for bridging from molecular dynamics to continua. *Neural Computing and Applications*. 2020;32(18):14359–73.
- [18] Ford E, Maneparambil K, Rajan S, Neithalath N. Machine learning-based accelerated property prediction of two-phase materials using microstructural descriptors and finite element analysis. *Computational Materials Science*. 2021;191:110328.
- [19] Bhosekar A, Ierapetritou M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers and Chemical Engineering*. 2018 Jan;108:250–267.
- [20] Phellan R, Hachem B, Clin J, Mac-Thiong JM, Duong L. Real-time biomechanics using the finite element method and machine learning: Review and perspective. *Medical Physics*. 2021;48(1):7–18.
- [21] Cilla M, Martinez J, Pena E, Martinez Mn. Machine learning techniques as a helpful tool toward determination of plaque vulnerability. *IEEE Transactions on Biomedical Engineering*. 2012;59(4):1155–61.
- [22] Molnar C. *Interpretable machine learning*; 2019.
- [23] Calka M, Perrier P, Ohayon J, Grivot-Boichon C, Rochette M, Payan Y. Machine-Learning based model order reduction of a biomechanical model of the human tongue.
- [24] Wang R, Kashinath K, Mustafa M, Albert A, Yu R. Towards physics-informed deep learning for turbulent flow prediction. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '20*. New York, NY, USA: Association for Computing Machinery; 2020. p. 1457–1466.
- [25] Sanchez-Gonzalez A, Godwin J, Pfaff T, Ying R, Leskovec J, Battaglia P. Learning to simulate complex physics with graph networks. In: III HD, Singh A, editors. *Proceedings of the 37th International Conference on Machine Learning*. vol. 119 of *Proceedings of Machine Learning Research*. PMLR; 2020. p. 8459–68.
- [26] Ye Y, Wang Z, Zhang X. An optimal pointwise weighted ensemble of surrogates based on minimization

- of local mean square error. *Structural and Multidisciplinary Optimization*. 2020 Aug;62(2):529-42.
- [27] Xu H, Zhang X, Li H, Xiang G. An ensemble of adaptive surrogate models based on local error expectations. *Mathematical Problems in Engineering*. 2021;2021:1-14.
 - [28] Acar E. Various approaches for constructing an ensemble of metamodels using local measures. *Structural and Multidisciplinary Optimization*. 2010 Dec;42(6):879-96.
 - [29] Goel T, Haftka RT, Shyy W. Comparing error estimation measures for polynomial and kriging approximation of noise-free functions. *Structural and Multidisciplinary Optimization*. 2008 Aug;38(5):429.
 - [30] Giselle Fernández-Godino M, Balachandar S, Haftka RT. On the use of symmetries in building surrogate models. *Journal of Mechanical Design*. 2019 01;141(6).
 - [31] Fuhg JN, Fau A. Surrogate model approach for investigating the stability of a friction-induced oscillator of Duffing's type. *Nonlinear Dynamics*. 2019;98(3):1709-29.
 - [32] Mo S, Lu D, Shi X, Zhang G, Ye M, Wu J, et al. A taylor expansion-based adaptive design strategy for global surrogate modeling with applications in groundwater modeling. *Water Resources Research*. 2017;53(12):10802-23.
 - [33] Freutel M, Schmidt H, Dürselen L, Ignatius A, Galbusera F. Finite element modeling of soft tissues: material models, tissue interaction and challenges. *Clinical Biomechanics*. 2014;29(4):363-72.
 - [34] Kingma DP, Ba JL. Adam: A method for stochastic optimization. In: Bengio Y, LeCun Y, editors. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings; 2015.
 - [35] Abaqus: Dassault Systèmes Simulia Corp., Providence, RI, USA; 2020. Available from: www.simulia.com.
 - [36] Chollet F, et al.. Keras. GitHub; 2015. Available from: <https://github.com/fchollet/keras>.
 - [37] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-scale machine learning on heterogeneous systems. 2015. Software available from tensorflow.org.
 - [38] Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Loizides F, Schmidt B, editors. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. IOS Press. IOS Press; 2016. p. 87-90.
 - [39] Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. *Nature*. 2020 Sep;585(7825):357–362.
 - [40] Hunter JD. Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*. 2007;9(3):90–95.

Multi-Stage Prediction of Feed System Time Series

Lihua SHEN^a, Biling WANG^{b1}, Hongjun LIU^a

^a*Shenyang Aerospace University, China*

^b*Avic Aerodynamics Research Institute of China, China*

Abstract. In order to reduce the tracking error of the computer numerical control (CNC) feed system and improve the CNC machining accuracy, a novel prediction model is proposed based on fuzzy C-means robust variational echo state network. Firstly, the feed speed time series is clustered, and then reconstructed for different categories. The multi-stage robust prediction models are established to realize the multi-state robust prediction of the CNC machining feed velocity to reduce the tracking error of the feed system. Finally, the reference and actual time series with different feed speed are used to verify the established models. The results show that the proposed method can reduce the tracking error and realize the effective prediction of the time series of the feed system.

Keywords. CNC, multi-stage, cluster, robust prediction, time series

1. Introduction

The modeling and prediction methods which are often used for error compensation include physical model driven prediction methods and data driven prediction methods. The physical model-driven forecasting methods need the specific structure of the physical model and corresponding parameters to be set in advance, while it is always difficult for complex control systems to establish accurate physical models. Consequently, physical model driven methods such as iterative control method [1] and its improvement [2] are subject to many restrictions in practical applications. The other type of methods to solve the above problems are the data-driven prediction methods. Two typical and commonly used error modeling prediction methods in this type of method are the neural network prediction method [3-5] and the support vector machine for regression (SVR) prediction method [6]. However, the error modeling based on wavelet neural network [3] and LSTM [4] is time-consuming and the BP [5] method is not robust. The SVR in [6] just predict time series without establishing specific models for different feed states.

As a data-driven model, neural network has strong data modeling capabilities and nonlinear approximation capabilities. Especially the robust neural network methods have found an increasingly wide utilization in time series prediction [7,8]. Compared with the traditional neural network, the echo state network (ESN) [9] has the advantages of fast training speed and considering the temporal correlation of data. It

¹ Corresponding author: Avic Aerodynamics Research Institute of China, China; E-mail: sacslh@126.com.

has been more and more used in complex system modeling and time series prediction, and has achieved good prediction results. And its improvement robust variational echo state network (RVESN) [10] is proved to be robust for data analysis with noise and outliers. In view of the characteristics of the time series of the feed system and the modeling ability of the RVESN in the time series prediction, a multi-stage feed system time series prediction method based on the fuzzy C-means (FCM) [11] and the RVESN named FCM-RVESN is proposed. Firstly, the first-order difference of the reference speed is extracted, and then FCM is used to cluster the data into stable processing state and acceleration-deceleration state. Subsequently two clusters are implemented phase space reconstruction respectively, and prediction models for the reconstructed time series are established based on RVESN. The rest of this paper is structured as follows. Section 2 gives a brief review of the preliminary works. Section 3 presents the proposed FCM-RVESN model. Section 4 gives the simulation results of time series from circle and linear motions of practical CNC feed system. In Section 5, the conclusions are given.

2. Robust Echo State Network

The ESN [9] always has a large reservoir to map data to higher-dimensional space. It often discards the initial transient samples in the prediction process to improve the accuracy of the network prediction. When the length of the transient state is $l-1$, suppose the state matrix of the reservoir and the target output matrix are \mathbf{A} and \mathbf{y} respectively. For the convenience of subsequent calculations, each row of \mathbf{A} is set as \mathbf{a}_k , and assume that N samples are retained, $\mathbf{A} \in R^{N \times r}$ then:

$$\mathbf{A} = [x(l), x(l+1), x(l+2), \dots, x(l+N-1)]^T \quad (1)$$

$$\mathbf{y} = [y(l), y(l+1), y(l+2), \dots, y(l+N-1)]^T \quad (2)$$

According to the echo state network structure, Eq. (3) is obtained:

$$\mathbf{Aw} = \mathbf{y} \quad (3)$$

The RVESN uses the robust Gaussian mixture distribution as the model output likelihood function. For any training sample, the formula is as follows:

$$p(y(k)) = \eta p(y(k)) + (1-\eta)p_0(y(k)) \quad (4)$$

Among them, η is an adaptive parameter that is automatically tuned as the Outliers and noise number and proportion of data set. $p(\mathbf{y})$ and $p_0(\mathbf{y})$ are as follows:

$$p(\mathbf{y} | \mathbf{w}, \beta) = \left(\frac{\beta}{2\pi} \right)^{N/2} \exp \left\{ -\frac{\beta}{2} \|\mathbf{y} - \mathbf{Aw}\|^2 \right\} \quad (5)$$

$$p_0(\mathbf{y}) = \left(\frac{\beta_0}{2\pi} \right)^{N/2} \exp \left\{ -\frac{\beta_0}{2} \|\mathbf{y} - \mathbf{Aw}\|^2 \right\} \quad (6)$$

The approximate posterior probability distribution of \mathbf{w} is Gaussian, and the covariance matrix and mean value are Σ and \mathbf{u} respectively:

$$\Sigma = \left(\sum_{k=1}^N [\beta_0(1 - E_z(z_k)) + \beta E_z(z_k)] \mathbf{a}_k^T \mathbf{a}_k + \text{diag}(\alpha_h) \right)^{-1} \quad (7)$$

$$\mathbf{p} = \sum \left(\sum_{k=1}^N [\beta_0(1 - E_z(z_k)) + \beta E_z(z_k)] y(k) \mathbf{a}_k^T \right) \quad (8)$$

3. FCM-RVESN

3.1. Feature Extraction and State Clustering

Through the analysis of the time series for linear and circle motions of the feed system, it is found that the first-order difference of the data is significantly different in the stable state and the acceleration-deceleration state. Therefore, features of the data at different feed speed are extracted by the first-order difference, and then FCM is used to cluster the time series based on the extraction results.

Suppose the time series generated at a certain feed rate has N samples with d dimensions, which is expressed as $\mathbf{x}(i) = [x_1(i), x_2(i), \dots, x_d(i)]^T \in R^d$, $N-1$ samples with d dimensions are obtained by implementing first difference.

After the above-mentioned first-order difference, FCM is utilized to cluster absolute value of the above data into two states. The time series to be clustered can be expressed as follows:

$$\begin{aligned} x_{fcm}(i) &= \text{abs}[x_i(i+1) - x_i(i)] \\ x_{fcm}(i+1) &= \text{abs}[x_i(i+2) - x_i(i+1)] \\ &\dots \\ x_{fcm}(N-1) &= \text{abs}[x_i(N) - x_i(N-1)] \end{aligned} \quad (9)$$

The FCM [11] defines the objective function to be optimized :

$$Jm = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m (d_{ij})^2 \quad (10)$$

In the Eq. (10), m is Weighted index, and $d_{ij} = \|x_i - c_j\|^2$ is the distance between the i th time series and the centre of the j th cluster, which is stable state or acceleration-deceleration state. In the objective function, C is number of the clusters, which is equal to two. N is the number of samples, u_{ij} is Membership degree that i belongs to the j th cluster, x_i is the i th sample, and c_j is the centre of the j th cluster. Finally, the speed time series are divided into two clusters: stable state and acceleration-deceleration state.

3.2. FCM-RVESN Prediction Model

Due to the characteristics of the time series, the actual speed is not only related to the reference speed at the current moment, but also related to the reference speed and actual speed at the previous moments. Therefore, it needs to be reconstructed in phase space. According to Takens theorem, as long as a suitable embedding dimension is found, the original time series can be reconstructed to obtain a dynamic system with a higher dimension than the input variable, and its prediction accuracy can be improved by reconstructing the time series.

After first-order difference and phase space reconstruction of the time series, the FCM-RVESN feed system time series prediction models of two states are established at different speed and the model structure is as figure 1:

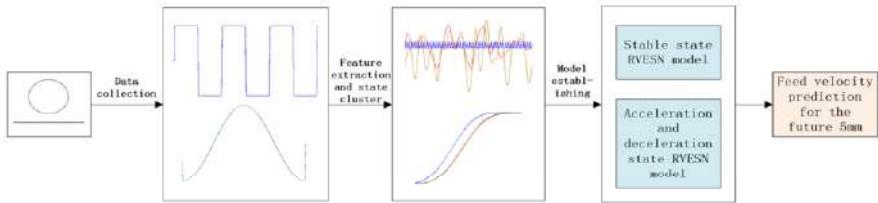


Figure 1. Structure of FCM-RVESN.

First, the reference speed and the actual speed of the semi-closed loop on each coordinate axis are collected at different feed speed in the process of linear and circular interpolation. The first-order difference is performed on the collected data to extract features, and then FCM is used to cluster the collected data. The time series in the stable state and the acceleration-deceleration state are obtained. The time series is reconstructed in different time processes and different states respectively. Subsequently, RVESN models for the reconstructed time series are established in different states.

4. Experimental Results

In the feed servo drive system, the excitation signal is set as $x(t)$, which is also the input reference speed, and $y(t)$ is the semi-closed loop actual response signal, which is the actual speed of the servo motor encoder. In the semi-closed loop speed prediction of the feed system, the previous $x(t)$ and $y(t)$ are used to predict $y(t+\delta)$. The value of the prediction horizon δ is set according to actual CNC system. For single-step prediction, δ is equal to 1. In each simulation experiment of this paper, δ is set as 5, which means that the actual speed of the fifth millisecond in the future is predicted. The reason why it is set to 5 is because the delay time of the servo is about 3 to 4 milliseconds between received the reference data and outputting response signal. The data used in this chapter is generated by linear and circular motions. Parameters used in the experiments are as follows. Data generated by the two motions will be analyzed respectively.

Table 1. parameters of the experiments

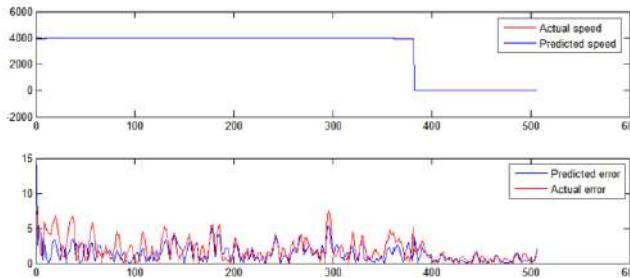
Size	of	Sparseness	Spectral	Regularization	Cluster	Prediction
reservoir			Radius	Coefficient of SVESN	Number	Horizon
100	0.05		0.95	2	2	5

4.1. Circle Time Series

The two-axis arc motion path is a circle with diameter of 100mm. By setting different feed speed as 4000mm/min and 5000mm/min, the time series at different preset feed speed can be obtained. In the case of circular motion. After decomposing the composite speed to the X-axis and Y-axis, the speed on the X-axis is generally a sine curve, and the speed on the Y-axis is generally a cosine curve. Parameters are set as table 1. The speed prediction results on X-axis are shown in the table 2 and figure 2, which demonstrate that the SVR has higher precision while predicting feed speed at 5000mm/min, but FCM-RVESN has higher precision compared with SVR, SVESN and actual system accuracy while predicting feed speed at 4000mm/min.

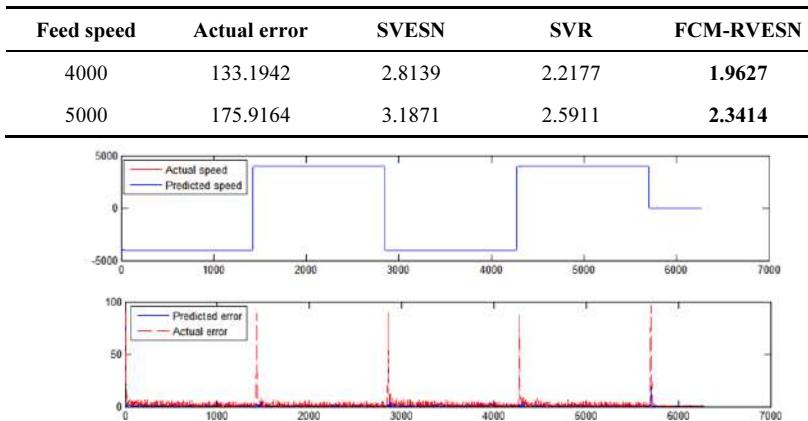
Table 2. comparison of prediction results (circle time series).

Feed	Actual	SVESN	SVR	FCM-RV
speed	error			ESN
4000	16.3927	14.2434	3.2185	2.4709
5000	23.8857	21.2286	1.5549	1.9804

**Figure 2.** Prediction results of circle (stable state at 4000mm/min)

4.2. Linear Time Series

The single-axis linear motion path length is 100mm. By setting feed speed as 4000mm/min and 5000mm/min, the time series at different preset feed speed can be obtained. The speed prediction results are shown in the table 3 and figure 3, which demonstrate that the FCM-RVESN has higher precision compared with SVR, SVESN and actual system accuracy in two state and it is a feasible model for the speed prediction of CNC feed system.

Table 3. comparison of prediction results (linear time series).**Figure 3.** Prediction results of linear time series (acceleration-deceleration state at 4000mm/min)

5. Conclusion

In this paper, the time series from the feed system of CNC machine tools is analyzed, clustered and a multi-stage FCM-RVESN prediction method is proposed at different feed speed. The actual speed of practical CNC feed system is forecasted, and the prediction results are compared with SVR, SVESN and actual measured error. It is found that the proposed method has higher precision in three datasets for all the datasets used in the experiments.

Acknowledgements

The authors would like to thank the financial support from the PhD Start-up Fund of Shenyang Aerospace University (120419026/502) and the fund of “Research on load prediction method of wind Tunnel balance calibration based on intelligent algorithm”.

References

- [1] Ahn HS, Chen YQ, Moore KL. Iterative learning control: brief survey and categorization. *IEEE Transactions on Systems Man & Cybernetics Part C*, 2007, 37(6): 1099-1121
- [2] Tsai MS, Yen CL, Yau HT. Integration of an empirical mode decomposition algorithm with iterative learning control for high-precision machining. *IEEE/ASME Transactions on Mechatronics*, 2013, 18(3): 878-886
- [3] Ustundag BB, Kulaglic A. High-performance time series prediction with predictive error compensated wavelet neural networks. *IEEE Access*, 2020, 8: 210532-210541.
- [4] Jiang Y, Chen J, Zhou H et al. Nonlinear time-series modeling of feed drive system based on motion states classification. *Journal of Intelligent Manufacturing*, 2020, 31(8).
- [5] Yin Q, Tan F, Chen H et al. Spindle thermal error modeling based on selective ensemble BP neural networks. *The International Journal of Advanced Manufacturing Technology*, 2019, 101(5): 1699-1713.

- [6] Lei M, Jiang G, Yang J et al. Thermal error modeling with dirty and small training sample for the motorized spindle of a precision boring machine. *The International Journal of Advanced Manufacturing Technology*, 2017, 93(1): 571-586.
- [7] Li D, Han M, Wang J. Chaotic time series prediction based on a novel robust echo state network. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, 23(5): 787-799
- [8] Altan A, Karasu S. The effect of kernel values in support vector machine to forecasting performance of financial time series. *The Journal of Cognitive Systems*, 2019, 4(1): 17-21.
- [9] Jaeger H, Haas H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 2004, 304(5667): 78-80
- [10] Shen L, Chen J, Zeng Z et al. A novel echo state network for multivariate and nonlinear time series prediction. *Applied Soft Computing*, 2018, 62: 524-535.
- [11] Ghosh S, Dubey SK. Comparative analysis of k-means and fuzzy c-means algorithms. *International Journal of Advanced Computer Science and Applications*, 2013, 4(4).

Multi-Class Imbalanced Corporate Bond Default Risk Prediction Based on the OVO-SMOTE-Adaboost Ensemble Model

Jie SUN¹ and Jingmei ZHU

School of Accounting, Tianjin University of Finance and Economics, Tianjin, China

Abstract. Corporate bond default risk prediction is important for regulators, issuers and investors in the bond market. We propose a new approach for multi-class imbalanced corporate bond risk prediction based on the OVO-SMOTE-Adaboost ensemble model, which integrates the one-versus one (OVO) decomposition method, the synthetic minority over-sampling technique (SMOTE) and the Adaboost ensemble method. We categorize corporate bond default risk into three classes: very low default risk, relatively low default risk and high default risk, which is more scientific than the traditional two-class bond default risk, and carry out empirical experiments by respectively using DT, SVM, Logit and MDA as basic classifiers. Empirical results show that the prediction performance of the OVO-SMOTE-Adaboost (DT) model is overall better than the other three ensemble models such as OVO-SMOTE-Adaboost (SVM), OVO-SMOTE-Adaboost (Logit) and OVO-SMOTE-Adaboost (MDA). In addition, the OVO-SMOTE-Adaboost (DT) model greatly outperforms the OVO-SMOTE (DT) model, which is a single classifier model based on OVO and SMOTE without Adaboost. Therefore, the OVO-SMOTE-Adaboost (DT) model has satisfying performance of multi-class imbalanced corporate bond default risk prediction and is of great practical significance.

Keywords. Corporate bond default risk prediction, multi-class imbalanced classification, OVO-SMOTE-Adaboost ensemble model

1. Introduction

The bond market is an indispensable part of the financial market system, which not only provides a platform for fund-raisers and investors to carry out financing and investment activities, but also plays an important role in deepening the reform of financial structure, promoting the marketization of interest rates and improving the government's function of macroeconomic regulation. In recent years, China's bond market has developed rapidly, with the total volume of bond issuance and the number of bond issuers both increasing, and the market size is growing, becoming the second largest market in the world. In 2014, China's bond market saw its first material default, then rigid payment was gradually broken in China's bond market. From 2014 to the present, the number of bond default cases has been increasing and the nature of bond default becomes more and

¹ Corresponding Author, School of Accounting, Tianjin University of Finance and Economics, Tianjin, China; E-mail: sunjiehit@gmail.com.

more complex and diverse, which not only harms the interests of investors, but also impedes the development of the bond market. Therefore, how to timely and effectively predict the default risk of corporate bonds is an important problem in the bond market. The Global Artificial Intelligence Technology Conference 2021 (GAITC2021), held in Hangzhou on June 5, aims to build a comprehensive platform that seamlessly links the global strength of artificial intelligence and presents intelligent technology in a three-dimensional manner, further promoting technological innovation, application innovation, achievement transformation and industrial chain collaboration in the field of artificial intelligence. In view of corporate bond default risk prediction, artificial intelligence technologies can extract more potentially valuable information from financial data than traditional statistical methods, and can predict the default risk of corporate bonds more effectively, which is of great significance to the healthy development of bond market.

2. Literature review

Prior literature has proposed many statistical or artificial intelligence models for predicting corporate financial distress or credit risk. Altman [1] used multiple discriminant analysis (MDA) to predict enterprise bankruptcy by introducing multiple financial ratios as variables, and constructed the Z-Score model, which was widely used for bond credit scoring. In recent years, many scholars have used artificial intelligence methods to achieve better prediction results. Li and Sun [2] found that the regression tree classification model shows better performance in financial distress prediction than MDA. Lee [3] introduced support vector machine (SVM) into enterprise credit rating prediction and used five-fold cross-validation method to determine the SVM model's best parameter values. They compared the performance of the SVM model with other models such as neural networks (NNs) and MDA, and the empirical results show that the SVM model has higher classification accuracy.

The application of single classifiers is easily affected by data structure. In order to solve the limitation of single classifier, some scholars have proposed multi-classifier ensemble methods, which refer to combining multiple single classifiers and finally integrating the output of multiple classifiers for prediction. Common ensemble methods include Bagging (Bootstrap aggregating), Boosting, Adaboost (adaptive Boosting), and so on. Sun et al. [4] proposed a financial distress prediction model based on weighted majority voting combination of multiple classifiers, which can greatly improve the accuracy and stability of financial distress prediction and outperform single classifiers. Tsai et al. [5] used the Bagging and Boosting algorithms to build ensemble models based on basic classifiers such as logistic regression, SVM, decision tree (DT), NNs, and so on. Li et al. [6] used NNs, SVM, random forest, Chi-squared automatic interaction detector and classification and regression tree (CART) to predict credit ratings, and the results showed that random forest has the highest accuracy. Sun et al. [7] proposed a time-weighted Adaboost-SVM ensemble model to predict dynamic financial distress, and empirical results show that its predictive performance is better than single SVM, batch-based ensemble with local weighted scheme, Adaboost-SVM and Timeboost-SVM. However, when the training data is highly class-imbalanced, both ensemble models and single classifier models usually have poor performance for predicting the minority class.

In order to solve the problem caused by imbalanced data, scholars have proposed different solutions. For example, Sun et al. [8] used SVM as the basic classifier, and

proposed a multi-classifier ensemble model. They divided the majority class into several subsets and combined each of them with the minority class to construct several class-balanced training datasets, and then used them to train multiple SVM classifiers for combination. The results show that the model can significantly improve the predictive ability for minority samples. Liu et al. [9] used a class-oriented feature selection method to solve the problem of multi-class imbalance. Sun et al. [10] proposed a DT ensemble model based on synthetic minority over-sampling technique (SMOTE) and Bagging with differentiated sampling rates, and their empirical results demonstrate the effectiveness of this method in credit assessment of enterprises in the circumstance of class imbalance.

In addition, most of the current literature simply divides corporate status into default and non-default and builds binary classification models to predict bond default risk, ignoring the diversity of corporate bond status. Hence, it is necessary to define the multi-class risk of corporate bond default and build multi-class bond default risk prediction models. However, multi-class prediction is more complex than the traditional two-class prediction, and we need to transform a multi-class classification problem into multiple two-class classification problems with a certain decomposition method. Common decomposition methods have one-versus-one (OVO), one-versus-all (OVA) and binary tree method. Li et al. [11] built a multi-class classifier through the OVO decomposition method. Zhou et al. [12] used the OVO and OVA decomposition methods for multi-class classification problems, and built different multi-class classification models based on three basic classifiers. Sun et al. [13] combined SVM with OVO, OVA, error correction output coding (ECOC) to build a multi-class financial distress prediction model, and empirical results show that OVO-SVM outperforms OVA-SVM and ECOC-SVM overall. However, the above studies mainly integrate decomposition methods with single classifiers to solve the multi-class classification problem, neglecting the advantage of ensemble classifiers. How to construct more effective ensemble models for multi-class corporate bond default risk prediction needs to be further investigated. This paper integrates the SMOTE oversampling method and the Adaboost ensemble method with the OVO decomposition method to build ensemble models for multi-class imbalanced corporate bond default risk prediction.

3. OVO-SMOTE-Adaboost Ensemble Model for Multi-Class Imbalanced Corporate Bond Default Risk Prediction

The framework of the OVO-SMOTE-Adaboost ensemble model for multi-class imbalanced corporate bond default risk prediction is shown in Figure 1, and the algorithm is described as follows.

(1) Decompose a multi-class sample into several two-class subsamples by OVO.

The initial training set $Train$ contains three types of samples, the first with a very low risk of bond default, the second with a relatively low risk of bond default, and the third with a high risk of bond default. Based on the OVO decomposition method, the initial training set $Train$ is decomposed into three two-class training subsets, $Train_1$ for the first class and the second class, $Train_2$ for the first class and the third class, $Train_3$ for the second class and the third class.

(2) Train the SMOTE-Adaboost ensemble model for two-class bond default risk prediction.

Use training sets $Train_j$ ($j=1, 2, 3$) to train three two-class bond default risk prediction models based on the SMOTE oversampling method and the Adaboost ensemble method, respectively, as follows:

- a) Initialize the sample weights, each with a weight of $w_j^{1i}=1/n$, and the initial weight distribution of the training set is represented as $D_j^1=[w_j^{11}, w_j^{12}, \dots, w_j^{1n}]=[\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}]$, where n is the sample number of the training set.
- b) Iterate $t=1 \dots T$
 - i. Construct a training dataset $Train_j^t$ by resampling from $Train_j$ according to the sample weight distribution D_j^t ;
 - ii. Use the SMOTE oversampling method to balance the training dataset $Train_j^t$ to get a balanced training dataset $Train\text{-balance}_j^t$;
 - iii. Use the balanced training dataset $Train\text{-balance}_j^t$ to train a weak classifier based on a certain classification algorithm (DT, SVM, Logit, and MDA will be used in our empirical study, respectively);
 - iv. Test the weak classifier with $Train_j$, and calculate the error rate $e_j^t=P(h_j^t(x_i)) \neq y_i)=\sum_{i=1}^n w_j^{ti} I(h_j^t(x_i) \neq y_i)$, where y_i is the class label for the training sample and $h_j^t(x)$ is the basic classifier, and the following is the calculation rule for I .

$$I = \begin{cases} 1, & h_j^t(x_i) = y_i \\ 0, & h_j^t(x_i) \neq y_i \end{cases}$$

vi. Calculate the voting weight $\alpha_j^t=0.5 \ln \frac{1-e_j^t}{e_j^t}$;

vii. Update the sample weight $w_j^{(t+1)i} = w_j^{ti} \exp(-\alpha_j^t \mu_j^{ti}) / Z_j^t$, where Z_j^t is the

normalization constant, $Z_j^t = 2 \sqrt{e_j^t(1 - e_j^t)}$.

$$\mu_j^{ti} = \begin{cases} 1, & h_j^t(x_i) = y_i \\ -1, & h_j^t(x_i) \neq y_i \end{cases}$$

- c) Constitute a strong ensemble classifier H_j through weighted voting combination of T weak classifiers obtained by iterative training.

(3) Combine the results of two-class bond default risk prediction to obtain the results of multi-class bond default risk prediction.

Input the feature values of a testing sample into the three strong ensemble classifiers for binary classification, $H_1(x)$, $H_2(x)$ and $H_3(x)$, respectively, and combine their results by the simple majority voting principle to finally output the multi-class prediction result $H(x)$.

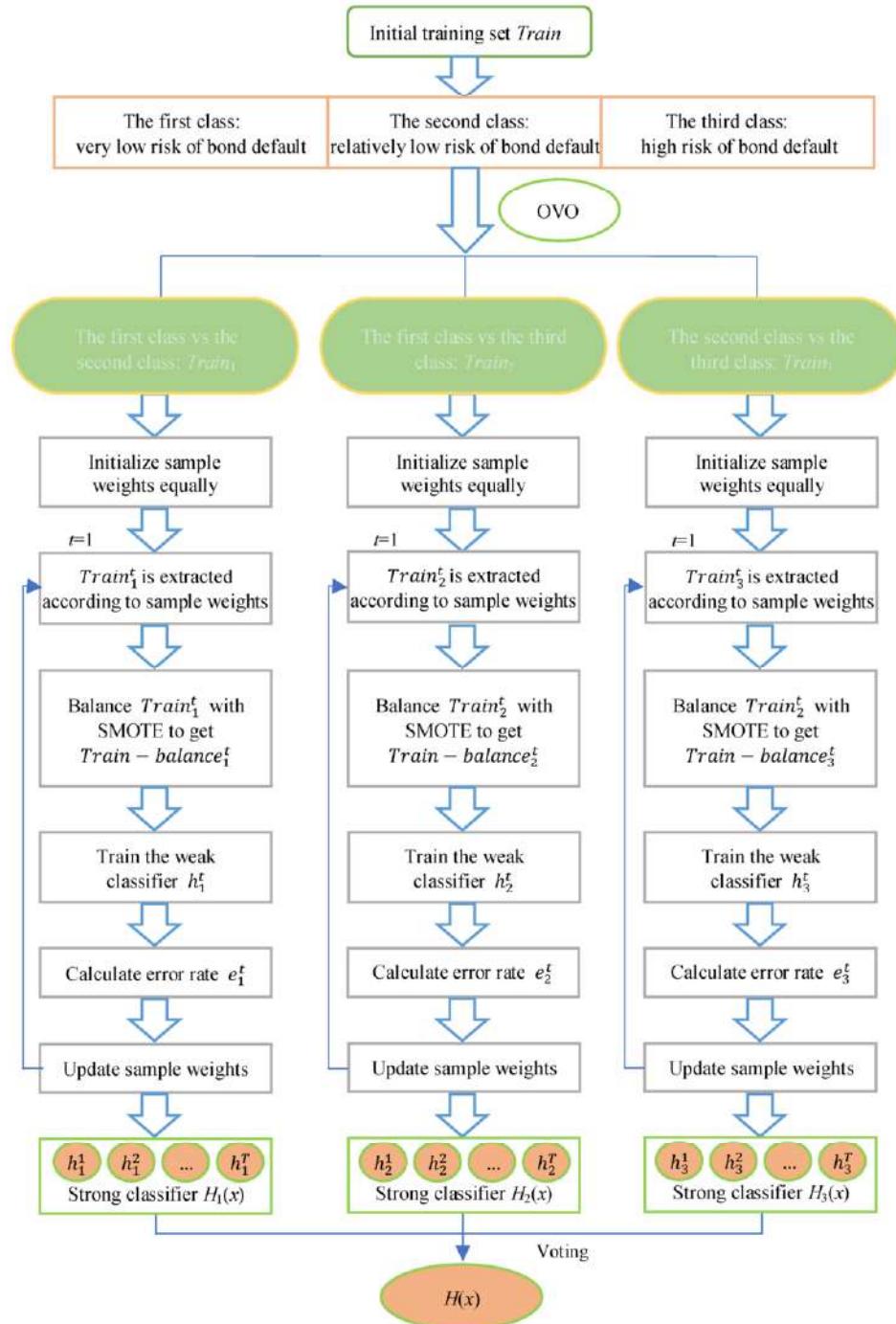


Figure 1. The framework of the OVO-SMOTE-Adaboost ensemble model.

4. Design of Empirical Research

4.1. Data Collection and Preprocessing

4.1.1. Data Collection

In 2014, China had its first default bond, breaking the rigid payment, so we collect data of enterprises that publicly issue bonds during 2014 to 2020 from the Wind database. Excluding the financial industry, we collect the credit ratings of bond issuers and their financial indicator data corresponding to the fiscal year that is two years earlier than the credit rating year. Then we delete the duplicate data of the companies issuing multiple bonds, the outliers and the data with missing values. We categorize the risk of bond default into three classes according to the credit ratings of the bond issuers: the first class is very low default risk, and the corresponding bond credit ratings are AAA and AA⁺; the second class is relatively low default risk, and the corresponding bond credit ratings are AA, AA⁻, A⁺, A and A⁻; the third class is high default risk, and the corresponding bond credit ratings are B, C and substantive default. The final sample consists of 3927 observations, among which 1534 belong to the first class, 2185 belong to the second class and 208 belong to the third class.

4.1.2. Data Preprocessing

In order to ensure the reliability of the results and improve the convergence speed and accuracy of the model, the original indicator data need to be normalized. The Min-Max normalization method is used to normalize the original data into a specific interval according to Eq. (1):

$$X' = \frac{X - \text{min}}{\text{max} - \text{min}} \quad (1)$$

where “max” is the maximum value of the sample data and “min” is the minimum value of the sample data. The normalization processing maps the data uniformly into the interval of [0,1].

4.2. Feature Selection

Initial features are composed of 26 financial indicators covering the five aspects of the operation capability, solvency, profitability, cash flow and development capability. We use the SPSS software to conduct multi-independent-sample Kruskal-Wallis test for the 26 financial indicators to verify mean differences among the three classes, and delete indicators with no significant mean differences among different classes. The correlation between each pair of indicators is analyzed, and the strong correlation indicators with a correlation coefficient greater than 0.5 is eliminated, resulting in 16 financial indicators as the model input variables, which were working capital turnover, accounts receivable turnover, fixed asset turnover, asset-liability ratio, current ratio, cash ratio, interest coverage, return on total assets, net profit margin on sales, ratio of operating profit to total operating revenue, ratio of net profit from operating activities to pre-tax profit, ratio of net cash flow from operating activities to operating revenue, proportion of net cash flows from operating activities, cash operating index, growth rate of operating profit and monetary capital growth rate.

4.3. Design of Empirical Comparison

For comparative analysis, we integrate the OVO decomposition method with the SMOTE oversampling method (OVO-SMOTE) to build single classifier models for multi-class imbalanced corporate bond default risk prediction, which are compared with the OVO-SMOTE-Adaboost ensemble models. We divide the preprocessed data set into the training set and the testing set in terms of the proportions of 80% and 20%. Then we use the training set to train the OVO-SMOTE single classifier models and the OVO-SMOTE-Adaboost ensemble models based on DT, SVM, MDA and Logit, respectively, and use the testing set to test their performance of multi-class bond default risk prediction. The training and testing process repeats 30 times, and we record 30 results and calculate the average for each model. The empirical research is implemented in the MATLAB 2016a software. We use the CART algorithm to build DT classifiers, and the minimum leaf tree is 50 after many experiments. The kernel function of SVM is set as the radial basis function (RBF), and the parameters c and g in the RBF SVM model are determined to be 32 and 2 respectively by grid optimization.

4.4. Model Evaluation Indicators

We use *Accuracy*, *Recall* (R), and *Macro-averaging Recall* (*Macro- R*) as indicators to evaluate model performance. For binary classification, TP refers to the number of positive samples that are correctly predicted as the positive; FP refers to the number of positive samples that are incorrectly predicted as the negative; TN refers to the number of negative samples that are correctly predicted as the negative; and FN refers to the number of negative samples that are incorrectly predicted as the positive. The indicators of *Accuracy* and *Recall* are calculated according to Eq. (2) and (3):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

$$R = \frac{TP}{TP+FN} \quad (3)$$

This study focuses on three-class corporate bond default risk prediction. Suppose N denotes the total number of testing samples, and $N1$, $N2$ and $N3$ are the testing sample numbers of the first, second and third class, respectively. $TC1$, $TC2$ and $TC3$ respectively represent the sample numbers of the first, second and third class that are correctly predicted as their true class. We should calculate the indicator of *Accuracy* as equation (4), the indicator of *Recall* for each class as Eq. (5)-(7), and the indicator of *Macro- R* as Eq. (8).

$$Accuracy = \frac{TC1+TC2+TC3}{N} \quad (4)$$

$$R1 = \frac{TC1}{N1} \quad (5)$$

$$R2 = \frac{TC2}{N2} \quad (6)$$

$$R3 = \frac{TC3}{N3} \quad (7)$$

$$Macro_R = \frac{R1+R2+R3}{3} \quad (8)$$

where $R1$ is the *Recall* for the first class, $R2$ is the *Recall* for the second class, $R3$ is the *Recall* for the third class.

5. Empirical Results and Analysis

5.1. Performance of Corporate Bond Default Risk Prediction based on OVO-SMOTE-Adaboost Ensemble Model

We repeat the whole experiment for 30 times to obtain 30 empirical results. In each experiment, we divide the initial data set into the training set and the testing set and use the training set to build the OVO-SMOTE-Adaboost ensemble models for corporate bond default risk prediction based on DT, SVM, Logit and MDA, respectively. We denote them as OVO-SMOTE-Adaboost (DT), OVO-SMOTE-Adaboost (SVM), OVO-SMOTE-Adaboost (Logit) and OVO-SMOTE-Adaboost (MDA), respectively, and use the testing set to test each model's performance. The average values of the model evaluation indicators obtained in 30 experiments by the OVO-SMOTE-Adaboost ensemble models are shown in Table 1.

As can be seen from Table 1, the OVO-SMOTE-Adaboost (DT) model has the highest *Accuracy* at 70.99%, showing overall better prediction performance than the other three models. According to the indicator of *Recall*, none of the above four models can achieve the highest value of *Recall* for all the three classes. For $R1$, the OVO-SMOTE-Adaboost (SVM) model has the highest value of 72.72%, followed by the OVO-SMOTE-Adaboost (DT) model with a value of 71.32%. For $R2$, the OVO-SMOTE-Adaboost (DT) model has the highest value of 70.75%, which is much higher than those of the other three models with a difference of 8.42%, 11.55% and 17.54%, respectively. For $R3$, the OVO-SMOTE-Adaboost (DT) model has the highest value of 71.14%, which is higher than those of the other three models with a difference of 5.21%, 1.95% and 0.57%, respectively. In terms of *Macro-R*, the OVO-SMOTE-Adaboost (DT) model is the highest at 71.07%. By comparing the results, the OVO-SMOTE-Adaboost (DT) model has optimal and stable performance for multi-class imbalanced corporate bond default risk prediction, and the OVO-SMOTE-Adaboost (SVM) model is overall better than the other two.

Table 1. The performance of the OVO-SMOTE-Adaboost ensemble models.

Ensemble model	Accuracy	R1	R2	R3	Macro-R
OVO-SMOTE-Adaboost (DT)	70.99%	71.32%	70.75%	71.14%	71.07%
OVO-SMOTE-Adaboost (SVM)	66.57%	72.72%	62.33%	65.93%	66.99%
OVO-SMOTE-Adaboost (Logit)	60.82%	62.03%	59.20%	69.19%	64.47%
OVO-SMOTE-Adaboost (MDA)	55.79%	57.48%	53.21%	70.57%	60.42%

5.2. Comparative Analysis between the Ensemble Models and the Single Classifier Models

For comparative analysis, we also train and test the single classifier models, i.e., OVO-SMOTE (DT), OVO-SMOTE (SVM), OVO-SMOTE (Logit) and OVO-SMOTE (MDA). The evaluation indicators of each model obtained in the 30 experiments are also recorded and averaged, and the results are compared with the empirical results of each ensemble model.

The results of OVO-SMOTE-Adaboost (DT) vs. OVO-SMOTE (DT) are shown in Figure 2. The OVO-SMOTE-Adaboost (DT) model outperforms the OVO-SMOTE (DT) model for each evaluation indicator, and the differences of *Accuracy*, *R1*, *R2*, *R3* and *Macro-R* between the two models are 10.33%, 13.65%, 8.76%, 2.36% and 8.26%, respectively. The empirical results of the two models show that the integration of the Adaboost ensemble method with the OVO decomposition method and the SMOTE oversampling method can greatly improve the performance of multi-class imbalanced corporate bond default risk prediction based on the DT classifier.

The results of OVO-SMOTE-Adaboost (SVM) vs. OVO-SMOTE (SVM) are shown in Figure 3. The OVO-SMOTE-Adaboost (SVM) model is 0.62% more accurate than the OVO-SMOTE (SVM) model, with 1.41% lower for *R1*, 2.02% higher for *R2*, 0.97% higher for *R3*, and 0.52% higher for *Macro-R*. The empirical results of the two models show that the integration of the Adaboost ensemble method with the OVO decomposition method and the SMOTE oversampling method slightly reduces the prediction ability for the first class, and slightly improves the prediction ability for the second and third class when SVM is used as the basic classifier. The two models have overall similar performance of multi-class imbalanced corporate bond default risk prediction.

The results of OVO-SMOTE-Adaboost (Logit) vs. OVO-SMOTE (Logit) are shown in Figure 4. The two models show little difference of multi-class imbalanced corporate bond default risk prediction, with -0.02% for *Accuracy*, with 0.1% for *R1*, -0.08% for *R2*, -0.16% for *R3* and -0.05% for *Macro-R*. The empirical results of the two models show that the integration of the Adaboost ensemble method with the OVO decomposition method and the SMOTE oversampling method cannot enhance the performance of multi-class imbalanced corporate bond default risk prediction when Logit is used as the basic classifier.

The results of OVO-SMOTE-Adaboost (MDA) vs. OVO-SMOTE (MDA) are shown in Figure 5. The OVO-SMOTE-Adaboost (MDA) model was 1.72% higher than the OVO-SMOTE (MDA) model for *Accuracy*, 4.67% higher for *R1*, 0.26% lower for *R2*, 0.65% higher for *R3*, and 1.69% higher for *Macro-R*. The empirical results of the two models show that the integration of the Adaboost ensemble method with the OVO decomposition method and the SMOTE oversampling method can improve the overall

performance of multi-class imbalanced corporate bond default risk prediction when MDA is used as the basic classifier.

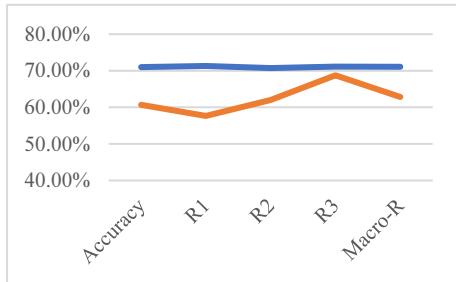


Figure 2. OVO-SMOTE-Adaboost (DT) vs. OVO-SMOTE (DT).

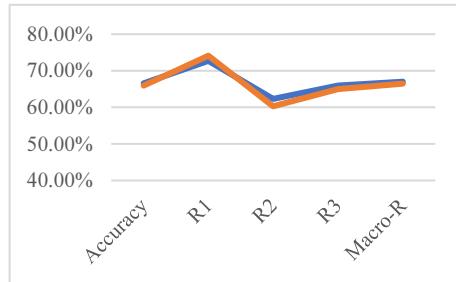


Figure 3. OVO-SMOTE-Adaboost (SVM) vs. OVO-SMOTE (SVM).

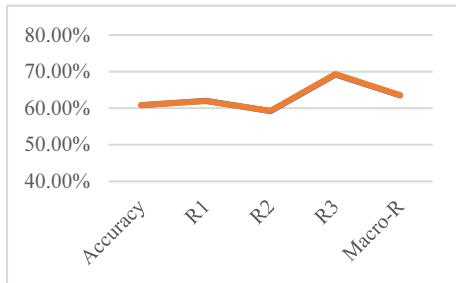


Figure 4. OVO-SMOTE-Adaboost (Logit) vs. OVO-SMOTE (Logit).

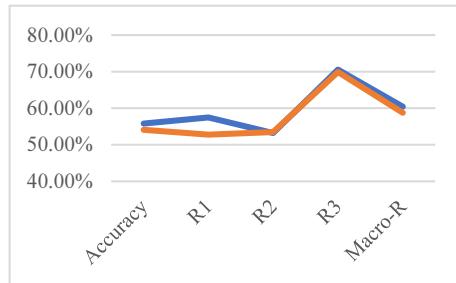


Figure 5. OVO-SMOTE-Adaboost (MDA) vs. OVO-SMOTE (MDA).

5.3. Discussion on Model Application

Corporate bond default has a negative impact on the development of bond market. Therefore, the corporate bond default risk prediction model based on integration of the OVO decomposition method, the SMOTE oversampling method and the Adaboost ensemble method is of great practical significance. Empirical results have shown that the OVO-SMOTE-Adaboost (DT) model has satisfying performance for multi-class imbalanced corporate bond default risk prediction. In practice, it can support the policy making of regulators and the decision makings of bond issuers and investors. First, for regulators, the OVO-SMOTE-Adaboost (DT) model for multi-class corporate bond default risk prediction is helpful for regulators to detect the risks in the bond market in advance, and to issue and implement regulatory policies in time to promote the healthy development of the bond market. Second, for bond issuers, the occurrence of bond default makes them harder to raise capital in the future, thus increasing their financing difficulty. The OVO-SMOTE-Adaboost (DT) model for multi-class corporate bond default risk prediction can help bond issuers to identify their own default risks in a timely manner and take management measures to prevent bond default. Finally, for bond investors, the OVO-SMOTE-Adaboost (DT) model for multi-class corporate bond default risk prediction can reduce the degree of information asymmetry between bond issuers and investors, and enable investors to identify and avoid investment risks in a timely manner.

6. Conclusion

In this paper, the OVO decomposition method, the SMOTE oversampling method and the Adaboost ensemble method are combined to establish the OVO-SMOTE-Adaboost ensemble models for multi-class imbalanced corporate bond default risk prediction. DT, SVM, Logit and MDA are used as the basic classifiers, respectively. The empirical results of the proposed ensemble models are compared with those of corresponding single classifier models, which only combine the OVO decomposition method and the SMOTE oversampling method without the Adaboost ensemble method. The conclusions are as follows:

(1) The OVO-SMOTE-Adaboost ensemble models achieve satisfying performance for multi-class imbalanced corporate bond default risk prediction. The OVO-SMOTE-Adaboost (SVM) model has better recognition ability for the bonds with very low default risk, the OVO-SMOTE-Adaboost (DT) model has better recognition ability for the bonds with relatively low default risk and the bonds with high default risk. The OVO-SMOTE-Adaboost (DT) model is overall superior to the other three ensemble models and is more suitable for multi-class imbalanced corporate bond default risk prediction.

(2) The OVO-SMOTE-Adaboost ensemble models show overall better performance for multi-class imbalanced corporate bond default risk prediction than the OVO-SMOTE single classifier models. In detail, the OVO-SMOTE-Adaboost (DT) model greatly outperforms the OVO-SMOTE (DT) model, and the performance of the OVO-SMOTE-Adaboost (MDA) model is slightly better than that of the OVO-SMOTE (MDA) model. There is little difference in performance of the other two groups of models. Therefore, it is of great significance to integrate the Adaboost ensemble method with the OVO-SMOTE model for multi-class imbalanced corporate bond default risk prediction, particularly when DT is used as the basic classifier.

This study still has the following limitations: (1) We only use financial features for multi-class imbalanced corporate bond default risk prediction, neglecting the impact of non-financial features on model performance. (2) We use cross-industry data to train and test the models and do not consider the different characteristics of corporate bond default risk prediction across different industries.

References

- [1] Altman EI. Financial ratios discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*. 1968 Sep;23(4):589–609.
- [2] Li H, Sun J, Wu J. Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods. *Expert Systems with Applications*. 2010 Feb;37(8):5895–5904.
- [3] Lee YC. Application of support vector machines to corporate credit rating prediction, *Expert Systems with Applications*. 2007 Apr; 33(1):67–74.
- [4] Sun J, Li H. Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications*. 2008 Oct; 35: 818–827.
- [5] Tsai CF, Sue KL, Hu YH, Chiu A. Combining feature selection, instance selection, and ensemble classification techniques for improved financial distress prediction. *Journal of Business Research*. 2021 Mar; 130: 200–209.
- [6] Li JP, Mirza N, Rahat B, Xiong D. Machine learning and credit ratings prediction in the age of fourth industrial revolution. *Technological Forecasting and Social Change*. 2020 Dec; 161: 120309.
- [7] Sun J, Fujita H, Chen P, Li H. Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble. *Knowledge-Based Systems*. 2016 Mar; 120: 4–14.

- [8] Sun J, Lee YC, Li H, Huang Q. Combining B&B-based hybrid feature selection and the imbalance-oriented multiple-classifier ensemble for imbalanced credit risk assessment. *Technological and Economic Development of Economy*. 2015 May; 21(3):351-378.
- [9] Liu Z, Wang R, Tao W, Cai X. A class-oriented feature selection approach for multi-class imbalanced network traffic datasets based on local and global metrics fusion. *Neurocomputing*. 2015 Nov; 168: 365-381.
- [10] Sun J, Lang J, Fujita H, Li H. Imbalanced enterprise credit evaluation with DTE-SBD: Decision tree ensemble based on SMOTE and bagging with differentiated sampling rates. *Information Sciences*. 2018 Jan; 425: 76-91.
- [11] Li L, Yang C, Xie Q. 1D embedding multi-class classification methods. *International Journal of Wavelets, Multiresolution and Information Processing*. 2016 Apr; 14(2).
- [12] Zhou L, Tam KP, Fujita H. Predicting the listing status of Chinese listed companies with multi-class classification models. *Information Sciences*. 2016 Jan; 328: 222-236.
- [13] Sun J, Fujita H, Zhenbbg Y, Ai W. Multi-class financial distress prediction based on support vector machines integrated with the decomposition and fusion methods. *Information Sciences*. 2021 Jul; 559: 153-170.

An Effective Deep Neural Network Architecture for Cross-Subject Epileptic Seizure Detection in EEG Data

Imene JEMAL ^{a,c,1}, Amar MITICHE ^a, Lina ABOU-ABBAS ^{b,c}, Khadidja HENNI ^{b,c}
and Neila MEZGHANI ^{b,c}

^aINRS-EMT, Université du Québec, Montréal, Canada

^bCentre de Recherche du CHUM, Montréal, Canada

^cCentre de Recherche LICEF, Université TÉLUQ, Montréal, Canada

Abstract. For several decades, the detection of epileptic seizures has been an active research topic. The performance of current patient-specific algorithms is satisfactory. However, due to significant variability of EEG data between patients, cross-subject seizure characterization and detection remains a challenging task. The purpose of this study is to propose and investigate a modified convolutional neural network (CNN) architecture based on separable depth-wise convolution for effective automatic cross-subject seizure detection. The architecture is conceived with a reduced number of trainable parameters to reduce the model complexity and storage requirements to easily deploy it in connected devices for real-time seizure detection. The performance of the proposed method is evaluated on two public datasets collected in the Children's Hospital Boston and the University of Bonn respectively. The method achieves the highest sensitivity–false positive rate/h of 91.93%–0.005, 100%–0.057 for the CHB-MIT and Ubonn datasets respectively.

Keywords. Epilepsy, Seizure detection, Deep learning, Convolutional Neural Network, EEG

1. Introduction

Epilepsy is a neuro-degenerative disorder manifested by recurrent unprovoked seizures. It is the second most frequent neurological disease [1]. Most often, EEG records are the basis for a diagnosis. The visual inspection of hours of EEG data is impractical because it is time consuming and requires interpretations by experts. Therefore, several studies have been conducted to develop computer-aided diagnostic systems that can automatically detect epileptic seizures [2, 3]. Several EEG-based epilepsy detection models have been proposed. However, epileptic patterns are highly variable across seizures and across patients, which makes real-time application of these models in clinical settings quite a challenging task. Models go along two veins: General cross-subject modeling which apply to patients at large, and patient-specific modeling which applies to patients individ-

¹Corresponding Author: Imene Jemal, INRS-EMT, Université du Québec, Montréal, Canada; E-mail: imene.jemal@inrs.ca

ually. Patient-specific modeling is generally impracticable because it requires recording sufficient seizure onsets for each patient, separately from others. Cross-subject modeling does not have to treat each patient separately, but it faces the major problem of adapting the detection algorithm to unseen data of new patients, mainly due to the significant cross-subject EEG pattern variability [4].

Recent studies of deep learning (DL) [5] for epilepsy detection, which automatically encodes EEG features characteristic of epilepsy, have been much more potent than traditional feature selection and classification methods [6]. One of the first deep learning studies on epilepsy detection [7] used a convolutional neural network (CNN) for feature extraction in an image-based representation of EEG signals, followed by Long Short Memory units (LSTM) for classification. The method was evaluated on the CHB-MIT dataset for subject-specific (sensitivity = 95-100%) and cross-subject (sensitivity = 85%) models. Along this vein, [8] used CNNs to distinguish interictal epileptiform discharges from normal activity. The method achieved the higher accuracy of 87.51% on the CHB-MIT dataset. The study in [9] compared time domain and frequency domain EEG representations in CNN feature coding for subject-specific seizure detection. On the CHB-MIT and the Freiburg datasets, frequency domain modeling performed significantly better (97.5% vs 95.4% accuracy). The investigation of [10] proposed a pyramidal one-dimensional convolution neural network architecture, achieving higher detection sensitivity, specificity, and accuracy of 89%, 99% and 98.2% respectively. The experiments, however, were carried out on a relatively small amount of data from 5 patients in the Ubonn database. Although it did not address inter-ictal and ictal period classification, the study in [11] used a relatively large EEG records dataset of 300 patients from the Temple University Hospital EEG database to compare CNN with conventional classification (support vector machine (SVM) and random forest (RF)) for distinguishing healthy from epileptic patients in cross-subject EEG data. Performance was better in terms of area under the precision-recall curve (AUPR) with the tiny visual geometry group CNN architecture (AUPR = 0.9242), than SVM (AUPR = 0.8651) and RF (AUPR = 0.8578).

Convolutional neural networks have recently received a lot of attention in the field of seizure detection. CNNs [5] can learn effective nonlinear local features of increasing complexity as processing progresses from the input layer to the output layer. CNNs were first described by [12] as neural networks composed of a sequence of convolution and pooling layers. The original CNN was subsequently upgraded to have a larger architecture, called AlexNet, which was followed by even more complex structures keeping the original basic ideas. The investigation of [13] introduced the Inception-V1 architecture (GoogLeNet) with processing steps that express correlation between channels followed by spatial pattern learning. The architecture allowed richer pattern feature learning using less network parameters. Similar in concept, the Xception architecture [14] starts with depth-wise convolution applied on channels to be followed by a point-wise convolution to combine the coded features. It has the particularity of not using non-linearity between layers. The architecture showed better performances than Inception-V3 in classification tasks on the ImageNet dataset and a larger image classification dataset comprising 350 million images and 17,000 classes.

The purpose of this study is to investigate epilepsy detection in EEG data by a novel CNN architecture based on separable depth-wise layers. Unlike others, the initial layer of this CNN performs a convolution to learn a representation of the raw signal in terms of frequency components. This is in agreement with feature extraction by filter bank sig-

nal decomposition [15]. The architecture also includes separable depth-wise layers: this necessitates significantly fewer network parameters than the standard 2D convolution layer, which has the effect of: 1) lowering model complexity and subsequent execution, 2) reducing storage requirements to allow execution on connected devices and, 3) allowing model training on either small or large datasets. This architecture, pertinent to cross-subject modeling, is investigated here to distinguish between ictal and inter-ictal periods in EEG data. The cross-subject modeling can significantly increase the applicability of the algorithm because it allows processing of data from unseen subjects, unlike models that are patient-specific.

This CNN architecture was evaluated using the publicly available CHB-MIT and Ubonn databases. As described in detail subsequently, it reached high performance, with 91.82% (5 patients in CHB-MIT) and 99.60% (23 patients in Ubonn).

The remainder of this paper is organized as follows. Section 2 presents the method in detail. The experimental setup is presented in Section 3. Finally, Section 4 contains the results and a discussion.

2. Methods

In this section, we first explain the difference between standard convolution and depth-wise separable convolution operations. Afterward, a detailed description of the deployed architecture is presented.

2.1. Standard vs Separable depth-wise convolution

A standard convolution layer simply applies a convolution operation between the input and learnable weighted filters to obtain a new representation of the data called a feature-map (See figure 1a). The purpose of the learned filters is to simultaneously capture spatial correlation and cross-channel correlation.

A separable convolution layer splits the convolution kernel into two smaller kernels, which has the effect of reducing the number of parameters. A classic example is the decomposition of the Sobel edge detector kernel into two smaller kernels as shown in the following Equation 1:

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} * (-1 \ 0 \ 1) \quad (1)$$

The depth-wise separable convolution [16] is similar to the separable convolution. However, the convolution operation is decomposed differently into two steps : (1) a depth-wise convolution to learn local patterns for each channel and (2) a point-wise convolution to find linear combination of the extracted feature capturing the correlations between the channels (See figure 1b).

- Depth-wise convolution : This type of layer is so named because it takes into consideration the depth dimension (the number of channels) where the convolution of each channel with a different kernel is performed separately as shown in Figure 1b. This step allows learning filters for each channel.

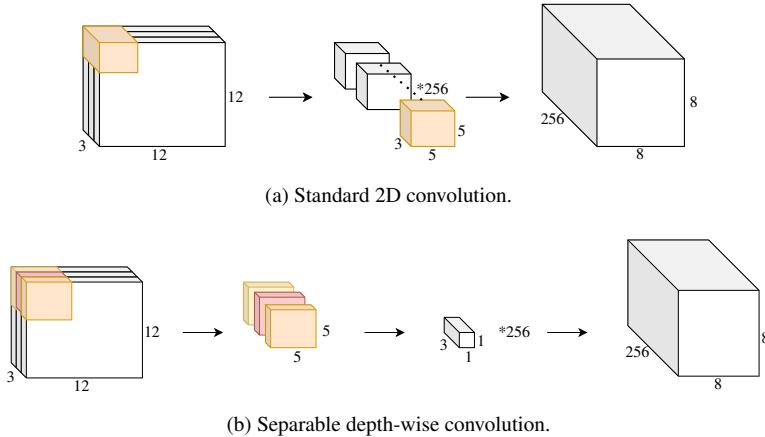


Figure 1. Illustration of different convolution layer types.

- **Point-wise convolution** : This convolution uses a 1×1 kernel with depth equal to the number of channels to iterate through each point to learn a linear combination of the feature-maps (output from the depth convolution). This step allows capturing the correlation across the channels.

Figure 1 shows an example of standard and separable depth-wise convolution. Normal convolution transforms the input 256 times using Kernels of size $5 \times 3 \times 3$ which lead to a total number of parameters of $5 \times 3 \times 3 \times 256$. On the contrary, separable depth-wise convolution applies a single transformation (kernel of size $5 \times 3 \times 3$) and simply elongate it to 256 channels using 256 kernels of size $3 \times 1 \times 1$. The number of parameters is reduced to $5 \times 3 \times 3 + 3 \times 1 \times 1 \times 256$.

2.2. Architecture design

In this section, we introduce a convolution neural network architecture inspired by [15, 17]. The architecture uses a reduced number of parameters, allowing it to be trained with very limited data as well as with larger data sets. Full details of the network architecture are summarized in the table 1.

- The network starts with a 2D convolution to learn F1 frequency filters. Indeed, this block is inspired by the concept of the filter bank, which is a set of band-pass filters that separate the input signal into several components, each corresponding to a single frequency sub-band of the original signal. This technique is usually performed before the feature extraction step [15].
- Subsequently, we alternate between separable depth-wise convolution layers and pooling layers. As mentioned above, the separable depth-wise convolution begin with the depth-wise convolution to learn specific filters for each channel (signal component in a specific frequency sub-band). This is followed by a point-wise convolution combining the learned features across channels. Pooling layers are used to reduce the dimensions. We applied batch normalization before the non-linear activation to stabilize the training. In order to regularize the model, we added a dropout layer.

- Finally, the features are flattened by the fully connected layer and fitted to a softmax classification.

Table 1. The detailed network architecture, where C = number of channels, T, T', T'' = signal duration, F1, F2, F3 = number of convolution kernels to learn, N = number of classes, respectively.

Layer	#Filters	Size	#Output	Activation
Input			(1,C,T)	
Conv2D	F1	(1,128)	(F1,C,T)	Linear
BatchNorm2d			(F1,C,T)	
Reshape			(F1*C,1,T)	
DepthwiseConv	F2	(1,32)	(32*F2*F1*C,1,T')	Linear
PointwiseConv	2	(1,1)	(2,1,T')	Linear
BatchNorm2D			(2,1,T')	
Activation			(2,1,T')	Relu
AveragePool2D		(1,8)	(2,1,T'//8)	
Dropout			(2,1,T'//8)	
DepthwiseConv	F3	(1,16)	(16*F3*2,1,T'')	Linear
PointwiseConv	2	(1,1)	(2,1,T'')	Linear
BatchNorm2D			(2,1,T'')	
Activation			(2,1,T'')	Relu
AveragePool2D		(1,4)	(2,1,T'//4)	
Dropout			(2,1,T'//4)	
Linear			(2*T'//4)	
Dense			N=2	Softmax

3. Experimental setup

3.1. EEG data

We evaluated the proposed architecture on two publicly available datasets of EEG data (see Table 2).

3.1.1. CHB-MIT database

The database collected at the Boston Children's Hospital [18] contains scalp EEG data of 23 patients. The EEG data were recorded through 19 electrodes distributed over the scalp according to the international standard 10/20. The signals were amplified and sampled with a frequency of 256 Hz. During 940 hours of EEG recording, 198 epileptic seizures were recorded. The time of onset of a seizure and its duration has been identified by experts. Signals were filtered to remove artifacts and noise. We divided the records into 5-seconds-long non-overlapping windows. The selected window length yielded 601,689 and 2,157 inter-ictal and pre-ictal samples respectively.

3.1.2. *UBonn university database*

The data were collected from 5 monitored patients at Bonn University [19]. It consists of five sets (denoted A-E) each containing 100 single-channel EEG segments of 23.6 seconds. Sets A and B contain surface EEG recordings from healthy people. Sets C and D were recorded from epileptic patients in seizure-free intervals. Set E is the only set that contains activity recorded during seizures. The data was sampled at a rate of 173.61 Hz. The segments were selected after a visual inspection for artifacts like muscle activity or eye blinking. For classifying seizure-free and seizure EEG segments, set A-D are labeled as normal EEG records and set E is reserved for seizure events.

Table 2. Public databases for seizure detection

Dataset	CHB-MIT	UBonn
Number of subjects	23	5
Number of seizure	198	100*
Total duration(hour)	940	3.24
Recording type	Scalp	Scalp and Intracranial
Number of channels	17	1**
Sampling frequency(Hz)	256	173.73

*100 seizures file each of 23.36 s duration.

**Multi-channel data was converted to a single channel.

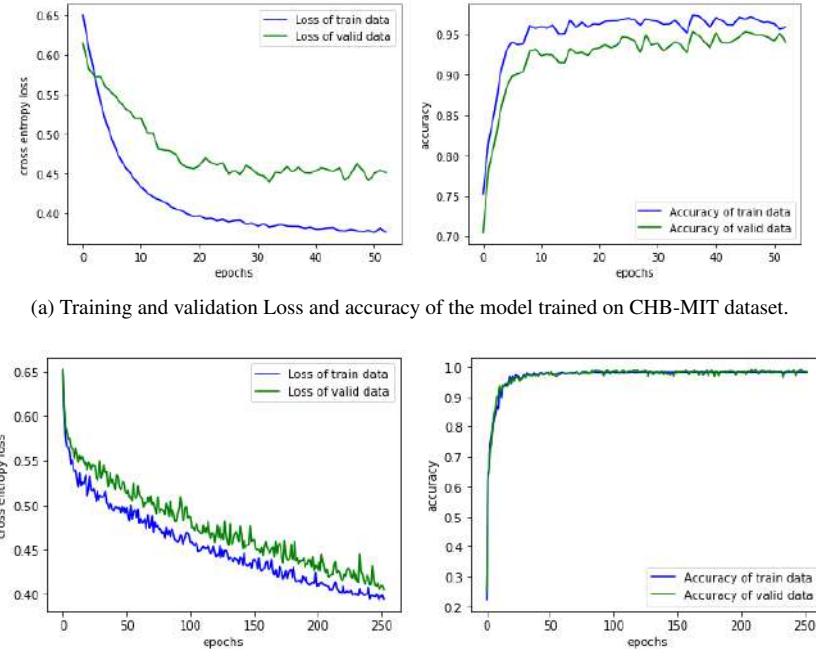
3.2. *Model training*

The CNN (Table 1) was trained using each dataset (Table 2) separately. Classifiers were implemented using Pytorch [20], while data pre-processing was performed using MNE-Python package [21]. The three-way holdout method was employed to optimise the performances in the hyper-parameter tuning step. In fact, we divided the data into three stratified (having same classes proportions) sets: the training set, the validation set, and the test set, each containing data from different patients. To tackle the problem of imbalanced data set, we used a cost-sensitive cross-entropy loss. The Adam optimizer was proposed as a gradient-based method with $\beta_1 = 0.9$, $\beta_2 = 0.999$ and a learning rate of 0.002. We used the early stopping criteria to prevent over-fitting where training runs up to 500 epochs, or until the validation loss does not decrease anymore for at least 20 epochs.

4. Results and Discussion

We evaluated the performance of the proposed cross-subject method on the CHB-MIT and Ubonn university public databases. Figure 2 shows the learning curves: the cross-entropy loss and the model accuracy curves, for training and validation data, as a function of epochs. The training and validation loss decreases to a point of stability, with a small gap between the loss values indicating that the models are well fitted.

Table 3 shows the accuracy, sensitivity, F1-score, and the false positive rate per hour obtained by evaluating the neural network with unseen data from new patients. The results reveal that the proposed architecture performs very well. Indeed, for the CHB-



(b) Training and validation Loss and accuracy of the model trained on Ubonn.

Figure 2. Learning curves for all three datasets.

MIT data set, our method achieved a maximum accuracy rate of 91.82% with a sensitivity of 91.93% and a false alarm rate of 0.005/hour. With the Ubonn database, results were boosted to an accuracy of 99.60%, a sensitivity of 100%, and a false alarm rate of 0.057/hour.

Table 3. Model's performances for different datasets.

Database	Accuracy	Sensitivity	F-measure	FRP
CHB-MIT	91.82	91.93	95.73	0.005
UBonn	99.60	100	99.75	0.057

The results clearly show that CNNs are capable to extracting discriminative features in EEG data to allow cross-subject classification of inter-ictal and ictal data intervals. The hyper-parameters related to the network structure (number of layers, size of convolution filters), the parameters of the activation and regularization functions, as well as the training parameters (learning rate and batch size, and optimization algorithm parameters), have been carefully chosen with the three-fold out method by observing the train and validation learning curves. Although not done in this study, performance can potentially benefit from further extensive fine-tuning of the hyperparameters.

A comparison of our method with other CNN-based solutions is given in Table 4. For a fair comparison, we focused on networks evaluated on the same data sets. However, unlike our cross-subject solution, all models are patient-specific except for [10]. Regarding the CHB-MIT dataset, our method performs significantly better than the patient-specific

CNN-based model in [8]. Although the method has slightly lower accuracy than [9], it has the significant advantage of the ability to generalize across patients. In addition, the method had a better performance, although slightly, on the Ubonn university data than [10].

Table 4. Comparison of Benchmarking of recent seizure detection CNN-based studies and our work.

study	Dataset	Model	Method	Acc
[8]	CHM-MIT	patient-specific	CNN on raw EEG signals	87.51
[9]	CHM-MIT	patient-specific	CNN on frequency domain signals	95.4
This work	CHM-MIT	cross-subject	CNN with separable depth-wise convolutions	91.82
[10]	Ubonn	cross-subject	pyramidal one-dimensional CNN	98.2
This work	Ubonn	cross-subject	CNN with separable depth-wise convolutions	99.60

Thanks to the specific choice of the type of layers in our study, the proposed architecture has no more than 2,700 parameters in total. the minimal number of used parameters promotes the deployment of this model in connected devices for real-time seizure detection. In comparison, other neural networks for epileptic seizure detection [10] [11], used 8,326 and 16,401 parameters.

Overall, the classification results show that this study CNN architecture, which uses depth-wise convolution layers, performs well for epileptic seizure detection, using both a small and a fairly large database. Future work will focus on evaluating the architecture on even larger datasets, as well as applying it to other EEG-based classification tasks, such as epilepsy prediction, rather than detection, and seizure types categorization.

References

- [1] World Health Organization. Epilepsy; World Health Organization: Geneva, Switzerland, 2019. [Online]. Available : <https://www.who.int/news-room/fact-sheets/detail/epilepsy>.
- [2] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 975–982.
- [3] L. Abou-Abbas, I. Jemal, K. Henni, A. Mitiche, and N. Mezghani, "Focal and generalized seizures distinction by rebalancing class dataand random forest classification," in International Conference on Bioengineering and Biomedical Signal and Image Processing. Springer, 2021, pp. 63–70.
- [4] I. Jemal, A. Mitiche, and N. Mezghani, "A study of eeg feature complexity in epileptic seizure prediction," Applied Sciences, vol. 11, no. 4, p. 1579, 2021.
- [5] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, Deep learning. MIT press Cambridge, 2016, vol. 1, no. 2.
- [6] Y. LeCun, Y. Bengio et al., "Convolutional networks for images, speech, and time series," The handbook of brain theory and neural networks, vol. 3361, no. 10, p. 1995, 1995.
- [7] P. Thodoroff, J. Pineau, and A. Lim, "Learning robust features using deep learning for automatic seizure detection," in Machine learning for healthcare conference. PMLR, 2016, pp. 178–190.
- [8] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial eeg data," in 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2016, pp. 1–6.

- [9] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on eeg signals and cnn,"Frontiers in neuroinformatics, vol. 12, p. 95, 2018.
- [10] I. Ullah, M. Hussain, H. Aboalsamhet al., "An automated system for epilepsy detection using eeg brain signals based on deep learning approach,"Expert Systems with Applications, vol. 107, pp. 61–71, 2018.
- [11] T. Uyttenhove, A. Maes, T. Van Steenkiste, D. Deschrijver, andT. Dhaene, "Interpretable epilepsy detection in routine, interictal eeg data using deep learning," in Machine Learning for Health.PMLR,2020, pp. 355–366.
- [12] Y. LeCun, L. D. Jackel, L. Bottou, C. Cortes, J. S. Denker, H. Drucker,I. Guyon, U. A. Muller, E. Sackinger, P. Simardet al., "Learning algorithms for classification: A comparison on handwritten digit recognition,"Neural networks: the statistical mechanics perspective, vol. 261,no. 276, p. 2, 1995.
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan,V. Vanhoucke, A. Rabinovichet al., "Going deeper with convolutions.arxiv 2014,"arXiv preprint arXiv:1409.4842, vol. 1409, 2014.
- [14] F. Chollet, "Xception: deep learning with depthwise separable convolutions (2016),"arXiv preprint arXiv:1610.02357, 2016.
- [15] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter,K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball,"Deep learning with convolutional neural networks for eeg decoding and visualization,"Human brain mapping, vol. 38, no. 11, pp. 5391–5420,2017.
- [16] L. Sifre and S. Mallat, "Rigid-motion scattering for texture classification,"arXiv preprint arXiv:1403.1687, 2014.
- [17] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," inProceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1251–1258.
- [18] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, 2009.
- [19] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E.Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,"Physical Review E, vol. 64, no. 6,p. 061907, 2001.
- [20] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin,A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.
- [21] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier,C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonenet al., "Meg and eeg data analysis with mne-python,"Frontiers in neuroscience, vol. 7,p. 267, 2013.

Prediction of Alzheimer's Disease Based on Coordinate-Dense Attention Network

Yongmei TANG^{a,b}, Xiangyun LIAO^{b,1}, Weixin SI^b and Zhigang NING^a

^a*University of South China, Hengyang, China*

^b*Shenzhen Institute of Advanced Technology Chinese Academy of Sciences, Shenzhen, China*

Abstract. Alzheimer's disease (AD) is a degenerative disease of the nervous system. Mild cognitive impairment (MCI) is a condition between brain aging and dementia. The prediction will be divided into stable sMCI and progressive pMCI as a binary task. Structural magnetic resonance imaging (sMRI) can describe structural changes in the brain and provide a diagnostic method for the detection and early prevention of Alzheimer's disease. In this paper, an automatic disease prediction scheme based on MRI was designed. A dense convolutional network was used as the basic model. By adding a channel attention mechanism to the model, significant feature information in MRI images was extracted, and the unimportant features were ignored or suppressed. The proposed framework is compared with the most advanced methods, and better results are obtained.

Keywords. Alzheimer's disease, deep learning, convolutional neural network

1. Introduction

Alzheimer's disease is a common brain disorder characterized by the loss of connections between neurons in the brain and the shrinkage of the rest of the brain. Studies show that more than six million Americans may have Alzheimer's disease, second only to heart disease and cancer as the leading cause of death among the elderly [1]. Patients with mild cognitive impairment have memory disorder and are more likely to develop Alzheimer's disease. Therefore, MCI plays an important role in the early diagnosis and intervention of AD [2]. Accurate and efficient prediction of sMCI and pMCI can timely intervene and prevent the occurrence of Alzheimer's disease and reduce the potential patient population.

Due to the success of untreatable Alzheimer's disease, intervention and treatment in the early stages of Alzheimer's disease have become critical. Li *et al.* [3] use principal component analysis (PCA) to obtain features, and then use unsupervised deep learning training and finally use SVM classification. In addition to SVM, others also commonly use logistic regression (LR) [4] and random forest (RF) [5]. Many studies using MRI area (ROI) are interested in classification. Calvini *et al.* [6] around the medial temporal lobe and hippocampus were extracted to classify, most based on ROI measures are interested in artificial selection area, this inevitably can add human factors in the process of feature extraction. Some potential details may be ignored to affect the results.

¹ Corresponding Author, Xiangyun LIAO, Shenzhen Institute of Advanced Technology Chinese Academy of Sciences, Shenzhen, China; E-mail: xiangyun-l@163.com.

With the continuous development of deep learning, excellent results have been achieved in various fields, and it has also been applied to the diagnosis of Alzheimer's disease. Qiu *et al.* [7] connected Fully Convolutional Networks (FCN) and Multi-Layer Perception (MLP) to predict AD state by combining MRI data and non-imaging data, to achieve visualization of disease risk. Alejandro *et al.* [8] using transfer learning (TL) technology, ResNet feature extractor and SVM classifier were used to identify sagittal MRI images. In this paper, DenseNet was used as the basic model to predict AD disease by adding channel attention mechanism and extracting significant channel differences according to the subtle differences between sMCI and pMCI lesions, and found that adding attention mechanism has a good effect on the early prediction of AD.

2. Related Work

General Classification of Alzheimer's disease includes methods based on voxel morphological measurement and region of Interest measurement (ROI). Klöppel *et al.* [9] segmentation of gray matter images from MRI images, extraction of voxels, and use support vector machine to assign weight to voxels. Vemuri *et al.* [10] added demographic information, such as age and gender, to sMR scan information from the structural abnormality index (STAND) of tissues in sMR scan, and added genetic information, APOE genotype information. Finally, support vector machine SVM was used to optimize classification. However, the 3D image data processing method based on voxel is difficult to calculate due to its large amount of computation.

Regions of Interest (ROI) based classification methods generally rely on prior knowledge to divide regions, and the hippocampus, entorhinal cortex and medial temporal lobe are often used to classify Alzheimer's disease. Zhang *et al.* [11] combined three biomarkers, MRI, PET, and CSF, using an atlas based approach and automatic labeling by atlas distortion algorithm. Lauge *et al.* [12] using hippocampal structure as a classification feature, demonstrated that texture is more meaningful for early diagnosis. Feature extraction also focuses on the organizations and regions with high correlation with classification, but at the same time, some global information will be lost.

In recent years, deep learning technology has achieved good results in the classification of Alzheimer's disease. Kanghan *et al.* [13] based on the unsupervised learning of convolutional autoencoder (CAE) for the classification of AD and NC, proposed the end-to-end concept for classification. Lian *et al.* [14] proposed hierarchical full convolutional networks (H-FCN) to identify multi-scale discriminant locations, such as differentiated local plaques and regions in the brain. Compared with fixed interest area is extracted (ROI). Zhao *et al.* [15] used 3D multi-information generative adversarial network (mi-GAN) to generate predicted MRI images, and then used DenseNet to classify pMCI and sMCI.

3. Materials and Methods

3.1. Datasets and Preprocessing

This paper used data from a public data set of Alzheimer's Disease Neuroimaging (ADNI) (<http://adni.loni.usc.edu/>) download. In this study, follow-up data from ADNI were used to study the progression of patients with mild cognitive impairment (MCI)

over 6 months to 2 year period. Patients who developed and developed Alzheimer's disease during this period were considered progressive mild cognitive impairment (pMCI), and those who did not deteriorate were considered stable mild cognitive impairment (sMCI). T1-weighted MRI rapid gradient echo (MPRAGE) images of 969 subjects were downloaded from ADNI, including 210 AD, 274 NC, 183 pMCI and 302 sMCI. The demographic information of all subjects is shown in Table 1.

Table 1. Demographic Information of the Subjects Included in the Studied Datasets

Group	Gender(male/ Female)	Age
AD	210(110/100)	75.68+7.67
pMCI	183(105/78)	75.07+7.67
sMCI	302(168/134)	72.90+7.12
NC	274(129/145)	74.98+6.31

Downloaded from ADNI open source data set of the original MRI image preprocessing, use spm8 (<https://www.fil.ion.ucl.ac.uk/spm/>) software kit pictures into the DICOM format NIFTI format, then the skull stripping, cerebellum resection, and the MRI images after skull removal were registered into the MNI152 standard space by affine matching to eliminate the global linear differences. Finally, all images were resampled and cropped to $121 \times 145 \times 121$.

3.2. Methods

3.2.1. Dense Attention Network

This section will introduce a 3D dense channel attention network. The structure of the network is shown in Figure 1, which is mainly composed of a 3D dense Convolutional Network (DenseNet) [16], Coordinate Attention (CA) [17] and Squeeze-and-Excitation Attention (SE) [18] network. The problem of gradient explosion occurs when the network is too deep. The dense Connection solves this problem well. A densely connected convolution network has strong regularization effect and can reduce the overfitting on small training sets. From the perspective of feature channels, SE automatically obtains the importance of each channel through learning, which enhances the model's ability to extract important information. CA aggregates features from the two spatial directions and embed location information into channel attention, which can obtain not only channel information, but also direction and position information.

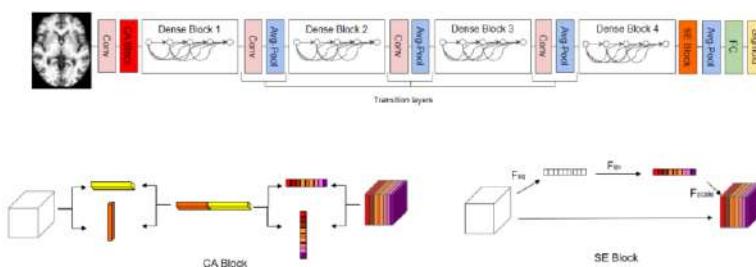


Figure 1. The architecture of dense attention connected convolution.

3.2.2. Dense Convolutional Network

The dense convolution network [16] does not directly add and fuse feature maps, but connects them on feature dimensions. Any input layer is connected with the subsequent layer, and the connected feature map is used as the input of the next layer:

$$X_L = H_L([X_0, X_1, \dots, X_{L-1}]) \quad (1)$$

In Eq. (1), L represents the number of network layers, and H_L represents a group of convolution operations at the L layer. Each layer is composed of batch normalization (BN), correction linear unit (RELU), pooling and convolution (CONV). $[X_0, X_1, \dots, X_{L-1}]$ represents the feature graph connection of the previous $L-1$ layer, and X_L represents the feature graph obtained after the convolution operation of the L layer. The Dense connection structure in a densely connected convolution network needs to keep the size of the feature image constant, so the Dense Block and Transition Layer are stacked. After the last dense block, the global average pooling layer is used to reduce the number of parameters. Each group of dense blocks is composed of the batch normalization layer (BN) and the convolution kernel of the RELU activation function is $1 \times 1 \times 1$ and $3 \times 3 \times 3$. The number of feature graphs output in each group is $R = 32$, which is used to control the number of network parameters. The transformation layer consists of a $1 \times 1 \times 1$ convolution layer and a $2 \times 2 \times 2$ pooling layer to reduce the resolution of feature maps.

3.2.3. Coordinate Attention

CA [17] proposed a novel attention network by embedding location information into channel attention, which decomposed attention into two features in different directions by global pooling, and then aggregated features along with these two directions respectively. Given the input feature x and compress them into $h \times 1 \times d$ and $1 \times w \times d$ in horizontal and vertical directions respectively, as shown in Eq. (2), Eq. (3), h , w , d represent the length, width and number of slices of the picture respectively. The horizontal and vertical pooling results were connected together, and the feature images were obtained through a $1 \times 1 \times 1$ convolution. The feature images were then divided into two groups of feature images in horizontal and vertical directions. The weight data were obtained by $1 \times 1 \times 1$ convolution and sigmoid respectively, and finally the obtained horizontal and vertical weights were multiplied to obtain the attention output feature.

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i, d) \quad (2)$$

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w, d) \quad (3)$$

3.2.4. Squeeze and Excitation Attention

SE [18] is an attention model based on channel dimensions. F_{sq} , F_{ex} and F_{scale} are the three key operations of the compression activation module. Firstly, the input feature graph is compressed to $1 \times 1 \times 1$ by F_{sq} operation to keep the number of feature images unchanged. Given the input feature x , the compression formula for channel C is Eq. (4). By activating F_{ex} operation, weight from 0 to 1 is generated for each feature graph. Redefine the relationship of each channel in the original feature graph, and F_{scale} adds weight to the picture. Channel attention networks often use global pooling to encode spatial

information and compress space into channel descriptors, ignoring location information. In this paper, SE is placed behind the last Dense Block.

$$Z_C = \frac{1}{H \times W \times D} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^D x_c(i, j, k) \quad (4)$$

4. Results and Discussion

4.1. Evaluation Metrics

We designed the classification of AD and CN and the predictive of sMCI and pMCI in two groups of controlled experiments. The performance of the classification is evaluated against six commonly used indicators. It includes accuracy, sensitivity, Specificity, Precision, Recall and F1 Score. These indicators are defined as Eq. (5) through Eq. (10):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (8)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (9)$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

TP indicates that the category is positive and the prediction is also positive, FP means the prediction is positive and the prediction is wrong, so it's a negative example. FN indicates that the prediction is negative and the prediction is wrong, so it is a positive example, TN indicates that the category is negative and the prediction is also negative.

4.2. Experimental Evaluation

In this part, we compared our experimental results with the most advanced convolutional neural networks, such as ResNet18 [19], DenseNet201, DenseNet264. We added attention to these networks to study the comparison between attentional mechanisms and those without. All experiments were implemented based on PyTorch with python3.8 programming language, executed on a server with an Intel Xeon Cascade Lake 8255C (2.5 GHz), Tesla V100-NVLINK-32G and Ubuntu18.04 operating system. The learning rate was attenuated from 0.0001 to 0.000001. In order to optimize the model, Adam was used as the optimization algorithm and the batch size was 8.

In this experiment, two groups of classification experiments will be carried out. Since the brain of AD and NC is quite different, it is easy to distinguish the two types. MCI is the early stage of AD, and there are only slight brain lesions. There is no significant difference in the changes of the lesion areas in the brain. The purpose of adding an attention mechanism is to make the model focus on the significant areas in the training process, excluding the influence of other organizational structures. sMCI and pMCI classification is a difficult task, and their classification accuracy is far lower than

that of AD and NC. The experimental results of these two groups of classification on different convolution models are shown in Table 2 and Table 3.

Table 2. Results of Classification for AD vs. NC.

Model	Accuracy	Sensitivity	Specificity	Precision	Recall	F1 Score
ResNet18	0.869	0.825	0.903	0.866	0.845	0.825
CSE_ResNet18	0.883	0.825	0.927	0.896	0.859	0.825
DenseNet201	0.876	0.730	0.987	0.978	0.836	0.730
CSE DenseNet201	0.890	0.809	0.951	0.927	0.864	0.809
DenseNet264	0.897	0.825	0.951	0.928	0.873	0.825
CSE DenseNet264	0.904	0.920	0.891	0.965	0.892	0.920

Table 3. Results of Classification for sMCI vs. pMCI.

Model	Accuracy	Sensitivity	Specificity	Precision	Recall	F1 Score
ResNet18	0.770	0.740	0.788	0.677	0.707	0.740
CSE_ResNet18	0.791	0.722	0.833	0.722	0.722	0.722
DenseNet201	0.791	0.685	0.855	0.740	0.711	0.685
CSE_DenseNet201	0.812	0.611	0.933	0.846	0.709	0.611
DenseNet264	0.819	0.740	0.866	0.769	0.740	0.754
CSE DenseNet264	0.826	0.796	0.844	0.754	0.774	0.796

As can be seen from the table, the classification accuracy of sMCI and pMCI in our experiment is 0.826, the sensitivity is 0.796, and the specificity is 0.844, which is superior to other models in overall classification. By adding attention mechanism to other models, the accuracy of sMCI and pMCI was improved, indicating that attention mechanism has a good effect on the prediction of MCI to AD transformation.

At the same time, we compare our method with the most advanced method in the present paper [20], [21], [14], [22] as shown in Table 4 for MCI to AD conversion of prediction of our method is better than the existing methods, and different papers with convolution neural network contrast experiment, it shows our network for locating sMCI and pMCI lesion area has obvious effect.

Table 4. Compare with the latest method on ADNI dataset.

References	AD vs NC			sMCI vs pMCI		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Cao et al. 2017	0.89	0.86	0.90	0.70	0.67	0.72
Liu et al. 2018	0.91	0.88	0.93	0.76	0.42	0.82
Lian et al. 2020	0.90	0.82	0.96	0.81	0.53	0.85
Zhu et al. 2021	0.92	0.91	0.93	0.80	0.77	0.82
Ours	0.90	0.92	0.89	0.83	0.80	0.84

4.3. Discussion

In order to understand how the network improves the classification effect, we conducted two groups of comparison experiments with and without attention in Densenet264. Figure 2 is the visual image of the experiment's features. pMCI and sMCI are the input of our model, and Conv1 represents the output of Conv1 without CA. CA Block represents the output of CA Block after adding the CA model. It was obvious from the images that compared with Conv1 images, CA Block images had more prominent

features in the hippocampus and cerebral cortex, while other useless features were ignored, which happened to be important diagnostic regions of Alzheimer's disease, proving that our model was superior to other models in feature extraction.

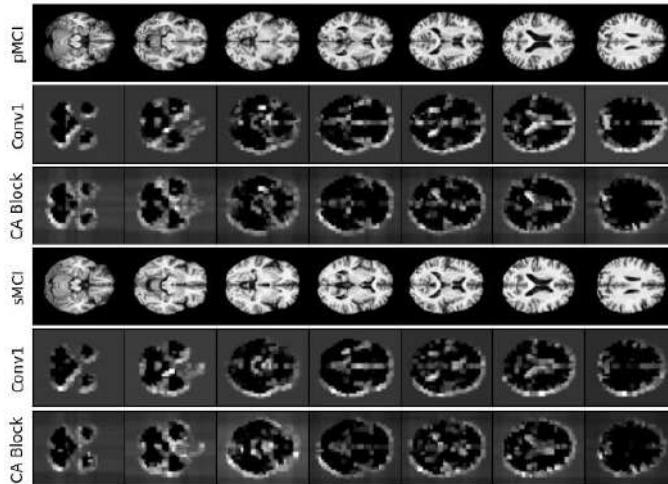


Figure 2. Visualization results of Densenet264 and CSE_Densenet264 feature maps.

5. Conclusion

In this paper, by adding attention mechanism to the model, significant information of MRI images can be extracted and combined with local and global information, so as to improve the accuracy of network classification. This method combines CA and SE attention and uses a dense attention network to automatically identify MRI lesions, for the early diagnosis and intervention treatment of AD, we put forward the method compared with the most advanced several methods has a better classification performance, especially in predicting the difficult MCI to AD transition task. MRI images can provide relatively simple information. In the future, we may add multimodal information, such as fluorodeoxyglucose positron emission tomography (FDG-PET), apolipoprotein-E (APOE) genotype, and age information as auxiliary information to further predict MCI to AD transformation.

Acknowledgment

This work is supported by multiple grants, including: The National Key Research and Development Program of China (2020YFB1313900), National Natural Science Foundation of China (61902386), Shenzhen Science and Technology Program (JCYJ20180507182415428).

References

- [1] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement*. 2007 Jul;3(3):186-91.
- [2] Ben Ahmed O, Mizotin M, Benois-Pineau J, Allard M, Catheline G, Ben Amar C. Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease diagnosis on structural MR images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. *Comput Med Imaging Graph*. 2015 Sep; 44: 13-25.
- [3] Li F, Tran L, Thung KH, Ji S, Shen D, Li J. A robust deep model for improved classification of AD/MCI patients. *IEEE J Biomed Health Inform*. 2015 Sep;19(5):1610-1616.
- [4] Liu Z, Maiti T, Bender AR. for the Alzheimer's Disease Neuroimaging Initiative. A role for prior knowledge in statistical classification of the transition from mild cognitive impairment to alzheimer's disease. *J Alzheimers Dis*. 2021 Aug 25.
- [5] Sarica A, Cerasa A, Quattrone A. Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: a systematic review. *Front Aging Neurosci*. 2017 Oct 6; 9: 329.
- [6] Calvini P, Chincarini A, Gemme G, Penco MA, Squarcia S, Nobili F, Rodriguez G, Bellotti R, Catanzariti E, Cerello P, De Mirti I, Fantacci ME; MAGIC-5 Collaboration; Alzheimer's Disease Neuroimaging Initiative. Automatic analysis of medial temporal lobe atrophy from structural MRIs for the early assessment of Alzheimer disease. *Med Phys*. 2009 Aug;36(8):3737-47.
- [7] Qiu S, Joshi PS, Miller MI, Xue C, Zhou X, Karjadi C, Chang GH, Joshi AS, Dwyer B, Zhu S, Kaku M, Zhou Y, Alderazi YJ, Swaminathan A, Kedar S, Saint-Hilaire MH, Auerbach SH, Yuan J, Sartor EA, Au R, Kolachalam VB. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain*. 2020 Jun 1;143(6):1920-1933.
- [8] Puente-Castro A, Fernandez-Blanco E, Pazos A, Munteanu CR. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Comput Biol Med*. 2020 May; 120:103764.
- [9] Klöppel S, Stonnington CM, Chu C, Draganski B, Seahill RI, Rohrer JD, Fox NC, Jack CR Jr, Ashburner J, Frackowiak RS. Automatic classification of MR scans in Alzheimer's disease. *Brain*. 2008 Mar; 131(Pt 3):681-9.
- [10] Vemuri P, Gunter JL, Senjem ML, Whitwell JL, Kantarci K, Knopman DS, Boeve BF, Petersen RC, Jack CR Jr. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *Neuroimage*. 2008 Feb 1;39(3):1186-97.
- [11] Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Alzheimer's disease neuroimaging initiative. multimodal classification of alzheimer's disease and mild cognitive impairment. *Neuroimage*. 2011 Apr 1; 55(3): 856-67.
- [12] Sørensen L, Igel C, Liv Hansen N, Osler M, Lauritzen M, Rostrup E, Nielsen M. Early detection of Alzheimer's disease using MRI hippocampal texture. *Hum Brain Mapp*. 2016 Mar;37(3):1148-61.
- [13] Oh K, Chung YC, Kim KW, Kim WS, Oh IS. Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Sci Rep*. 2019 Dec 3;9(1):18150.
- [14] Lian C, Liu M, Zhang J, Shen D. Hierarchical fully convolutional network for joint atrophy localization and alzheimer's disease diagnosis using structural MRI. *IEEE Trans Pattern Anal Mach Intell*. 2020 Apr; 42(4):880-893.
- [15] Zhao Y, Ma B, Jiang P, Zeng D, Wang X, Li S. Prediction of alzheimer's disease progression with multi-information generative adversarial network. *IEEE J Biomed Health Inform*. 2021 Mar;25(3):711-719.
- [16] Huang G, Liu Z, Laurens VDM, Weinberger KQ. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 2261-2269.
- [17] Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. *CVPR*. 2021.
- [18] Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-excitation networks. *IEEE Trans Pattern Anal Mach Intell*. 2020 Aug; 42(8):2011-2023.
- [19] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 .pp. 770-778.
- [20] Cao P, Liu X, Yang J, Zhao D, Huang M, Zhang J, Zaiane O. Nonlinearity-aware based dimensionality reduction and over-sampling for AD/MCI classification from MRI measures. *Comput Biol Med*. 2017 Dec 1; 91: 21-37.
- [21] Liu M, Zhang J, Adeli E, Shen D. Landmark-based deep multi-instance learning for brain disease diagnosis. *Med Image Anal*. 2018 Jan; 43:157-168.
- [22] Zhu W, Sun L, Huang J, Han L, Zhang D. Dual attention multi-instance deep learning for alzheimer's disease diagnosis with structural MRI. *IEEE Trans Med Imaging*. 2021 Sep;40(9):2354-2366.

Study on Dune Morphology in Computer Mapping Based on 3S Technology

Zhidan BA¹

College of Tourism and Geographical Sciences, Jilin Normal University, China

Abstract. Take the parabolic dune of Hobq desert in Inner Mongolia as research object. Based on the GIS platform by using differential GPS data and spatial interpolation to generate DEM, then using Multi-periods high resolution images to acquire the environmental background, at the same time combine with regional wind regime and vegetation condition to measure and analyze the morphology of the parabolic dune. The result shows that the parabolic dune showed U shape in plane, and dune arms point to the west which was also wind direction. The windward slope of longitudinal profile is gentler, while leeward slope is steeper. And cross section wasn't symmetric. The dune's average moving speed is 11.76 m/yr. Desert ridge line's medial axis direction is WNW-ESE, in accord with the direction of prevailing wind and resultant drift potential. *Artemisia Ordosicas* mainly distribute on leeward slope, two arms, and the plane ground between them, and the annual average vegetation coverage decreased 0.95%. In the long-term effect of resultant wind, the dune keeps moving forward and *Artemisia Ordosica* between two arms show gradual natural stage recovery which presented zonal distribution. 3S technology has already become important research method in modern Aeolian sand morphology.

Keywords. Parabolic dune, Morphological, 3S technology, Hobq

1. Introduction

Parabolic dune is u-shaped or horse shoe shaped in plane, the two arms of the dune extend to the upper wind drift, and the windward slope is gently and concave, while the leeward slope is steep and convex [1-2] and from the elongation of blowout hollows [3]. At present, research on the dynamics of parabolic dunes has focused on several key areas including dune distribution and morphology [4-6]; timing of past dune activity [7-8]; rates of migration [9]; internal structure [10]; and contemporary processes [11].

3S technology is short for remote sensing (RS), Geography information system (GIS) and Global positioning systems (GPS) [12], and it is an important method in modern geography research [13]. At present, the most common way to acquire the 3D data of Aeolian geomorphology in moderate-small scale is ground measurement by using GPS [14], electronic total station [15], 3D laser scanner [16] and LIDAR [17]. Digital elevation model (DEM) is the basis of geography analysis [18], effective method to research the surface morphological character, and its change law [19].

To address these issues, the spatial and temporal patterns of sediment transport at monthly to annual timescales were documented across an active parabolic dune. This

¹ Corresponding Author, College of Tourism and Geographical Sciences, Jilin Normal University, China; e-mail: bazhidan1999@163.com.

was accomplished by repeat measurements from RTK GPS over a four-year period and from a grid of erosion pin over a two-year period. These data were used to map the morph dynamics of the dune and to develop indices of sediment transport, which are compared with nearby meteorological data to examine the effects of variability in atmospheric conditions on sediment transport. In addition, this paper takes the typical parabolic dune in the edge of Hobq desert as the research object, and use 3S technology to do the research in morphological changes of the dune.

2. Study Area and Data Acquisition

Hobq desert is in the northern part of the east-west uplift belt in Ordos Plateau, and locate in Hangjinqi, Dalateqi and Zhungerqi of Yikezhao league in Inner Mongolia ($109^{\circ}30' - 110^{\circ}45'E$, $39^{\circ}48' - 40^{\circ}33'N$). It is higher in South than North and surrounded by the yellow river in West, North and East. It is temperate arid and semi-arid zone, 25 to 35 days of strong wind a year, annual mean temperature is $7.4^{\circ}C$, annual precipitation is 249 mm, average annual evaporation 2 450 mm, and frostless period is 156 days. The moving dunes are densely distributed in the northern part of the desert, while in the southern scab rock, flaky moving dune, shrub and *Nabkha* are distributed and among them there are some parabolic dunes.

We use Leica 1 230 differential GPS to do the morphological measurement; the instrument's marked precision in horizontal is $10mm + 1ppm$, and vertical is $20 mm + 1 ppm$. We collected the information, such as the coordinate system and central meridian about the research area before measuring. We embedded more than three marked stone where the blown sand does not influence around the dune as the control points which are used to check the random error when continuous measuring and in order to cooperate with the total station. To measure the control points in static model of the GPS simultaneous, Since all the surveying and mapping data accumulated in Early China were Beijing 54 coordinates, in order to make comparative analysis with the data results of existing coordinates, so change the WGS-84 coordinate system to Beijing-54. Using the dynamic model to measure the WGS-84 coordinates of the control points with moving station and then change them to Beijing-54 coordinates by seven-parameter method, finally measuring the detail points of the dune. The data used in this paper were collected in Oct 2009; May, Oct, and Dec 2010; May, Jul, Sep 2011. The error of the data which meet the precision requirement is less than 20 mm.

In order to research the change of vegetation coverage around the dune, the remote sensing data were needed to analyze it in large time-span, such as Quick Bird image in Aug 2005, Feb 2007, and May 2010. The reason is to try to ensure that the measurements are close to the dune morphology, because commercial companies do not have data from other times.

3. Results and Analysis

3.1. DEM Acquisition and its Features

It is necessary to generate DEM data by spatial interpolation, because the data acquired from the differential GPS is spatial discrete coordinates, and this paper use the

neighborhood interpolation method of 3D Analyst module of ArcGIS 9.3 to realize it, the neighborhood interpolation uses an algorithm that finds the subset of input samples closest to the query point and applies weights to these samples proportionally based on the region size to interpolate. We use the elevation points by spatial interpolation to generate the raster surface, and then generate contour and finally generate TIN DEM (Figure 1).

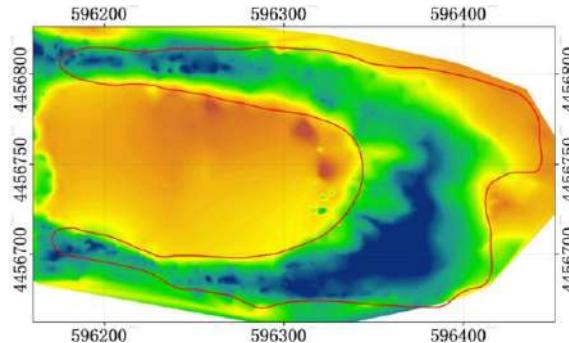


Figure 1. DEM of parabolic.

According to the study area investigation and DEM analysis, the parabolic dune showed U shape in plane, and dune arms where grow *Artemisia Ordosica* pointing to the west in accordance with up windward. The windward slope is concave, while the leeward is convex. Until now, the length of the dune is 262.45 m, width is 147.09 m, relative height is 5.73 m, the length of the arms is 168.66 m, the inner width between arms is 91.36 m, the plane area is 18 600.58 m², and the volume is 40 239.06 m³. The toe of the windward slope outcrops soil layer, and in the middle of it develops four blowouts which developed and expanded. While in the leeward slope, the gradient of sand fallen slope is increasing, and the slip face is becoming higher. According to the length-width ratio, it is leaf shape (LLWR is between 1.0 to 3.0) has the length-width ratio features of relatively mature.

In the research period, the dune has been moving to West. The toe of windward slope advanced 23.22 m (11.61 m/yr), while the leeward advanced 23.52 m (11.76 m/yr), and relative height has decreased 0.78 m (0.39 m). The plane area of dune has increased 1 901.70 m² (950.85 m²/yr), because nose moved along the prevailing wind direction, while arms relatively fixed. The volume decreased 840.78 m³ (420.34 m³/yr), because undulant underlying surface and a small amount of sand material being blown away.

3.2. Profile and its Morphological Parameters

Based on the DEM, extract and analyze the longitudinal and cross profile by means of using the 3D Analyst module in ArcGIS 9.3. That means using the Interpolate line tool to make sure the location and trend of profile line then generate the graph of profile line by using create profile graph tool.

As it is shown in the profile graph (Figure 2), the geomorphology units of longitudinal profile of the dune are flat ground in two arms, windward slope, dune top, and leeward slope. At present, the length of flat sandy land in two arms is 167.72 m where growth the *Artemisia Ordosica*. The length of windward slope is 46.54 m, the

surface of the slope is concave, the gradient is between 16° to 21° , and in the bottom distributes some small wind-erosion residual frusta, and in top of the dune has strong wind erosion. The length of leeward is 32.71 m; the gradient is between 27° to 36° , the main body of dune move forward in the form of slip face collapse.

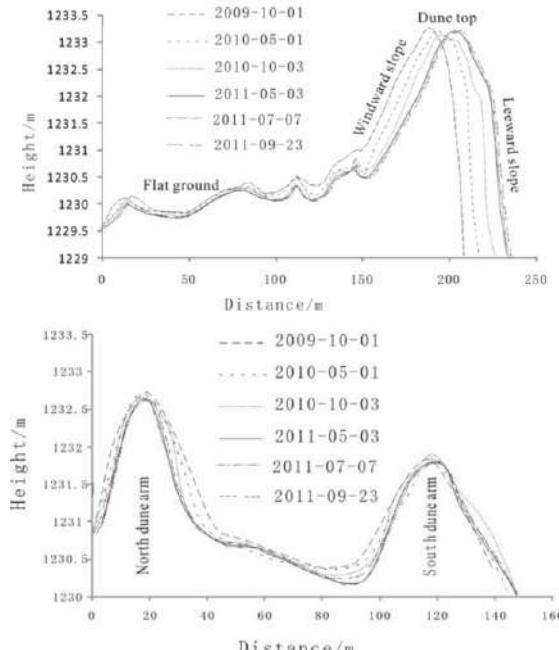


Figure 2. Profile of dune (left longitudinal profile; right, cross profile).

3.3. Ridge Line of the Dune and Wind Regime Analysis

Using the method mentioned above to extract the plan curvature of negative landform B. Fourthly, use the formula $[[[A+B]-[A-B].abs]/2]$ to acquire the plan curvature of DEM without error. Lastly, adjust the threshold to extract the ridge line.

Ridge line is U shaped, and it has the characteristics of typical parabolic dune's ridge line. The axis is WNW-ESE. The ridge line moves forward along with the dune, and it advanced 22.50 m (14.21 m/yr).

3.4. Morphological Changes and Vegetation Distribution

Vegetation distribution characteristic is a significant factor to control the erosion-deposition pattern and to effect its formation and development. A little change of vegetation coverage can make great effect in dune morphology. The previous method to measure the vegetation coverage was using the GPS or electronic total station dotting the vegetation boundary and then sketch out the boundary in computers [19], while the vegetation coverage acquired mainly by sampling surveys estimated by researchers in the field.

According to the Quick bird image and interpretation graph (Figure.3), vegetation coverage of the dune has decreased from 36.81% in 2005 to 32.04% in 2010. The main

reason is that the dune moved forward and arms are elongating which makes the area of vegetation is increasing though the area of dune is increasing.

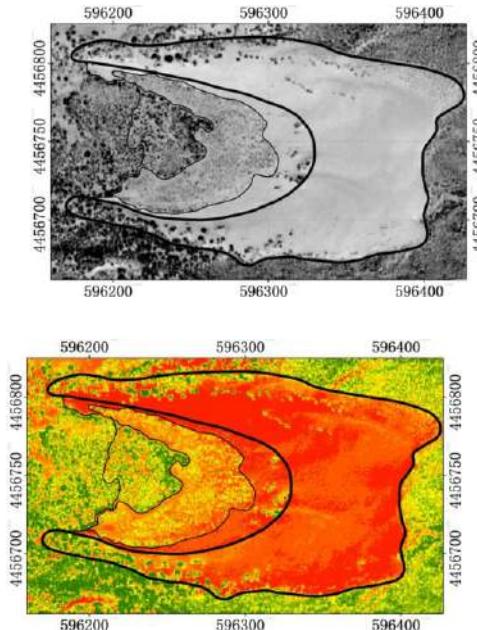


Figure 3. Quick bird panchromatic image and interpretation graph.

4. Conclusion

Based on the GIS platform, using the measured data acquired from differential GPS to generate DEM by spatial interpolation. The terrain factors and morphological parameters can be extracted effectively, in order to realize the quantitative research about Aeolian morphological changes.

The parabolic dune that we research showed U shaped in plane, the arms point to west which was also wind direction, windward slope is concave and leeward slope is convex. Because the dune is located in the edge of desert, the relative height is low; its move speed is fast and advanced 11.76 m annually. The plane area of dune has annual increased 950.85 m² due to nose moved along the prevailing wind direction, while arms relatively fixed.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (NO. 41871022).

References

- [1] Wu Z. Geomorphology of wind-drift sands and their controlled engineering, Science Press, Beijing, 1995.
- [2] Zhu ZD, Wu Z, Liu S et al. Introduction to Chinese desert, Science Press, Beijing, 1980.
- [3] Fernandez-Cascales L, Lucas A, Rodriguez S, et al., First quantification of relationship between dune orientation and sediment availability, Olympia Undae, Mars, Earth and Planetary Science Letters, 489(2018), 241-250.
- [4] Wang P, Ma Q, Zhu YP, et al., Particle size characteristics of surface sediments of scrub sand piles and parabolic dunes in the Tokay Desert, Xinjiang and their depositional environment, Arid Zone Geography, 10(2021), 1-13.
- [5] Yan NH, Liu HQ, et al., Research progress in morphology and evolution of the parabolic dunes, Journal of Desert Research, 30(2010), 801-807.
- [6] Kathryn LR, Geoff JV, Tim DF. Urban sediment supply to streams from hill slope sources, Science of the Total Environment, 653(2019), 684-697.
- [7] Dong YX, Huang DQ, A typical study of coastal barchans dune movement and morphological change, Geosciences, 34(2014), 863-869.
- [8] John TK, Andrew JM, Claire W. Spatial and temporal patterns of suspended sediment transport in nested urban watersheds, Geomorphology, 336(2019), 95-106.
- [9] Srivastava A, Thomas DSG, Durcan JA. Holocene dune activity in the thar desert, India, Earth Surface Processes and Landforms, 44(2019), 0197-9337.
- [10] Tripaldi A, Mehl A, Zarate MA. Parabolic mega dunes in a subtropical quaternary inland dune field, southwestern pampas, Argentina, Geomorphology, 321(2018), 103-116.
- [11] Cheng RT, Cai LF. Ship borne radar calibration technology based on differential GPS and radar beacon, Ship Electronic Engineering, 41(2021), 76-79+88.
- [12] Zhang LW, Lu HF. Summary of application status of 3S technology in coal mine geological hazards, Software Engineering and Applications, 10(2021), 549-556.
- [13] Emmanouil P, Konstantinos XS, Melpomeni Z, et al., Synergistic approach of remote sensing and gis techniques for flash-flood monitoring and damage assessment in thessaly plain area, Greece, Water, 11(2019), 448.
- [14] Zhong YF, Yi MS. Application of 3D laser scanner in mine survey, Low-carbon world, 11(2021), 66-67.
- [15] Baptista PC, Cunha TR, Matias AG, et al. New land-based method for surveying sandy shores and extracting DEMs: the INSHORE system, Environmental Monitoring and Assessment, 182(2011), 243-257.
- [16] Hugenholz CH. Topographic changes of a supply-limited inland parabolic sand dune during the incipient phase of stabilization, Earth Surface Processes and Landforms, 35(2010), 1674-1681.
- [17] Liu B, Yu Y, Jiang M. Review of advances in LIDAR detection and 3D imaging, Photoelectric Engineering, 46(2019), 21-33.
- [18] Keck H, Fel De V, Drahorad SL, et al. Biological soil crusts cause subcritical water repellency in a sand dune ecosystem located along a rainfall gradient in the NW Negev desert, Israel, Journal of Hydrology and Hydromechanics, 64(2016), 133-140.
- [19] Chen XL, Li ZZ, Chen SJ et al., The morphological character research of parabolic dune-based on the date of differential gps survey. Journal of Fujian Normal University (Natural Science Edition), 26(2010), 84-89.

A Study of Forest Swamp Mapping in Hani Wetland Integrating Sentinel-1 and Sentinel-2 Satellite Images

Jingfa WANG¹

College of Tourism and Geographical Science, Jilin Normal University, Siping, Jilin 136000, China

Abstract. As a unique wetland type, forest swamps play an important role in regional carbon cycling and biodiversity conservation. Taking Hani wetland in Jilin province as the research object, we integrated the application of Sentinel-1 radar and Sentinel-2 multispectral images, fully exploited the potential of Sentinel-1 multi-polarization band features and Sentinel-2 red edge index for forest swamp remote sensing identification, and applied the random forest method to realize the extraction of forest swamp distribution information of Hani wetland. The results show that when the optimal number of decision trees for forest swamp information extraction is 1200, the fusion of Sentinel-1VV and VH backscattering coefficient radar band features and Sentinel-2 red-edge band features can significantly improve the extraction accuracy of forest swamp distribution information, and the overall accuracy and Kappa coefficient of forest swamp information extraction in protected areas are as high as 89% and 0.85, respectively. The overall accuracy and Kappa coefficient of forest swamp information extraction in the protected area were 89% and 0.85, respectively. The landscape types of Hani Wetlands of International Importance are diversified, with natural wetlands, artificial wetlands and non-wetland landscape types co-existing. Among the natural wetland types, the forest swamp has the largest area of 27.1 km², accounting for 11.2% of the total area of the reserve; the river has the smallest area of 0.7 km², accounting for 0.3% of the total area of the reserve. The forest swamp extraction method provides data support for the sustainable management of Hani wetlands and case guidance for forest swamp mapping in other regions.

Keywords. forest swamp, Sentinel, random forest, mapping of remote sensing

1. Introduction

Wetlands are one of the most important types of terrestrial ecosystems and play an extremely important role in maintaining the balance of material cycles, protecting the diversity of biological species and maintaining ecological security [1-3]. Forested bogs are boggy wetlands with woody plants of more than 6 m as the dominant species, which are extremely rich in carbon stocks and respond significantly to global climate change [4-6]. Facing the severe challenges of both human activities and climate change, remote sensing monitoring of forest swamps can clarify the distribution boundaries of forest swamps in time, reflect the current status and transformation trends of forest

¹ Corresponding Author: College of Tourism and Geographical Science, Jilin Normal University, Siping, Jilin 136000, China; E-mail: jlnuwangjingfa@163.com.

swamp wetland landscape types, which is of great significance to effectively protect the health of forest swamp ecosystems and maintain the stability of regional ecosystems [7, 8]. Multi-source remote sensing data integration can maximize the fusion of different types of sensor remote sensing data, effectively improve the accuracy of feature type identification and extraction, and multi-source remote sensing data fusion is widely used in the extraction of wetland cover information research [9, 10]. With the rapid development of satellite remote sensing technology, the types of remote sensing data sources are constantly enriched. Sentinel-1 synthetic aperture radar (SAR) is different from passive optical remote sensing, which can effectively avoid the influence of clouds, rain, fog and no light at night, and has the advantages of high resolution, dual polarization, short period, strong penetration, multi-mode, all-weather, all-day and rapid product production, which can accurately determine the satellite position and attitude angle, and realize the high-resolution monitoring of global land, coastal zone, airline routes and large area coverage monitoring of global ocean [11, 12]. Sentinel-2 A and B satellite systems are equipped with Multispectral Instrument (MSI) sensors, which can carry out ground monitoring in 13 bands of visible, near-infrared, short-wave infrared and red-edge with a 5d cycle, especially the red-edge range of three bands has great advantages for vegetation growth and health monitoring, and the high spatial and temporal resolution and wide spectral resolution make Sentinel-2 remote sensing images commonly used in monitoring the evolution of terrestrial ecological environment, wetland landscape mapping, vegetation crop growth monitoring and disaster warning mapping [13, 14].

The Hani Wetland is the third wetland of international importance in Jilin Province and is extremely rich in freshwater, forest swamp and other wetland resources. It has the largest peat swamp deposit in northeast China, which is of great significance in terms of water conservation and provision, climate regulation and scientific research on peat swamp wetlands. This study takes Hani wetland in Jilin Province as the research object, integrates Sentinel-1 radar and Sentinel-2 multispectral images, fully exploits the potential of Sentinel-1 multi-polarization band feature and Sentinel-2 red edge index for forest swamp remote sensing identification, applies random forest method to realize the extraction of forest swamp distribution information in Hani wetland, and the forest swamp. The extraction method can provide case guidance for forest swamp mapping in other regions and provide data support for sustainable management of Hani wetlands.

2. Study Area and Data Processing

2.1. Study Area

Jilin Hani Wetland ($42^{\circ} 4' 12'' \sim 42^{\circ} 14' 30''$ N, $126^{\circ} 4' 9'' \sim 126^{\circ} 33' 30''$ E) is located in the southeast of Liuhe County, Tonghua City, Jilin Province, adjacent to Tonghua County in the south, and bordered by Baishan City and Jingyu County in the east, with an approximate "northeast-southwest" strip distribution, including three parts: core area, buffer area and pilot area. The study area is located in the temperate humid and semi-humid monsoon climate zone, with four distinct seasons, high temperature and humidity in summer, cold and dry in winter, rain and heat in the same period, annual average temperature 5.01°C , annual average precipitation 755.52 mm,

annual frost-free period 125-145 d. The geomorphological type is mainly basalt lava mountains, with high terrain in the east and low terrain in the west, with an average altitude of 884.50 m.

2.2. Data Source and Processing

The Sentinel-1 Ground Range Detected (GRD) polarized radar image data applied in this study were downloaded from ESA Opernicus, the image acquisition instrument was SAR-C, the collection mode was IW, the data size was about 1.64 G, and the data quality was high [15]. Sentinel-2 remote sensing images with rich spectral resolution are widely used in the fields of terrestrial vegetation, soil and water monitoring, grain yield estimation and wetland landscape mapping research, etc. The Sentinel-2 level-1C product was atmospherically corrected and radiometrically calibrated, and each band was super-resolved and resampled to 10 m [16].

In this study, 1942 training samples were obtained through visual interpretation and field GPS survey, of which 502 were the largest number of training samples from forested areas and 105 were the smallest number of training samples from paddy fields, which generally met the conditions of "uniform spatial distribution and sufficient number" and satisfied the experimental classification requirements. Based on the selected number of training samples, 50% of the training samples were selected as the set of validation points to test the final accuracy of the wetland classification results, with a total of 971 validation points, including 91 forest swamps.

With reference to the definition of wetlands in the international Convention on Wetlands and related literature, a classification system was established in line with the forest swamp wetland classification of Jilin Hani National Nature Reserve (Table 1).

Table 1. Classification system of swamp in the Hani Ramsar site.

The primary classification	The secondary classification	Definition and basis of division
Natural wetlands	River	A river having annual, seasonal, or intermittent river runoff
	Lake	A permanent, seasonal, or intermittent lake consisting of fresh, brackish, salt, or salt water
	Swamp	A freshwater marsh dominated by thickets or trees
	Marsh	A freshwater marsh consisting of dominant communities of aquatic or bog herbs
Artificial wetland	Paddy field	A ridge is enclosed to store water and plant paddy fields of one, two or three seasons
	Dry farmland	Cultivated land without irrigation facilities that relies mainly on natural precipitation for xerophytic crops
Non-wetlands	Woodland	Land covered by patches of natural, secondary, and artificial forests
	Grassland	Land on which herbaceous and shrub plants are predominant and suitable for animal husbandry production
	Construction land	Land for artificial construction of buildings and structures

3. Research Methodology

3.1. Random Forest

Random Forest (RF) classification algorithm is a newly emerged, flexible and highly accurate predictive machine learning classification algorithm for calculators, which is commonly used in the extraction of wetland information from geological remote sensing images. Random Forest classification algorithm has the advantages of no worry about over-fitting, suitable for large data sets and a large number of unknown features, autonomous estimation of optimal feature sets, direct processing of high-dimensional feature input samples without dimensionality reduction, strong resistance to fabrication, easy to understand the algorithm content and the possibility of merging processing, etc. [17, 18].

3.2. Waveform Extraction and Feature Set Combination

Based on Sentinel-1 radar and Sentinel-2 multispectral remote sensing images, Sentinel-1 VV and VH polarized backscatter coefficients were extracted, and band features such as Sentinel-2 spectral features, red-edge index features, vegetation index features, water body index features and texture features were obtained by Band Maths band calculation. .

3.3. Precision Evaluation

The study used the data of 971 validation points obtained from field measurements and the classification results to establish the confusion matrix of wetland information extraction in Hani Reserve 2020, and calculated four evaluation indexes of overall accuracy (*OA*), kappa coefficient, producer accuracy (*PA*) and user accuracy (*UA*) to accurately and reasonably evaluate the accuracy of this forest swamp information extraction.

4. Results and Analysis

4.1. Optimal Random Forest Classification Tree

The overall accuracy and Kappa coefficient of random forest extraction of wetland cover information in Hani Reserve showed a trend of fluctuating and stabilizing after first increasing at the stage of 10~100 classification trees (Figure 1), and the overall accuracy of wetland information extraction in the reserve increased by 2.6%, 5.5%, and 3.2% at the stage of 10~100 classification trees for combinations 1, 2, and 3, respectively, and the Kappa coefficients increased by 0.03, 0.06, and 0.03, respectively. 0.06 and 0.03, respectively, and the kappa coefficients increased by 0.03, 0.06, and 0.03, respectively, with the increase of the number of classification trees, the wetland information extraction time of protected areas also increased accordingly. The optimal number of random forest classification trees for forest swamp information extraction was different for the three combinations of protected areas, and the optimal random

forest classification trees were 100, 400, and 1200 for combinations 1, 2, and 3, respectively.

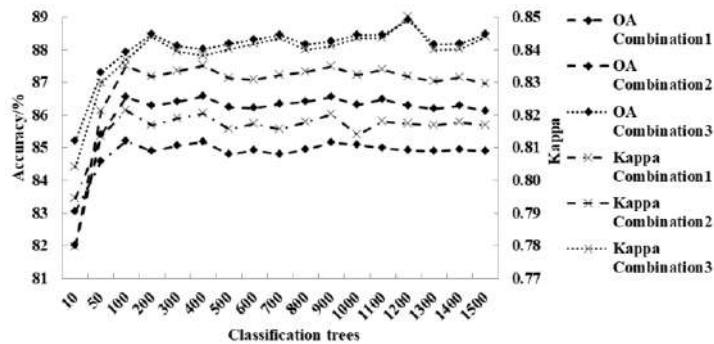


Figure 1. Varied classification accuracies among the different numbers of classification trees for different feature combinations.

4.2. Optimal Combination of Features

The combination of Sentinel-1 radar VV and VH backscatter coefficient bands with Sentinel-2 red-edge bands3 forest swamp information extraction accuracy is the highest, the overall accuracy and Kappa coefficient are as high as 88.9%, 0.85, where the producer accuracy of forest land, forest swamp, dry land, and grassland are as high as 97.9%, 93.8%, 90.6%, and 90.1%, and the user accuracies were as high as 96.6%, 89.5%, 89.4%, and 86.9% for woodland, paddy field, forest swamp, and grassland, respectively, with relatively high extraction quality for forest swamp, followed by combination 2 which only incorporates Sentinel-2 red-edge bands, and the smallest overall accuracy and Kappa coefficient for combination 1 wetland information extraction.

4.3. Spatial Patterns of Hani Internationally Important Wetlands

The total area of Hani Reserve is about 242.5 km² (Figure 2), among which, the natural wetland area is 30.6 km², and the forest swamp is the largest among the natural wetland types, 27.1 km², accounting for 88.6% of the natural wetland area and 11.2% of the total area of the reserve, mainly distributed in the relatively high altitude area in the northeast of the reserve; the river area is the smallest, 0.7 km², accounting for only 2.2% of the natural wetland area and 0.3% of the total area of the reserve. The area of rivers is the smallest, 0.7 km², accounting for only 2.2% of the area of natural wetlands and 0.3% of the total area of the reserve, running from northeast to southwest; the area of lakes and herbaceous marshes is 1.3 km² and 1.5 km² respectively, accounting for 4.4% and 4.9% of the area of natural wetlands and 0.6% and 0.6% of the total area of the reserve respectively. The area of artificial wetland paddy field is 0.9 km², which is 0.4% of the total area of the reserve, mainly distributed in the gentle terrain on both sides of the river; the area of non-wetland types in the reserve is relatively large, and the area of forest land is the largest, up to 192.3 km², which is 79.3% of the total area of the reserve; the area of construction land is the smallest, 3.6 km², which is 1.5% of the total area of the reserve; the area of dry land and grassland is 5.2 km², respectively. The area

of dry land and grassland is 5.2 km² and 9.9 km² respectively, accounting for 2.2% and 4.1% of the total area of the reserve.

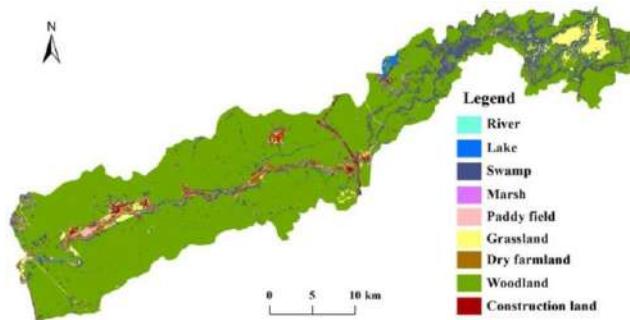


Figure 2. Land cover distribution patterns in the Hani Ramsar wetland site.

5. Conclusion

The accuracy of the random forest method for extracting forest swamp and wetland landscape information is governed by various factors such as the type of study area, the number of classification trees, and the selection of training and validation samples, etc. The optimal random forest classification tree for extracting forest swamp and wetland landscape information of Hani International Important Wetland is 1200.

The fusion of Sentinel-1 VV and VH backscatter polarization and Sentinel-2 red-edge feature bands can significantly improve the accuracy of forest swamp landscape information extraction, the overall accuracy of Hani internationally important wetland landscape information extraction is 88.9%, the Kappa coefficient is 0.85, and the accuracy of forest swamp extraction producers and users are as high as 93.8% and 89.4%, respectively.

The wetland landscape types in the reserve are diverse, with natural wetlands, artificial wetlands and non-wetland landscape types co-existing. Among the natural wetland types, the forest swamp has the largest area of 27.1km², accounting for 11.2% of the total area of the reserve, and is mainly distributed in the relatively high altitude area in the northeast of the reserve, while the river has the smallest area.

Acknowledgements

This research was funded by National Natural Science Foundation of China (NO.41871022).

References

- [1] Jin BCA., Jiang MY, Bai WR et al. Fluorescence spectra characteristics of reclaimed water to replenish constructed wetlands using EEM-PARAFAC, Spectroscopy and Spectral Analysis 41(2021), 1240-1245.
- [2] Islam AMT, Talukdar S, Mahato S, et al. Machine learning algorithm-based risk assessment of riparian wetlands in Padma River Basin of Northwest Bangladesh. Environmental Science and Pollution Research 28(2021), 34450-34471.

- [3] Song F, Su FL, Mi CX, et al. Analysis of driving forces on wetland ecosystem services value change: A case in Northeast China, *Science of The Total Environment* 751(2021), 141778.
- [4] Xulu S, Mbatha N, Peerbhay K, et al. Detecting harvest events in plantation forest using sentinel-1 and-2 data via google earth engine, *Forests* 11(2020), 1283.
- [5] Dong D, Wang C, Yan JH, et al. Combing Sentinel-1 and Sentinel-2 image time series for invasive *Spartina alterniflora* mapping on Google Earth Engine: a case study in Zhangjiang Estuary. *Journal of Applied Remote Sensing* 14(2020), 044504.
- [6] Mohammadi A, Kamran KV, Karimzadeh S, et al. Flood detection and susceptibility mapping using sentinel-1 time series, alternating decision trees, and bag-adtree models. *Complexity*, (2020), 4271376.
- [7] Pyankov SV, Maximovich NG, Khayrulina EA et al. Monitoring acid mine drainage's effects on surface water in the kizel coal basin with sentinel-2 satellite images. *Mine Water and the Environment*, 40(2021), 606-621.
- [8] Yang QQ, Yuan QQ, Li TW, et al. 2020. Mapping PM2.5 concentration at high resolution using a cascade random forest based downscaling model: Evaluation and application, *Journal of Cleaner Production*, 277(2020), 123887.
- [9] Braun A. Retrieval of digital elevation models from Sentinel-1 radar data - open applications, techniques, and limitations, *Open Geosciences*, 13(2021), 532-569.
- [10] Rumora L, Miler M, Medak D. Contemporary comparative assessment of atmospheric correction influence on radiometric indices between Sentinel-2A and Landsat 8 imagery, *Geocarto International*, 36(2021), 13-27.
- [11] Yahiaoui I, Bradai A, Douaoui A, et al. Performance of random forest and buffer analysis of Sentinel-2 data for modelling soil salinity in the Lower-Cheliff plain (Algeria), *International Journal of Remote Sensing*, 42(2021), 128-151.
- [12] Harper KA, Gray L, Querry ND. Spatial patterns of vegetation structure and structural diversity across edges between forested wetlands and upland forest in Atlantic Canada, *Canadian Journal of Forest Research*, 51(2021), 1189-1198.
- [13] Santos CF, Teixeira RC, Raizer J, et al. Post-fire phyllostomid assemblages in forest patches of the Pantanal wetland, *Mammalia*, 85(2021), 155-158.
- [14] Bonansea M, Ledesma M, Bazan R, et al. Evaluating the feasibility of using Sentinel-2 imagery for water clarity assessment in a reservoir, *Journal of South American Earth Sciences*, 95(2019), 102265.
- [15] Yousef B, Hossein H, Abbas M. Investigating the capabilities of multispectral remote sensors data to map alteration zones in the Abhar area, NW Iran, *Geosystem Engineering*, 24(2021), 18-30.
- [16] Clauss K, Ottiger M, Leinenkugel P, et al. 2018. Estimating rice production in the Mekong Delta, Vietnam, utilizing time series of Sentinel-1 SAR data, *International Journal of Applied Earth Observation and Geoinformation*, 73(2018), 574-585.
- [17] Chen ZT, Wang JF, Gao DY, et al. Dynamic spatial fusion of cloud vertical phase from CALIPSO and cloudsat satellite data. *Photogrammetric Engineering and Remote Sensing*, 87(2021), 61-67.
- [18] Chen R 2021. Application of UAV-low altitude remote sensing system in sea area supervision, *Earth Sciences Research Journal*, 25(2021), 65-68.

The RBF Hyperparameter in Metamodel-Assisted Evolutionary Search

Yoel TENNE ¹

Mechanical Engineering Department, Ariel University, Israel

Abstract. RBF metamodels, which are commonly used in expensive optimization problems, rely on a hyperparameter which affects their prediction. The optimal hyperparameter value is typically unknown and hence needs to be estimated by additional procedures. As such this study examines if this overhead is justified from an overall search effectiveness perspective, namely, if changes in the hyperparameter yield significant performance differences. Analysis based on extensive numerical experiments shows that changes are significant in functions with low to moderate multimodality but are less significant in functions with highly multimodality.

Keywords. Evolutionary algorithms, Metamodels, RBF, Expensive functions

1. Introduction

Simulation-driven design optimization which is common in engineering and science defines presents unique challenges such as non-analytic ‘black-box’ functions, computationally expensive evaluations, and complicated function features [1]. A framework which has been extensively used for the solution of such problems is the metamodel (or surrogate) assisted evolutionary algorithm (EA) where the former is an interpolant trained based on sampled vectors (candidate designs). A common metamodel variant is the radial basis functions (RBF) predictor which is an aggregation of basis functions. The latter rely on a hyperparameter which defines their shape and therefore varying this parameter affects the metamodel prediction accuracy and possibly the overall search effectiveness. The optimal hyperparameter value is typically unknown and hence needs to be estimated. This raises the question if the required computational overhead is justified, namely, does it yield significant performance gains. Previous studies have examined the impact of hyperparameter calibration solely on the metamodel prediction accuracy [2,3]. However, the overall search effectiveness is affected in an intricate manner in which metamodel accuracy is just one component [4]. Accordingly this study takes a different approach and examines the impact of hyperparameter variation on the overall search effectiveness. The remainder of this paper is organized as follows: Section 2 surveys relevant studies, Section 3 describes the algorithm and numerical experiments, Section 4 presents the results and analysis, and lastly Section 5 concludes this paper.

¹Corresponding Author: Mechanical Engineering Department, Ariel University, Israel, E-mail: y.tenne@ariel.ac.il

2. Previous Studies

Metamodel-assisted EAs have been used extensively in problems ranging from drug formulation to aircraft design [1]. Extensions ensembles of multiple metamodels [5] and [6] for incorporation of machine learning techniques.

The issue of the RBF hyperparameter calibration has been explored in multiple studies. Early examples include [7] which used a heuristic fixed value and [8] which used RMS error minimization. Related studies also used calibration based on error minimization but with a leave-one-out estimation approach [9,10,11]. Calibration based on a training-testing approach was explored by [2,12]. The studies mentioned have gauged the parameter calibration based solely on the metamodel prediction accuracy.

3. Experiments Layout

To study the impact of the hyperparameter numerical experiments were performed with an RBF metamodel-assisted EA with different hyperparameter settings. A representative EA was used which begins by sampling a random set of vectors and it then repeatedly: i) ranks the vectors by fitness, ii) selects parents by stochastic universal selection (SUS) and recombines them with a uniform crossover operator (probability of 0.7) to produce offspring, and iii) mutates the offspring (Breeder genetic algorithm operator, probability of 0.1) [13]. The population size was 20 and the generations limit was 1000.

Also a standard RBF metamodel was used which is defined as

$$m(\vec{x}) = \sum_{i=1}^n \alpha_i \phi_i(\vec{x}), \quad \phi_i(\vec{x}) = \exp(-\kappa \|\vec{x} - \vec{x}_i\|_2) \quad (1)$$

where $\phi_i(\vec{x})$ are Gaussian basis functions, κ is the hyperparameter, and $\vec{x}_i, i = 1 \dots n$ are the sampled vectors. The coefficients α_i are obtained from the Lagrange interpolation conditions $m(\vec{x}_i) = f(\vec{x}_i)$, namely

$$\begin{pmatrix} \phi_1(\vec{x}_1) & \phi_2(\vec{x}_1) & \dots & \phi_n(\vec{x}_1) \\ \vdots & & & \\ \phi_1(\vec{x}_n) & \phi_2(\vec{x}_n) & \dots & \phi_n(\vec{x}_n) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} f(\vec{x}_1) \\ f(\vec{x}_2) \\ \vdots \\ f(\vec{x}_n) \end{pmatrix} \quad (2)$$

The complete RBF–EA algorithm operates as follows: it begins by sampling an initial random set of vectors and uses them to train an RBF metamodel. Next the main loop begins where the EA uses the metamodel predictions. Every g generations the best s vectors in the EA population are evaluated with the true objective function. The metamodel is then retrained and the EA population is re-evaluated with the new metamodel. The process stops when the limit of true (expensive) function evaluations or maximum EA generations has been reached. Following [14] the settings $g = 4$ and $s = 20\%$ were used. To complete this description Algorithm 1 gives the full pseudocode.

```

sample an initial random population and evaluate;
repeat
  train an RBF metamodel based on sampled vectors;
  run one generation of the EA;
  foreach g generation do
    evaluate the best %s vectors with the true function;
    retrain the metamodel and re-evaluate the population;
  end
until max true (expensive) evaluations or max generations;

```

Algorithm 1: The implemented RBF-EA algorithm.

4. Experiments

The above algorithm was applied to an established set of test functions (Ackley, Griewank, Powell, Rastrigin, and Rosenbrock) in dimensions 5 and 10 [15]. Four hyperparameter settings were tested: 0.01, 0.1, 1, and 10, and for each function–hyperparameter combination 30 runs were repeated yielding a total of $5 \times 4 \times 30 = 600$ runs.

Tables 1 and 2 show the resultant statistics (mean, median, standard deviation, min, max) of the best true function value found. The Mann-Whitney *p*-value is also included where the null hypothesis examined was $P(b > r) > 0.5$ and *b* are the results of the variant with the best mean and *r* are the results of one of the other variants in turn. Statistical significance is accepted for $p < 0.05$. From analysis of the results it follows that:

- Changing the hyperparameter yielded statistically significant performance differences for the Griewank, Powell, and Rosenbrock functions both in the 5D and 10D cases.
- Changes were not statistically significant for the Ackley and Rastrigin functions both in the 5D and 10D cases.
- The $\kappa = 0.1$ setting was the best performing in 6 cases, followed by $\kappa = 0.01$ and $\kappa = 1$ settings in 2 cases each.

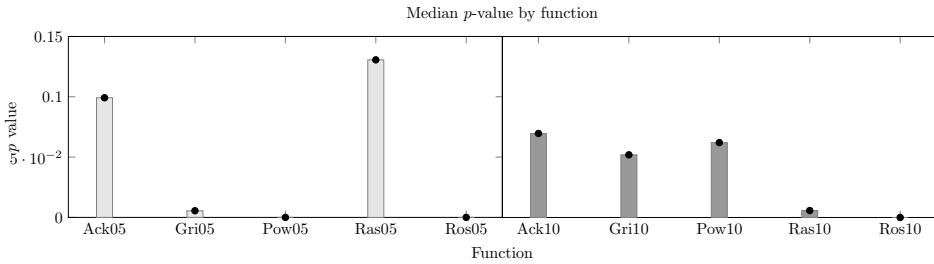
It follows that changes to the hyperparameter were had a significant impact with low–medium multimodality functions. This is attributed to their simpler shape which allowed, given a suitable hyperparameter, for an adequate approximation. In contrast, high multimodality functions (Ackley, Rastrigin) hyperparameter changes had a markedly smaller impact on performance. This is attributed to their complex shape and consequentially poor metamodel accuracy which is less sensitive to different hyperparameter values. Figure 1 shows the median *p*-value by function (out of the three Mann-Whitney tests) and the division between the low and high multimodality functions is evident. Also, moderate hyperparameter values ($\kappa = 0.1$) typically yielded the best results since this avoided generating basis functions which are too narrow or too flat and hence cannot adequately approximate the test functions. Lastly, in the 10D cases the metamodel accuracy also degraded due to the higher dimension, namely it had a similar impact as multimodality.

To augment the analysis additional tests were performed with an application of airfoil shape optimization as a representative real-world simulation-driven problem. The goal is to find an airfoil which maximizes its lift to drag ratio

Table 1. Statistics for the 5D functions

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-05	Mean	1.675e + 01	1.813e + 01	1.815e + 01	1.728e + 01
	SD	3.202e + 00	1.802e + 00	1.544e + 00	2.135e + 00
	Median	1.768e + 01	1.838e + 01	1.845e + 01	1.785e + 01
	Min	7.716e + 00	1.300e + 01	1.304e + 01	1.109e + 01
	Max	2.066e + 01	2.027e + 01	2.005e + 01	2.015e + 01
Griewank-05	<i>p</i> value		9.918e - 02	8.689e - 02	4.238e - 01
	Mean	3.916e + 01	3.084e + 01	2.480e + 01	3.267e + 01
	SD	1.884e + 01	1.522e + 01	1.353e + 01	1.486e + 01
	Median	3.961e + 01	2.750e + 01	2.167e + 01	2.907e + 01
	Min	7.940e + 00	8.968e + 00	5.288e + 00	9.673e + 00
Powell-05	Max	8.665e + 01	7.769e + 01	6.050e + 01	7.108e + 01
	<i>p</i> value	1.281e - 03	3.926e - 02		5.497e - 03
	Mean	2.885e + 01	1.540e + 01	1.364e + 02	1.227e + 02
	SD	2.523e + 01	1.357e + 01	1.225e + 02	1.211e + 02
	Median	2.116e + 01	8.740e + 00	9.514e + 01	7.374e + 01
Rastrigin-05	Min	1.129e + 00	4.463e - 01	1.281e + 00	2.206e + 01
	Max	1.073e + 02	4.944e + 01	5.200e + 02	5.196e + 02
	<i>p</i> value	1.279e - 02		1.607e - 08	2.658e - 10
	Mean	4.263e + 01	3.902e + 01	4.299e + 01	4.212e + 01
	SD	1.136e + 01	1.103e + 01	1.311e + 01	9.976e + 00
Rosenbrock-05	Median	4.243e + 01	4.025e + 01	4.409e + 01	4.073e + 01
	Min	1.966e + 01	1.706e + 01	7.767e + 00	2.789e + 01
	Max	6.739e + 01	5.802e + 01	6.808e + 01	6.398e + 01
	<i>p</i> value	1.306e - 01		1.044e - 01	1.797e - 01
	Mean	3.900e + 01	9.920e + 00	9.758e + 01	1.902e + 02
	SD	5.200e + 01	1.293e + 01	9.729e + 01	7.089e + 01
	Median	1.857e + 01	4.961e + 00	8.194e + 01	1.964e + 02
	Min	3.040e + 00	3.772e - 01	3.574e + 00	7.153e + 01
	Max	2.396e + 02	5.355e + 01	3.977e + 02	3.336e + 02
	<i>p</i> value	2.108e - 05		7.188e - 07	1.436e - 11

The best mean is emphasized.

**Figure 1.** The median p -values of comparisons for each function.

$$f = -\frac{c_L}{c_D} \quad (3)$$

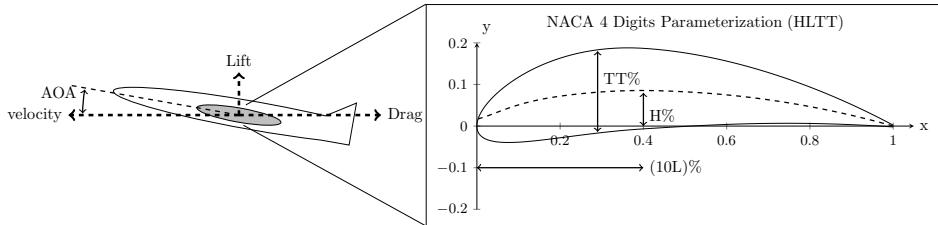
where the lift (c_L) and drag (c_D) coefficients were obtained from the Xfoil numerical aerodynamic simulation code [16]. Airfoils were represented with the NACA 4 digit parameterization and standard flight conditions were used (Mach number=0.7, Reynolds number= 10^7 , angle of attack= 2°) and Figure 2 shows the problem layout and all other algorithm parameters were as before.

Table 3 shows the resultant statistics from which it follows that changes to the hyperparameter were not statistically significant (as with some of the test functions in the pre-

Table 2. Statistics for the 10D functions

Function	Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Ackley-10	Mean	1.973e+01	2.011e+01	1.972e+01	1.983e+01
	SD	1.217e+00	4.317e-01	5.775e-01	4.632e-01
	Median	2.004e+01	2.016e+01	1.965e+01	1.985e+01
	Min	1.498e+01	1.923e+01	1.800e+01	1.824e+01
	Max	2.095e+01	2.083e+01	2.055e+01	2.059e+01
	<i>p</i> value	6.964e-02	3.118e-03		2.577e-01
Griewank-10	Mean	1.482e+02	1.129e+02	1.156e+02	1.244e+02
	SD	3.170e+01	3.293e+01	3.056e+01	3.174e+01
	Median	1.529e+02	1.046e+02	1.093e+02	1.313e+02
	Min	8.037e+01	5.000e+01	6.540e+01	5.481e+01
	Max	1.979e+02	1.939e+02	2.057e+02	1.710e+02
	<i>p</i> value	6.429e-05		3.894e-01	5.194e-02
Powell-10	Mean	6.392e+01	6.374e+01	1.482e+02	8.858e+01
	SD	5.593e+01	6.411e+01	9.066e+01	7.040e+01
	Median	5.303e+01	4.660e+01	1.514e+02	6.497e+01
	Min	8.446e+00	7.498e+00	1.346e+01	5.057e+00
	Max	2.532e+02	3.408e+02	3.492e+02	2.581e+02
	<i>p</i> value	4.941e-01		2.717e-05	6.208e-02
Rastrigin-10	Mean	1.057e+02	1.190e+02	1.232e+02	1.093e+02
	SD	1.925e+01	1.618e+01	1.582e+01	1.734e+01
	Median	1.083e+02	1.184e+02	1.240e+02	1.091e+02
	Min	6.402e+01	8.002e+01	7.702e+01	6.161e+01
	Max	1.310e+02	1.564e+02	1.532e+02	1.441e+02
	<i>p</i> value	5.733e-03		1.461e-04	4.008e-01
Rosenbrock-10	Mean	1.427e+02	1.162e+02	1.231e+03	1.107e+03
	SD	1.132e+02	6.780e+01	4.217e+02	3.513e+02
	Median	1.263e+02	1.037e+02	1.227e+03	1.019e+03
	Min	3.131e+01	2.922e+01	2.908e+02	4.837e+02
	Max	6.257e+02	3.315e+02	2.141e+03	1.954e+03
	<i>p</i> value	1.997e-01		1.588e-11	1.436e-11

The best mean is emphasized.

**Figure 2.** The airfoil problem with the NACA 4 digit parameterization.

vious section) which indicates that the objective function has a complicated landscape. This was confirmed during the tests by the fact that the simulation failed to evaluate numerous candidate airfoils, namely, the function is discontinuous and possibly highly multimodal. Results also show that the best performance was observed with $\kappa = 0.01$ which is consistent with the earlier results.

5. Conclusion

The RBF metamodel-assisted EA has been widely used in the literature with simulation-driven optimization problems. This metamodel relies on a hyperparameter which affects

Table 3. Statistics for the airfoil problem

Statistic	$\kappa = 0.01$	$\kappa = 0.1$	$\kappa = 1$	$\kappa = 10$
Mean	-1.896e + 02	-1.811e + 02	-1.728e + 02	-1.682e + 02
SD	5.471e + 01	5.700e + 01	6.368e + 01	5.712e + 01
Median	-2.084e + 02	-1.931e + 02	-1.939e + 02	-1.601e + 02
Min	-2.586e + 02	-2.564e + 02	-2.662e + 02	-2.564e + 02
Max	-7.424e + 01	-5.714e + 01	-1.436e + 01	-8.499e + 01
<i>p</i> value		1.976e - 01	1.855e - 01	8.457e - 02

its prediction. This study examined if changes to the hyperparameter affected the overall search effectiveness and therefore if calibration is justified. Results from an extensive set of numerical experiments show that hyperparameter changes had a significant impact with low–medium multimodality functions and lesser impact in high multimodality functions. Moderate hyperparameter values (around 1) typically performed best.

References

- [1] Tenne Y, Goh CK, editors. Computational Intelligence in Expensive Optimization Problems. vol. 2 of Evolutionary Learning and Optimization. Berlin: Springer; 2010.
- [2] Durantin C, Rouxel J, Désidéri JA, Gliére A. Multifidelity surrogate modeling based on radial basis functions. *Structural and Multidisciplinary Optimization*. 2017;56:1061–1075.
- [3] Bengisen PG. A New Regression-Based Approach to Estimate the Shape Parameter of MQ-RBFs in a Free Convection Problem. *Journal of Computing and Information Science in Engineering*. 2019;20(1):011009.
- [4] Ong YS, Zhou Z, Lim D. Curse and Blessing of Uncertainty in Evolutionary Algorithm Using Approximation. In: Proceedings of the IEEE Congress on Evolutionary Computation (CEC). Piscataway, New Jersey: IEEE; 2006. p. 2928–2935.
- [5] Tenne Y. An Adaptive-topology Ensemble Algorithm for Engineering Optimization Problems. *Optimization and Engineering*. 2015;16(2):303–334.
- [6] Tenne Y. Frameworks for simulation-driven optimization. *Journal of Applied Mathematics and Computer Science*. 2017;27(1):105–118.
- [7] Hardy RL. Theory and Applications of the Multiquadric-Biharmonic Method. *Computers and Mathematics with Applications*. 1990;19(8/9):163–208.
- [8] Kansa EJ, Carlson RE. Improved Accuracy of Multiquadric Interpolation Using Variable Shape Parameters. *Computers and Mathematics with Applications*. 1992;24(12):99–120.
- [9] Rippa S. An Algorithm for Selecting a Good Value for the Parameter c in Radial Basis Function Interpolation. *Advances in Computational Mathematics*. 1999;11(2–3):193–210.
- [10] Hickernell FJ, Hon YC. Radial Basis Functions Approximation of the Surface Wind Field from Scattered Data. *International Journal of Applied Science and Computation*. 1998;4(3):221–247.
- [11] Sóbester A, Leary SJ, Keane AJ. On the Design of Optimization Strategies Based on Global Response Surface Approximation Models. *Journal of Global Optimization*. 2005;33:31–59.
- [12] Marchetti F. The extension of Rippa's algorithm beyond LOOCV. *Applied Mathematics Letters*. 2021;120:107262.
- [13] Chipperfield A, Fleming P, Pohlheim H, Fonseca C. Genetic Algorithm TOOLBOX For Use with MATLAB, Version 1.2. Sheffield; 1994.
- [14] Ratle A. Kriging as a surrogate fitness landscape in evolutionary optimization. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*. 2001;15(1):37–49.
- [15] Suganthan PN, Hansen N, Liang JJ, Deb K, Chen YP, Auger A, et al. Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization; 2005. KanGAL 2005005.
- [16] Drela M, Youngren H. XFOIL 6.9 User Primer. Cambridge, MA: Massachusetts Institute of Technology; 2001.

Online Life Forecasting Method and State Assessment Architecture of Power Cable Based on Recurrent Neural Network and Internet of Things

Changcheng SONG^{a,1}, Binglei XUE^a, Fumu LU^a, Feng LAN^a, Zheng YANG^b, Chunlei ZHANG^c, Peiran YUAN^d

^a State Grid Shandong Electric Power Company Economic & Technology Research Institute, Jinan 250021, China

^b State Grid Jibei Comprehensive Energy Services Co. Ltd, Beijing 100045, China

^c State Grid Jibei Electric Power Co. Ltd Maintenance Branch, Beijing 102487, China

^d State Grid Jibei Electric Power Co. Ltd Langfang Power Supply Company, Langfang 065000, China

Abstract. For the purpose of increasing the accuracy of power cable life forecasting and status assessment, improving its life cycle management process, this paper proposes a power cable online life forecasting method and status assessment system based on recurrent neural network and Internet of Things (IoTs). Power cable electrical insulation online monitoring system is established on the first place. Then, recurrent neural network and fuzzy analytical hierarchy process are used in the IoTs based power cable online status assessment architecture to proceed life forecasting and status assessment process. Lastly, example analysis is presented to verify the effectiveness and superiority of the methodology introduced in this paper. It is shown that artificial intelligence and IoTs will also have broad development and application prospect when combined with power cable life cycle management.

Keywords. power cable, life cycle management, recurrent neural network, Internet of Things (IoTs), life forecasting, status assessment

1. Introduction

Power cable is a cable product used to transmit and distribute high-power electrical energy in power systems. Its practical application has a history of more than 100 years. Compared with overhead transmission lines, cables have the advantages of less land occupation, high reliability, large distributed capacitance, less maintenance workload, and low probability of electric shock. Moreover, with the development trend of power systems, cables have the advantages of more favorable conditions for the development of extra high voltage, large capacity and long distance. Therefore, the proportion of cables in the total number of transmission lines is gradually increasing during the development of the power system. XLRE cables have become the first choice for

¹Corresponding Author: Changcheng Song, State Grid Shandong Electric Power Company Economic & Technology Research Institute, Jinan 250021, China; E-mail: 18366118336@163.com.

power cables due to their small quality, large transmission capacity, convenient operation and maintenance, and good electrical and heat resistance properties, since it was researched and developed in 1957 [1].

However, the quality of power cables is affected by the production process, laying methods, operating conditions, and time of use, and the quality is uneven. In addition, the cable production industry itself has serious problems of repeated production, repeated construction, and overcapacity. It will bring great hidden dangers to its safe and reliable operation and bring great inconvenience to residents' lives and economic losses caused by power outages, if the cable is not properly monitored, maintained and managed. Therefore, the full life cycle management of power cables from design and development, production, use and decommission and material recovery can comprehensively consider initial investment and operation and maintenance costs, and improve equipment reliability and efficiency [2]. Life forecasting and condition assessment play a pivotal role in the life cycle management of power cables. It is necessary for electric power companies to conduct real-time condition monitoring of power cables in operation to assess the reliability of their operating conditions, forecast their remaining service life and prevent its possible failures or abnormal conditions, which can achieve the purpose of reducing production costs and improving benefits.

At present, there are few researches on power cable life forecasting and condition assessment around the world. Normal distribution and Weibull distribution is used in reference [3-5] to describe commonly used equipment life and reliability probability distributions. Reference [6] and [7] assessed the operating state of power transformers, and its theories and methods have reference value for the state evaluation of power cables. A differential scanning calorimetry system is used in reference [8] to measure the relationship between the heat flow rate of the sample and the reference and the temperature with the temperature change, thereby obtaining the thermal aging equation, and then forecasting the cable life. Bayesian method used in reference [9] to forecast the insulation reliability of cables. The statistical data of cable insulation aging is used in reference [10] to establish a cable reliability life calculation model that takes into account the failure rate. A status evaluation model is established by setting three levels of cable status in reference [11] based on lognormal distribution, and a corresponding maintenance plan based on the evaluation results is proposed. However, the above-mentioned work is mainly based on a specific index. Moreover, for power cables in actual operation of the project, it is impossible to forecast and comprehensively assess their life and status in real time with the above-mentioned methods, or the results of forecasting and evaluation are difficult to reflect the coupling with time. In addition, most of the current methods for cable life forecasting and state evaluation are to assess and calculate the cable in the laboratory after the cable is withdrawn from operation, which is of low practical value. With the continuous increase of data volume and the development of real-time measurement and cloud computing technology, the advantages of artificial intelligence methods and Internet of Things technology have gradually emerged. Artificial intelligence and Internet of Things technology have a good development and application prospects to be applied to the life cycle management of power cables.

An online forecasting method of power cable insulation aging life based on recurrent neural network and a power cable state assessment architecture based on the Internet of Things are proposed in this paper. A power cable online state evaluation architecture based on the Internet of Things is established by building the power cable online state evaluation index system in Section 2. The online forecasting method of

power cable insulation aging life based on recurrent neural network is put forward in Section 3, including the online forecasting method of power cable insulation aging index based on recurrent neural network and the calculation method of power cable insulation aging life based on fuzzy analytic hierarchy process. The effectiveness of the method is analyzed and proved in Section 4. The conclusion is drawn in Section 5.

2. Evaluation Index System and Architecture of Power Cable Online Status

2.1. Power Cable Online Status Evaluation Index System

It is generally confirmed that the online index system describing the cable condition assessment mainly determines the cable life by affecting the insulation aging process of the cable. As shown in Figure 1, the online index system mainly includes the following four indices:

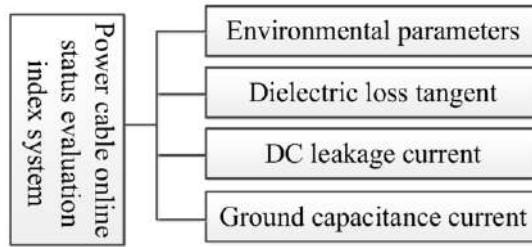


Figure 1. Power cable online status evaluation index system.

1) Environmental parameters E : It mainly includes average surface temperature, average soil moisture, soil pH, road pressure and soil sand quality, which are important factors affecting dielectric insulation.

2) Dielectric loss tangent $\tan \delta$: It is a characteristic parameter reflecting the dielectric loss of the cable insulation and an important index for evaluating the insulation performance. The $\tan \delta$ will gradually increase with the gradual aging of the cable. The late aging degree will increase significantly, and the $\tan \delta$ will increase significantly until the cable breakdown occurs.

3) DC leakage current I : There will be a nano-ampere DC leakage current in the insulation layer generated by the water tree of the power cable under the action of AC voltage, which is an important parameter for cable insulation testing. Therefore, the water tree aging of the power cable can be found by detecting the DC leakage current of the cable, which reflects the degree of development of the cable insulation aging.

4) Grounding capacitance current I_c : This is also an important parameter for cable insulation monitoring. In the aging process of power cable, with the increase of capacitance, the grounding capacitance current of the cable increases.

2.2. Power Cable Online Status Assessment Architecture Based on the Internet of Things

In order to assess the operating status of power cables in real time, it is necessary to conduct real-time online monitoring and analysis of the above-mentioned cable operating indices, which is difficult to meet the requirements of online status evaluation because of huge manual measurement workload. To solve this problem, an online state assessment architecture is proposed in this paper, based on the Internet of

Things technology, as shown in Figure 2. The data collected by the perception layer is uploaded to the data layer through the communication network for data aggregation, integration and analysis, and application services such as management decisions are provided to users based on the results of data analysis and state evaluation.

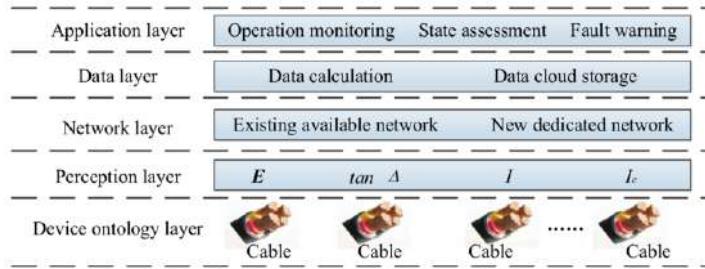


Figure 2. Power cable online state assessment system architecture.

2.2.1. Perception Layer

The perception layer is composed of various sensors to collect data based on actual cable equipment operation, including environmental parameters E , dielectric loss tangent $\tan \delta$, DC leakage current I , grounding capacitance current I_c and other data in the index system. The collection terminal supports data collection methods such as periodic, fixed time, current time, etc. The collected data is uploaded to the cloud service platform through the network for data analysis and processing, and the requirements for fast and accurate cable online status evaluation are realized.

2.2.2. Network Layer

The network layer uploads the monitoring information data packets collected by the collection terminal through the transmission network to the cloud service platform master station to realize data interaction between on-site collection and the data center. The communication network can directly use the existing network, or build a dedicated network according to the actual situation on site.

2.2.3. Data Layer

The data layer is based on the cloud service platform, which integrates and analyzes the data uploaded by each cable device, and obtains the evaluation and analysis results of the operation status of the equipment. Because cloud computing has the characteristics of high flexibility, strong reliability, strong computing power, and strong scalability [12,13], data processing speed and efficiency will be greatly improved; it also depends on the development of cloud storage technology [14], massive monitoring data is stored in the virtual server of the cloud service platform, greatly reducing the cost of data storage.

2.2.4. Application Layer

The application layer connects the analyzed and processed data to each application module to realize the user's functional requirements for cable operation monitoring, status evaluation, fault warning, and life forecasting. Thanks to the open architecture of the Internet of Things and the characteristics of integration and decoupling [15], users

can freely combine and customize background applications according to actual needs, expand their functional applications, and adapt to diverse needs.

3. Online Forecasting Method of Power Cable Insulation Aging Life Based on Recurrent Neural Network

3.1. Online Forecasting Method of Power Cable Insulation Aging Index Based on Recurrent Neural Network

Neural network is a complex network system that simulates the behavior of the human brain, which is composed of a large number of simple neurons widely connected and layered, as shown in Figure 3. Neural networks are often used to achieve automation, solve complex calculation and forecasting problems, as an important part of artificial intelligence technology [16]. Currently commonly used neural networks mainly include BP neural network, Hopfield network and so on.

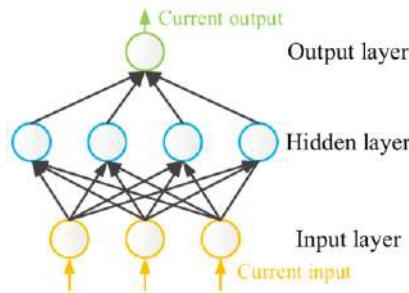


Figure 3. Common neural network structure.

In the above-mentioned more traditional neural network structure, the neurons between the layers are generally fully connected, and the neurons in each layer are not connected. The output vector at the current moment is only related to the input vector at the current moment. However, when this structure is processing time series data (such as real-time data of power cable operating status), the calculation and forecasting results cannot meet the requirement because it cannot reflect the connection and change law of data and time.

Aiming at the shortcomings of common neural networks, a power cable life forecasting method is proposed in this paper, based on recurrent neural networks (RNN). Different from common neural networks, there are connections between neurons in the hidden layers of recurrent neural networks, as shown in Figure 4. The output at the current moment depends not only on the input at the current moment, but also on the system state at the previous moment. In other words, the recurrent neural network will memorize the past system state and use it to calculate the output vector at the current moment, which can better deal with time series problems.

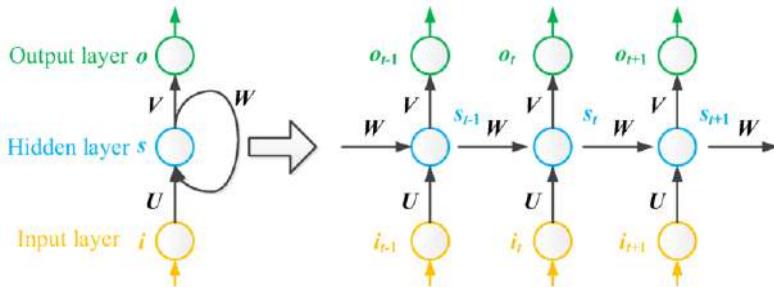


Figure 4. Recurrent neural network structure.

The calculation method of recurrent neural network is:

$$s_t = f(Ug_i + Ws_{t-1} + b) \quad (1)$$

In the formula, f is the activation function of the neural network, and b is the bias vector. When $t=0$, $s_{-1}=0$ by default. The calculation method of the output vector o_t at time t is

$$o_t = \text{softmax}(Vs_t + c) \quad (2)$$

In the formula, c is the bias vector. The parameter vectors U , V , and W of each layer in each time interval are unchanged, which means that the work done by the recurrent neural network at different times is unchanged, which also significantly reduces the number of parameters that need to be learned. The output of the recurrent neural network is related to the input at all previous moments, but the actual calculation takes into account that the connection between the two moments when the interval is too long tends to decrease, so the neural network is generally cut to have a maximum length. In other words, the output at a certain moment is related to all the inputs in a specific period of time before.

For a given power cable, I and I_c can be calculated by installing a current sensor on the cable grounding wire, measuring the cable grounding wire current signal, and performing analog-to-digital conversion and discrete Fourier transform. The dielectric loss tangent $\tan \delta$ can be obtained by setting a voltage sensor to obtain the operating voltage of the power cable. After filtering and amplification, analog-to-digital conversion and fast Fourier transform, the fundamental wave of the two signals can be obtained, and the phase difference of the two fundamental waves can be further calculated to obtain the dielectric loss tangent $\tan \delta$. Environmental parameters can be directly measured by installing temperature, humidity, pressure and other sensors around the cable. However, considering that environmental factors also affect other indices to determine the remaining aging life of the cable, they are not considered in the application of the method.

For indices such as $\tan \delta$, I , and I_c , the real-time measurement and calculation results are used to form time series data, and the time interval between two adjacent moments is Δt . According to the currently measured data, the recurrent neural network is used to forecast the future index change data. The cut-off length of the neural network is set to l , which means that the index data at a certain moment is strongly

correlated with the index data at the previous l moments. For example, the learning process of $\tan \delta$ of the recurrent neural network is shown in the following formula:

$$\mathbf{s}_t = f(\mathbf{U} \mathbf{g}_{\tan \delta, t} + \mathbf{W} \mathbf{g}_{t-1} + \mathbf{b}) \quad (3)$$

$$\mathbf{o}_{\tan \delta, t} = \text{softmax}(\mathbf{V} \mathbf{s}_t + \mathbf{c}) \quad (4)$$

In the formula, $\mathbf{i}_{\tan \delta, t} = [\tan \delta_{t-l}, \tan \delta_{t-l+1}, \dots, \tan \delta_{t-1}]^T$ is the input vector of the dielectric loss tangent at time t , $\mathbf{o}_{\tan \delta, t} = \tan \delta_t$ is the output vector of the dielectric loss tangent.

3.2. Calculation Method of Power Cable Insulation Aging Life Based on Fuzzy Analytic Hierarchy Process

Fuzzy analytic hierarchy process is often used to solve the problem of weight distribution among multiple evaluation indices [17], and its steps are as follows:

1) The final decision-making goal and each influencing factor are layered according to the relationship to obtain a hierarchical structure diagram.

2) Construct the fuzzy judgment matrix $A = (a_{ij})_{n \times n}$, the element a_{ij} in the matrix represents the importance of index i relative to index j , which can be obtained by evaluating the index importance according to the 0.1~0.9 scale method [17], and satisfy $a_{ij} + a_{ji} = 1 (i, j = 1, 2, \dots, n)$, and the scaling rules are shown in Table 1.

Table 1. 0.1~0.9 scaling rule.

Scale result	Scale value
Both indices are equally important.	$a_{ij} = a_{ji} = 0.5$
$i(j)$ is slightly more important than $j(i)$.	$a_{ij}(a_{ji}) = 0.6 \quad a_{ji}(a_{ij}) = 0.4$
$i(j)$ is more important than $j(i)$.	$a_{ij}(a_{ji}) = 0.7 \quad a_{ji}(a_{ij}) = 0.3$
$i(j)$ is much more important than $j(i)$.	$a_{ij}(a_{ji}) = 0.8 \quad a_{ji}(a_{ij}) = 0.2$
$i(j)$ is extremely important than $j(i)$.	$a_{ij}(a_{ji}) = 0.9 \quad a_{ji}(a_{ij}) = 0.1$

3) In order to meet the requirements of consistency of thinking, the fuzzy judgment matrix is transformed by the following formula to obtain the consistency matrix B :

$$b_i = \sum_{m=1}^n a_{im}, \quad i \in \{1, 2, \dots, n\} \quad (5)$$

$$b_{ij} = (b_i - b_j) / [2(n-1)] + 0.5 \quad (6)$$

4) The weight k_i of the evaluation index i is calculated according to the following formula:

$$k_i = \frac{1}{n} - \frac{1}{n-1} + \frac{2}{n(n-1)} \sum_{m=1}^n b_{im} \quad i \in \{1, 2, \dots, n\} \quad (7)$$

In this research, the measured index data in section 2.1 should be normalized at first to obtain the relative aging degree of the cable represented by various indexes, as shown in the following formula:

$$u_i = \frac{|U_i - U_{if}|}{|U_{io} - U_{if}|} \quad (8)$$

In the formula, U_i represents the measured value of index i , U_{if} represents the qualified threshold of index i , U_{io} represents the initial value of index i , and u_i is the normalized index value that characterizes the relative aging of the cable. The larger the value of u_i , the lighter the aging of the power cable. Finally, the calculation method of the power cable life index is shown in the following formula:

$$y = \sum_{i=1}^n k_i u_i, \quad i \in \{1, 2, \dots, n\} \quad (9)$$

In the formula, the larger the value of y , the longer the remaining life of the cable. The calculation method of the remaining life of the power cable is shown in the following formula:

$$T_s = \frac{y}{1-y} T_R \quad (10)$$

In the formula, T_s represents the remaining life of the cable, while T_R represents the operating time of the cable.

The remaining aging insulation life of the cable can be obtained online according to the method shown in equations (5) ~ (10) after the various index data are forecasted. The overall architecture of the method is shown in Figure 5. After obtaining the forecasting results of various indices, the root mean square error (RMSE) between the forecasted value of the index and the true value should be calculated according to formula (11) to assess the accuracy of the forecasting.

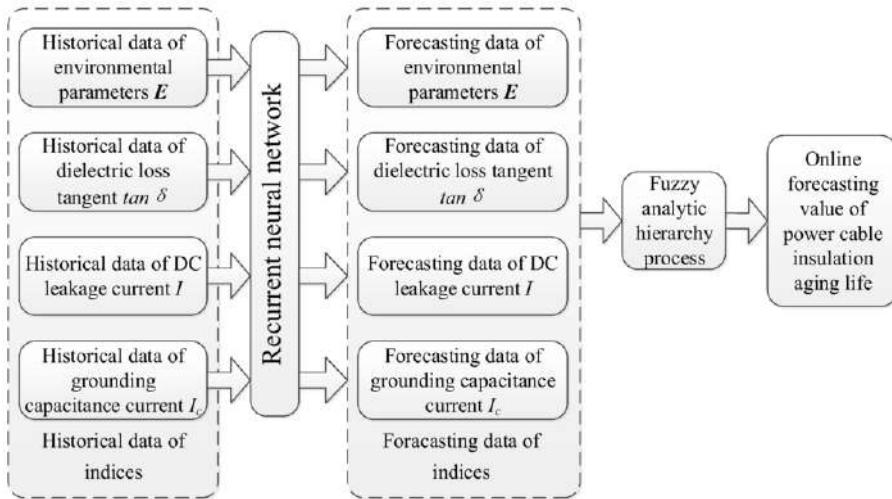


Figure 5. Overall structure of the forecasting method.

$$RMSE(X) = \sqrt{\frac{1}{N} \sum_{t=1}^N [X(t) - \hat{X}(t)]^2} \quad (11)$$

In the formula, X is the set of forecasted values at each time of a certain index, N is the number of forecasted time, $X(t)$ is the forecasted value of index X at the t -th time, and $\hat{X}(t)$ is the true value of the index X at the t -th time.

4. Case study

A 110kV voltage grade XLPE power cable put into operation on Sept. 16, 2007 is used as the test cable in this paper. The neural network toolbox in MATLAB R2014a is used to build a recurrent neural network (RNN) architecture on a computer with a CPU model of Intel(R) G4600@3.60GHz and a RAM model of 8GB. The real-time monitoring data of the power cable between January 1, 2010 and December 31, 2017 is selected for neural network training and learning. The sampling interval is 1 hour, and a total of 70128 sets of data are used.

The cut-off length of the recurrent neural network is set to 48, which means that the index value at a certain moment is related to the index value in the previous 48 hours. The index data between January 1, 2018 and December 31, 2019 is forecasted and the forecasting results are compared with the common BP neural network (BPNN) forecasting results. The online forecasting results of various indices are shown in Figure 6. In order to show the changes of the indices more clearly, the figure only lists the index values at the beginning of each month from January 1, 2018 to December 31, 2019. The comparison of the RMSE between the forecasted value and actual value of each index is shown in Table 2.

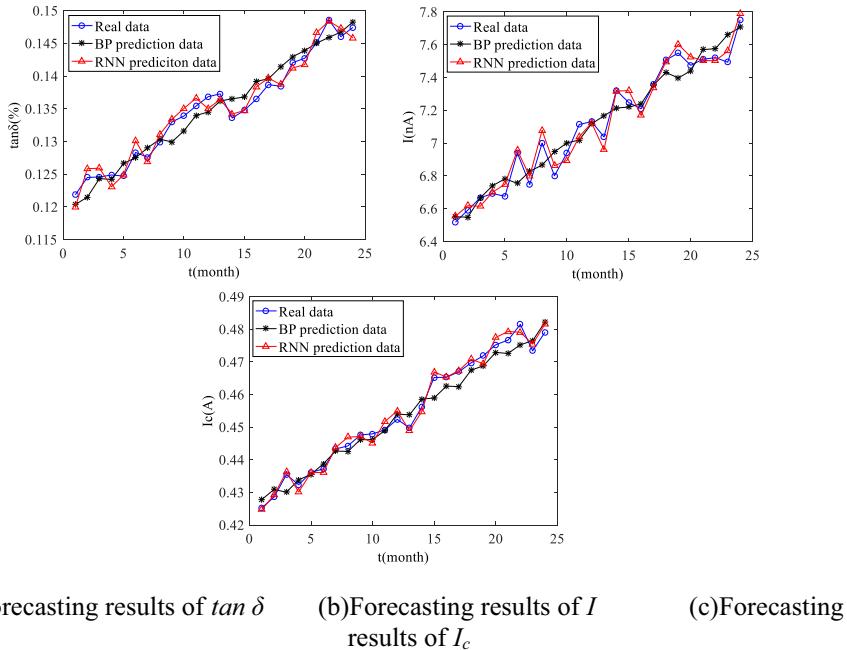


Figure 6. Online forecasting results of insulation aging indices for power cables.

Table 2. Comparison of RMSE of index forecasting results.

	$\tan \delta$	I	I_c
RNN	0.0012	0.048	0.0018
BPNN	0.0018	0.093	0.0032

As Shown in Figure 6 and Table 2, the forecasting method based on RNN proposed in this paper can accurately forecast various insulation aging indices of power cables online, and the accuracy is higher than that of the common BPNN. The effectiveness and superiority of this method are proved.

The method proposed in section 3.2 is used to calculate the remaining life of the power cable. The dielectric loss tangent $\tan \delta$ is the most important index reflecting the degree of power cable aging; the measurement of DC leakage current I is subject to external interference and is less important; the importance of grounding capacitance current I_c is between the above two indices. The fuzzy judgment matrix is constructed:

$$A = \begin{bmatrix} 0.5 & 0.8 & 0.7 \\ 0.2 & 0.5 & 0.7 \\ 0.3 & 0.3 & 0.5 \end{bmatrix}$$

According to formulas (5) and (6), The consistency matrix is calculated as:

$$B = \begin{bmatrix} 0.5 & 0.65 & 0.725 \\ 0.35 & 0.5 & 0.575 \\ 0.275 & 0.425 & 0.5 \end{bmatrix}$$

According to formula (3), the weight matrix of each index is calculated as:

$$K = [0.4583 \ 0.3083 \ 0.2333]^T$$

Finally, the forecasting results of various indices combined, according to formulas (7) ~ (10). The remaining insulation aging life of the power cable is about 11 years and 4 months by the end of 2019.

5. Conclusion

An online forecasting method of power cable insulation aging life based on recurrent neural network and a power cable status assessment architecture based on the Internet of Things are proposed in this paper, aiming at the disadvantages of poor comprehensiveness and low practicability of the existing research on power cable life forecasting and condition assessment. The power cable online state evaluation index system is constructed. Combined with the power cable online state assessment architecture based on the Internet of Things, the method of recurrent neural network is used to forecast the various insulation aging indices of power cables online, and to assess the operating state of the power cable. According to the forecasted results, the fuzzy analytic hierarchy method is used to calculate the remaining insulation aging life of the cable, so as to systematically manage the life cycle of the power cable. Simulation results proves the effectiveness and superiority of the method proposed in this paper. This method and assessment framework apply artificial intelligence and Internet of Things technology to the life forecasting and state assessment of power cables, which have good development and application prospects for the life cycle management of power cables.

References

- [1] A Madariaga, J Martí'n, I Zamora, S Ceballos, O Anaya-Lara. Effective assessment of electric power losses in three-core XLPE cables [J]. IEEE Transactions on Power Systems, 2013, 28(4): 4488-4495.
- [2] S Raghavan, B Chowdhury. State diagram-based life cycle management plans for power plant components [J]. IEEE Transactions on Smart Grid, 2015, 6(2): 965-972.
- [3] A H Etemadi, M Fotuhi-Firuzabad. Design and routine test optimization of modern protection systems with reliability and economic constraints [J]. IEEE Transactions on Power Delivery, 2012, 27(1): 271-278.
- [4] C Zhou, M Michel, D Hcpburn, et al. Online partial discharge monitoring in MV underground cables [J]. IFT Measurement and Science, 2009, 3(5): 353-363.
- [5] H Qiao, P Chris. Estimation of the three parameter Weibull probability distribution [J]. Mathematics and Computers in Simulation, 1995, 39(1-2):173-185.
- [6] S Li, S Liu, Y Huang. Application of the levenber-marquardt method to determine the rainfall parameters [J]. Arch & Tech, 2004, 40(1): 85-90.
- [7] E Pang, G Tang, H Yu. Comparative research on transformer trust and safety index [J]. Power System Protection and Control, 2014, 2(23): 139-145.
- [8] J Crine. Cable life predictions from breakdown and accelerated aging tests: influence of frequency and high fields [C]. 2009 IEEE Electrical Insulation Conference, May 31, Montreal, Canada 2009.
- [9] Chioldo E, Fabiani D. Bayes inference for reliability of HV insulation systems in the presence of switching voltage surges using a weibull stress-strength model [C]. IEEE Bologna PowerTech Conference, June 23-26, Bologna, Italy, 2003.
- [10] Stotzel M, Zdrallek M. Reliability calculation of MV-distribution networks with regard to ageing in XLPE-insulated cables [J]. IEE Proceeding-Generation Transmission and Distribution, 2001, 148 (6): 597-602.
- [11] Q Hong, D Wei, W Yuan, H Yuan, H Guo, Z Wang. Research on cable condition evaluation method based on lognormal distribution [J]. Power System Protection and Control, 2018, 46(2): 79-84 (in Chinese).
- [12] P Porambage, J Okwuibe, M Liyanage, M Ylianttila, T Taleb. Survey on multi-access edge computing for internet of things realization [J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 2961 - 2991.

- [13] Z Zhang, X Zhang. Realization of open cloud computing federation based on mobile agent [C]. 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, Nov. 20-22, Shanghai, China, 2009.
- [14] Q He, Z Li, X Zhang. Analysis of the key technology on cloud storage [C]. 2010 International Conference on Future Information Technology and Management Engineering, Oct. 9-10, Changzhou, China, 2010.
- [15] H Dang, N Krommenacker, P Charpentier, D Kim. Toward the internet of things for physical internet: Perspectives and challenges [J]. IEEE Internet of Things Journal, 2020, 7(6): 4711 - 4736.
- [16] Z Shi. Artificial intelligence [M]. Beijing: China Machine Press, 2016 (in chinese).
- [17] Z Wang. A Representable uninorm-based intuitionistic fuzzy analytic hierarchy process [J]. IEEE Transactions on Fuzzy Systems, 2020, 28(10): 2555-2569.

Virtual Power Plant Platform for Demand Response Based on Microservice Architecture

Guodong JIANG^{a,1}, Zuixi XIE^b, Tao HAN^a, Hongwei DU^a, Rongqiang FENG^a,
Shouquan LUO^c, Xueqiong WU^a, Jiaqi ZHANG^a

^a*NARI TECHNOLOGY CO., LTD, Nanjing 211106, China*

^b*Three Gorges Electric Energy (Guangdong) Co., Ltd, Shenzhen 518000, China*

^c*State Grid Electric Power Research Institute, Nanjing 211106, China*

Abstract. It is a trend to use virtual power plant technology to realize demand response and participate in electricity trading. We design and implement the software control platform of virtual power plant for demand response. For this software platform, we analysed the requirements and got the overall architecture of the platform. On this basis, we design and implement the microservice architecture, interface design, basic application function design, advanced application function design, hardware architecture, communication architecture and security encryption of the platform. Finally, we summarize the application of the platform, and put forward the direction of further research and development.

Keywords. virtual power plant, demand response, microservice, function design, communication architecture, security encryption

1. Introduction

Virtual power plant (VPP) is the aggregation of a variety of distributed energy resources (DER). Through advanced control, communication and metering technology, it can aggregate distributed generation, energy storage, flexible load and other adjustable resources, participate in the dispatching of power grid as a whole. It makes full use of the coordination and complementarity of DER's regulating characteristics to realize the reasonable optimal allocation and utilization of resources. As a special power plant, a virtual power plant presents the overall function and effect of a traditional power plant. It can be controlled and managed like a traditional power plant. It can submit power generation plans to the power grid, participate in the electricity market, implement peak load regulation and frequency regulation and other auxiliary services. As an energy management system, internal energy management of VPP has multiple functions such as self-coordination, self-management and self-control [1].

¹Corresponding author: Guodong Jiang, NARI TECHNOLOGY CO., LTD, Nanjing 211106, China; E-mail: jiangguodong@sgepri.sgcc.com.cn.

Distributed energy resources involved in demand response include distributed generation, air conditioning load, distributed energy storage, electric vehicle charging pile, etc. Many DERs are small in scale, large in number and scattered in dispersive layout. Due to legal and technical reasons, traditional demand response technology cannot use these resources. At the same time, the traditional demand response has low automation level, low accuracy, poor reliability, and the time scale of load regulation response is long, which is difficult to meet the timeliness requirements of power system safe operation [2,3].

The virtual power plant is used to participate in the demand response, so that all kinds of users, especially small and medium capacity users, can participate in the demand response through the virtual power plant to obtain additional economic benefits; For power grid companies, it is beneficial to utilize small-scale flexible load resources, improve economic benefits and reduce construction investment; For the nation and society, the construction of virtual power plant can significantly reduce the combustion of fossil fuels and greenhouse gas emissions, promote the efficient and energy-saving operation of power grid, and effectively promote the optimization and transformation process of power grid [4].

This paper describes a virtual power plant platform for demand response, which is a flexible and extensible system. Demand response business has two characteristics: rapidly changing of functions and rapid altering of customers. The platform adopts microservice architecture, which has the advantages of loose coupling, high cohesion, good interface adaptability, high security and high reliability. So, it can meet the requirements of demand response. This platform can realize the scalability and high availability of the system for customers to join and exit at any time; For the rapid changing of functions, it can realize rapid development and rapid deployment. We analyzed the requirements of the software platform and designed the overall architecture, software architecture and hardware architecture; the software function and system interface are designed; The communication architecture and security architecture are described.

2. Requirement Analysis

Based on the requirements and current situation of Chinese power system, the VPP platform monitors and analyses DERs and traditional power, including traditional power, DG, controllable load, energy storage, etc. After that, the VPP platform sends the calculated instructions to the dispatching system D5000 and the load control system, and these two systems send the instructions to the AGC and AVC modules of the dispatching system after calculation. After its calculation, AGC and AVC modules send instructions to their DERs for execution. The centralized control platform of VPP does not directly send control instructions to DERs. These are shown in Figure 1.

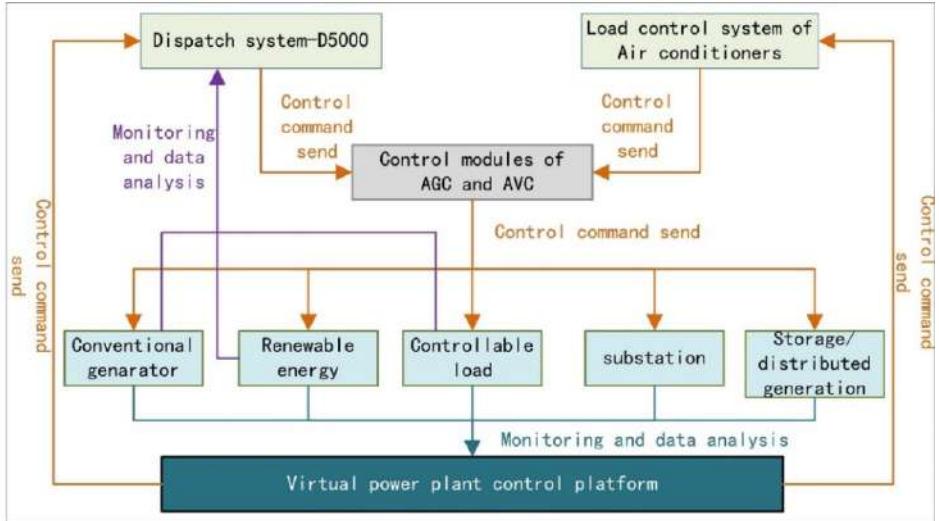


Figure 1. Virtual power plant platform diagram.

The VPP participates in the electricity market of demand response as an independent identity, and the following process is shown in Figure 2. Firstly, the power dispatching center releases the market demand, and then the VPP center interacts with it, including quotation, load control, settlement and so on. After the VPP platform receives the task, it decomposes and sends it to the end users, and the end users execute it. The VPP cooperates with the power grid dispatching center to complete the peak load regulation task, and makes settlement with the trading center.

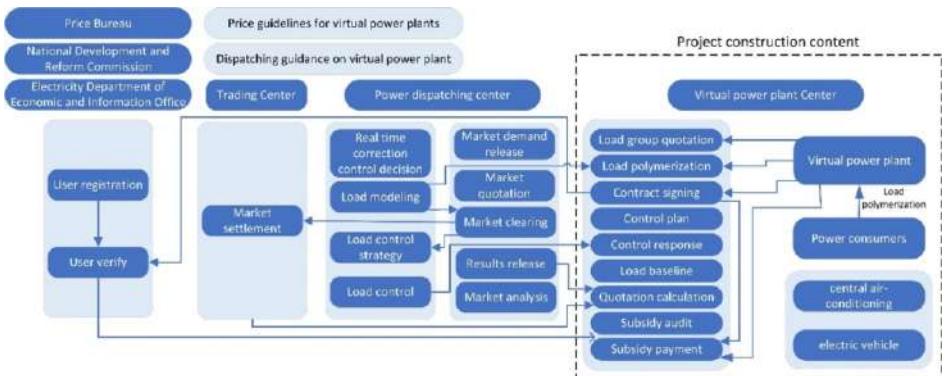


Figure 2. Operation flow chart of virtual power plant.

In order to complete this project, we have done the following three aspects.

One is to aggregate and evaluate the potential of DERs (mainly including air conditioning, small PV, etc.) added to the VPP. The aggregation parameters include the maximum and minimum power generation capacity (operating power constraints, cost and other indicators), the ramp rate of the VPP in each period, and various types of auxiliary service capacity (reserve capacity, power consumption, Cost and performance indicators etc. of the VPP).

The second is to establish a hierarchical, partitioned and multi-level dispatching and operation control architecture of VPP. Based on this framework, the multi-scale spatial-

temporal fluctuation characteristics of DER output are analyzed, and the control strategy of power balance and operation mode optimization is proposed. The VPP collaborative interactive operation control technology is adopted.

The third is to design and implement the overall architecture of the virtual power plant platform, including process design, hardware architecture, software architecture, interface design, etc. We connect the VPP with the dispatch platform D5000 and the load control system, so as to realize the execution of the dispatch platform D5000 instructions and the optimal control of the internal DERs.

3. Virtual Power Plant System

The system diagram of VPP is shown in Figure 3. The operation architecture of VPP is shown in the figure. According to the mode of unified management, distributed access and collaborative control management, based on the power Internet of Things, the VPP system is constructed through intelligent edge computing, perception terminal, data intelligent fusion analysis, auxiliary management decision and other technical methods [5].

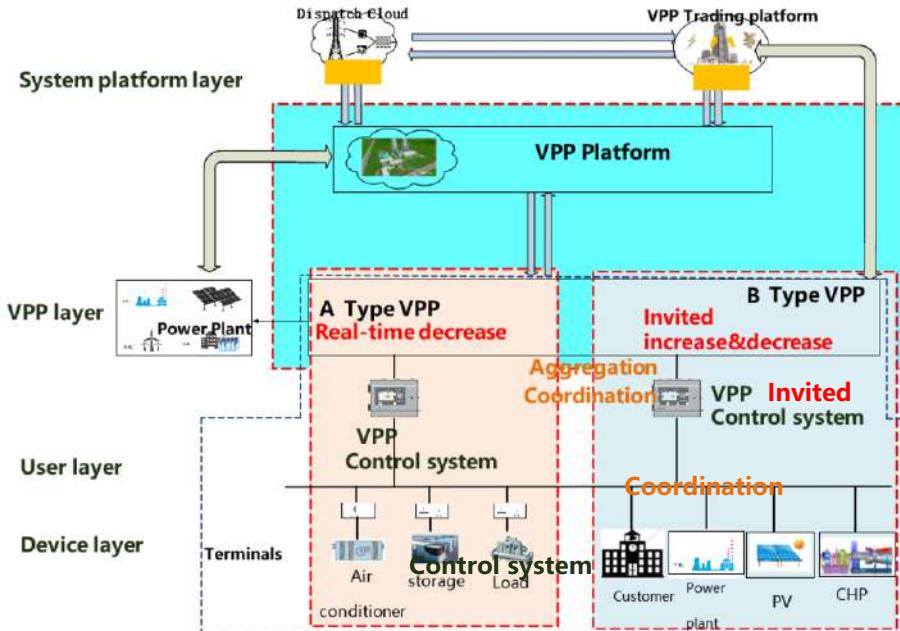


Figure 3. The system diagram of VPP.

The VPP system realizes the partition clustering of DERs and forms the external characteristics similar to traditional power plants. Then, the registration information, quotation information and contract information are submitted to the VPP platform. The VPP platform is responsible for reviewing the VPP access conditions, access scheme, external characteristics of the VPP, VPP files and response information, receiving dispatching cloud call information and reporting response information.

Relying on the existing unified power market trading platform, the VPP trading platform receives the dispatching requirements issued by the dispatching cloud, the access audit results submitted by the VPP platform, the external characteristics of the

VPP, and the implementation identification results. Combined with the registration, quotation, and contract information reported by the VPP platform, the VPP trading platform carries out market clearing, and at the same time reports release clearing results to the VPP's every entity, VPP platform and dispatching cloud.

The VPP system divides DERs into class A and class B. Class A VPP mainly include industrial and commercial residential building air conditioning, user side energy storage and industrial user load. These objects are adjusted and reduced in real time, and participate in the transaction of power grid peak shaving auxiliary services. Class B VPP are mainly large users, self-provided power plants, distributed photovoltaic power plants and triple generation power plants. They are invited to increase or decrease these kinds of objects and participate in the auxiliary service transaction of peak load cutting and valley filling.

VPP system is divided into four layers, namely Device layer, User layer, VPP layer and System platform layer.

Device layer: the VPP resources on the demand side. The project considers the load of building central air conditioning and split air conditioning and so on.

User layer: air conditioning intelligent control terminal is deployed in public buildings, so that the VPP layer can connect with it to control demand side resources.

VPP layer: build virtual unit control system. Monitoring and controlling the operation of VPP resources. It is a centralized control mode with flexible control resources. The VPP resources are actively aggregated, and the end users transfer the control right of resources to them, so that the VPP can control the demand side resources through price or other signals.

System platform layer: build VPP platform. There are several VPP units under the jurisdiction of VPP platform, and there is a decentralized control mode between VPP platform and VPP units. Its instructions include the upload and release of instructions and data monitoring. The VPP platform directly participates in the power market transactions and receives the control cloud response instructions.

4. Software Design

4.1. Microservice Architecture

Microservice architecture decouples solutions by decomposing functions into several services. By splitting a large application and service into several or even dozens of microservices, the microservice architecture implements the extension of a single component rather than the stack of the whole application. In this way, developers can create applications around a single business domain component, and develop, manage and iterate these applications independently, which greatly improves the scalability. In the distributed components, by Using Cloud Architecture and platform deployment, the delivery and management of products become easier [6].

Compared with traditional monolithic architecture, micro service has the following advantages: project can be delivered continuously, service can be expanded independently, compatibility is stronger, technology upgrading and simultaneous interpreting are quicker and more convenient.

Microservice architecture can effectively split applications, realize agile development and deployment, and solve the problems of traditional monolithic architecture such as tight coupling, inflexible deployment and poor scalability. Microservice architecture is a distributed overall solution, including service registration and discovery, service configuration, link monitoring, service gateway, load balancing, service fusing and other functions. This project selects the microservice framework spring cloud, and the architecture diagram is shown in Figure 4.

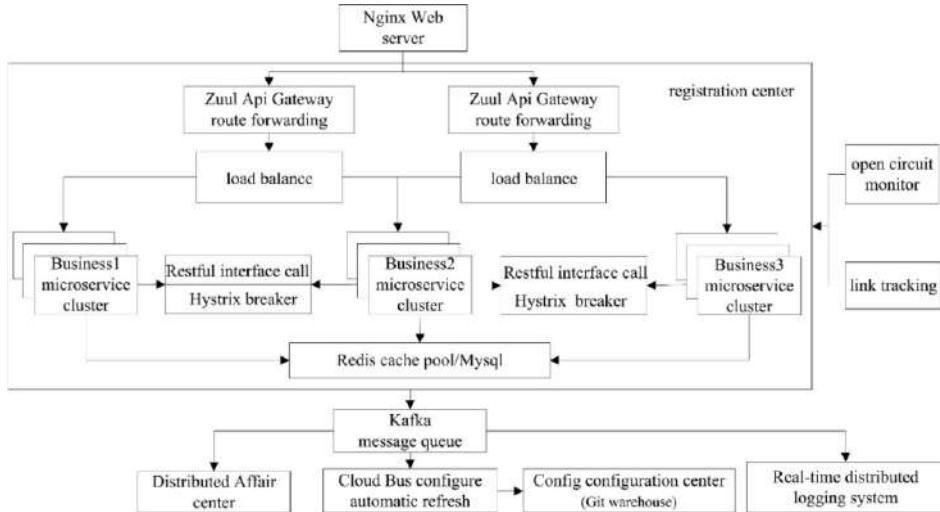


Figure 4. Overall architecture of microservice.

The running process of microservice architecture is: firstly, all microservices are deployed to each server and registered in the registry. When the front end sends an HTTP or HTTPs request to the Nginx server, the Nginx server checks its identity. After the verification is passed, Nginx forwards the request to a micro service in the registry through the router.

Microservices can call each other. When Microservice A calls Microservice B, it needs to obtain the registered address and port of Microservice B from the microservice registry. Since Microservice B may have multiple copies, it needs to select one of the interfaces exposed by the copy service of Microservice B to call through the load balancing polling mechanism. The breaker mechanism is introduced in the process of microservice call to avoid the microservice call exception caused by a microservice failure, which leads to the whole microservice crash. In this way, when one of the micro services fails, other micro services can run stably [7].

Redis is used to cache real-time data, and MySQL relational database is used to store static data and historical data. Microservices decouple the call relationship among microservices through Kafka message queue. The architecture can change the characteristics of microservices through configuration parameters, and manage and maintain the configuration information of each microservice through the configuration center. It can monitor the running of microservice in real time. Restful API interface is used when the third-party system wants to call its own microservices or the microservices inside the system call each other.

4.2. Platform Interface Design

This work mainly includes the interface design of the third-party system and data acquisition. The interface design is shown in Figure 5:

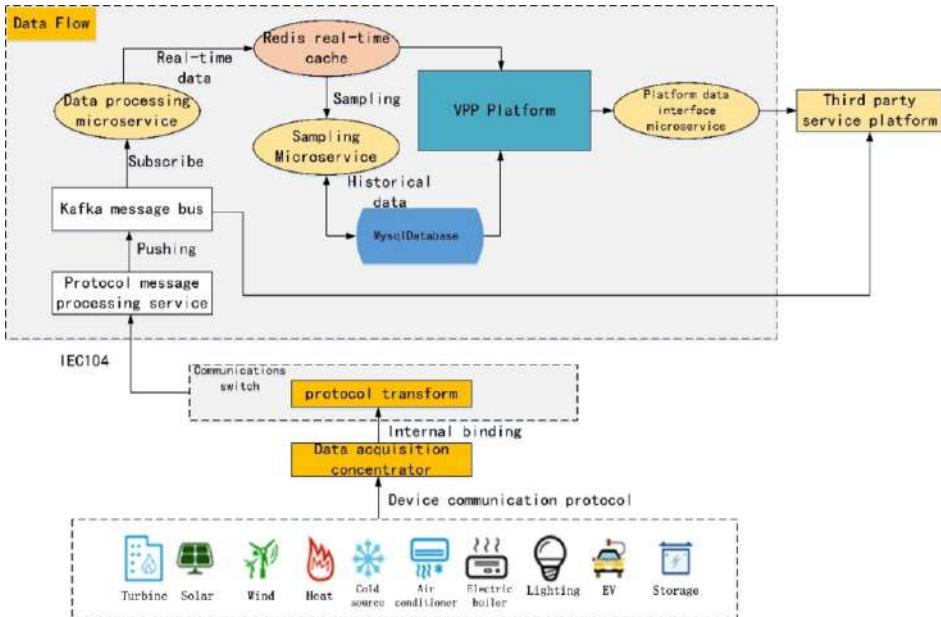


Figure 5. Platform Interface Design Drawing

The data collection of various energy equipment is connected to the platform through IEC104 protocol. After the protocol processing program processes the mature data, it pushes the mature data into the related topics of Kafka message bus. The microservice subscribing to this topic obtains the data for real-time analysis, stores the real-time data into the real-time database Redis, and the sampling micro service samples the real-time data and writes it into the history database MySQL. By defining the three-party interface microservice, real-time data, static data and historical data are obtained from Redis and MySQL, and the data is released to the third-party system through the standard restful interface.

4.3. Software Function Design

The software architecture of VPP platform is divided into four layers, namely hardware layer, platform layer, basic applications layer and advanced applications layer [8]. The software architecture is shown in Figure 6.

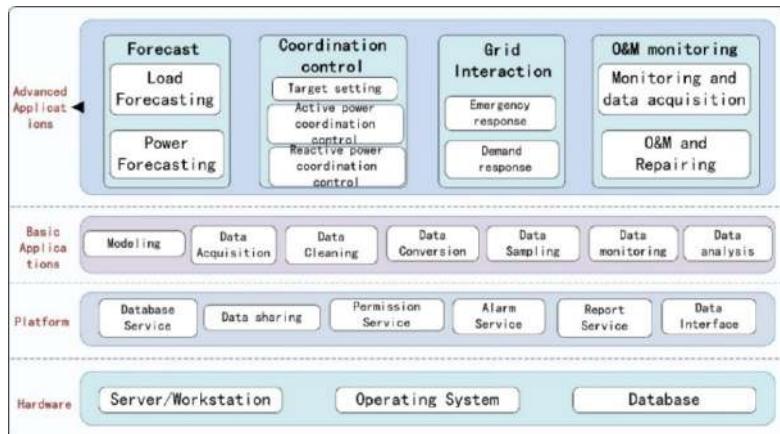


Figure 6. Software architecture of VPP platform.

4.3.1. Basic Application Functions

Modelling

It has a model covering the whole business area of VPP and meets the following functional requirements:

- (1) According to IEC61850, IEC61968, IEC61970 and other standards and specifications, a complete set of object model is constructed.
- (2) It has system hierarchical organization model object.
- (3) Equipment model object with system control and analysis.
- (4) Network model objects with physical associations.
- (5) Business model objects with logic.

Data acquisition

The communication mode of data acquisition and control can adopt serial communication, twisted pair communication, optical fiber communication and wireless communication, and the communication protocol can adopt Modbus, DL / t634.5101, DL / t634.5104, etc. Ethernet is used to communicate with grid dispatching automation system and load centralized control system. IEC104, Webservice and HTTP are used as communication protocols.

The acquisition of real-time operation information mainly includes but is not limited to the following contents:

- (1) Remote data of power generation, heating, cooling and gas supply monitoring system.
- (2) Operation data of power grid, cooling / heating network, gas supply network, etc.
- (3) Electric, heat, cold, gas and other load operation data.
- (4) Operation data of power storage, heat storage, cold storage, gas storage, charging pile and other equipment or systems.
- (5) Electricity, heat, cold and gas metering information.
- (6) Data acquisition of temperature, pressure, flow, heat and valve opening in heat exchange station.
- (7) Power quality data.
- (8) Meteorological and environmental monitoring data.
- (9) Other system data.

Data cleaning

Data cleaning is to solve the problem of data quality and make data more suitable for mining. The result of data cleaning is to deal with all kinds of disordered data and get standard, high-quality and continuous data, which provides data sources for big data statistics and big data mining. The main functions include:

- (1) Identify the integrity of supplementary data
- (2) Screening the uniqueness of data
- (3) Study and judge the legitimacy of data.

Data conversion

Data conversion is modeling based on device container, which provides support for multi-source big data standardization collection. No matter what equipment, system, protocol and communication mode are connected, the integrated energy management and control platform can be easily connected, and all data collection tasks can be completed in one data acquisition system, it is convenient for model synchronization between different platforms, realizing the same function processing logic and providing the same external services, so as to realize the unified operation, maintenance and display of different collection tasks. Be different from the conventional database architecture change, data conversion can adapt to the increasing amount of data, the original data architecture is unreasonable, cannot meet the requirements of all aspects, from the replacement of the database, the replacement of the data structure, into the conversion of the data itself. The main functions are as follows:

- (1) Support a variety of communication methods (analog, digital, dial-up, network) and communication structure (point-to-point, multi-point collinear, star, ring).
- (2) It supports online extension protocol types, provides standard protocols such as IEC60870-5-101 and IEC60870-5-104, supports industrial control communication protocols such as MODBUS, and also supports user-defined protocols. For authorized users, the specification library can be expanded by API functions provided by the system.
- (3) In addition to conventional measurement data, data types are also compatible with unstructured data such as pictures and video streams.
- (4) It supports wired and wireless communication; It supports two logical communication modes: fixed IP terminal and dynamic IP terminal.
- (5) Multi security mechanism: it has multiple security mechanisms, such as dual standby, multi machine switching, data encryption, etc.
- (6) Provide a series of convenient and practical data query interface. Including: hardware interface configuration, channel code monitoring, various types of data monitoring, device status monitoring, bit error rate monitoring, database configuration.

Data sampling

According to the massive data types sent by the front-end acquisition server, such as alarm information, video information, remote control information, analog quantity, state quantity, etc., the platform classifies them according to the data types, provides different optimization processes of data writing to the database, and writes the data to the real-time database or historical database respectively. To maximize the processing capacity of data server and application server. After the data is stored in the database, the platform provides data sampling and reading cycles according to real-time data and non-real-time data: historical data provides sampling cycles ranging from 5 minutes, 10 minutes, 15 minutes, 20 minutes, 30 minutes and 60 minutes. Real time data provides a sampling period of 5 seconds (Once the alarm information, switch position and protection status change, they will be read immediately.) The main functions are as follows:

- (1) Provide massive data writing
- (2) Provide data sampling and reading period ranging from 1 second to 60 minutes
- (3) Provide data warehousing correctness detection
- (4) It provides integrated association of graph, model and library, and generates graph from library and library from graph.
- (5) Provide database read permission by class-table-domain.

Data monitoring

On the one hand, data monitoring monitors the broadcast message and point-to-point message receiving sending transmission status in the process of data transmission, on the other hand, it monitors the network traffic and system resources caused by data transmission. The main functions are as follows:

- (1) Provide process status query, log query, message sending and message receiving query.
- (2) Provide network operation condition monitoring and network flow control.
- (3) Provide system resource monitoring alarm, including CPU, memory, hard disk monitoring.
- (4) Provide logging function, system exception log, system operation log, application operation log, etc.

Data analysis

The total power, electricity consumption and sales electricity of VPP and internal virtual units are analyzed; It makes a comparative analysis of ranking and comparison from the aspects of the trend of electricity, the nature of electricity consumption, the type of industry and the composition of electricity consumption.

4.3.2. Advanced Application Functions

Load forecasting

Load forecasting of VPP includes forecasting energy consumption, maximum load, minimum load, and the time interval of maximum load and minimum load. When drawing hourly load curves of typical days in different seasons, it should be superimposed hour by hour according to the type and nature of each load.

- (1) Combined with historical data, weather, temperature and other factors, the basic operation status of load resources is predicted;
- (2) Understand load resource control variables, control methods, influencing factors and customer comfort requirements;
- (3) Based on the basic operation state of load resources, load regulation methods and comfort requirements, the adjustable potential of load resources is analyzed [9].

Generation forecasting

Through the identification, correction and similar day clustering of historical meteorological and wind power, photovoltaic, other power generation data, the dynamic optimal historical data set is established. The prediction results of different prediction algorithms under different classification clusters are analyzed and compared, and the adaptive optimal weighted combination generation prediction model is established. Realize the short-term power generation prediction of virtual power plant, and provide the short-term prediction results with a time resolution of 15 minutes in the next 1 to 3 days.

Coordination control

Considering the security and stability constraints of power grid, load fluctuation and uncertainty of distributed generation output, the optimization mathematical model of VPP is established to maximize the economic, environmental and social benefits. It has the generation power prediction technology of distributed generation in VPP, focusing on the impact of climate conditions on the output of distributed generation.

Operation target setting

Operation objectives include two levels: power plant level operation objectives and component level operation objectives. The operation objectives of power plant level refer to the safe operation objectives of the units that can be controlled and regulated by the regional dispatching decision-making level, and the rated power and current of the microgrid; Component level operation target refers to the safe operation target of terminal components in virtual power plant, which generally includes the cut in and cut out wind speed of wind turbine, the rated capacity of photovoltaic inverter, etc. The system provides a friendly running target setting tool, which can be directly called to modify on the screen, and the user authority will be checked before the modification operation.

Through the distributed controllable units such as energy storage, photovoltaic, wind energy and load in the VPP, the stable operation of various modes such as economic mode, high reliability mode and isolated network mode can be realized, and the system can make automatic decision and control.

Active power coordination control

According to the real-time operation status, generation power prediction results, load prediction results and other information of the distribution network, under the constraints of system operation, the real-time optimal dispatching adjustment and control strategy are given for the active power output of each distributed energy in the VPP system, so as to improve the economy of real-time operation of the system.

According to the actual operation status of DGs and load, the power control strategy of each period in the future dispatching cycle is formulated to form a generation dispatching plan in the dispatching cycle, which is distributed to the controllers of each DER in the VPP system for execution. The objectives of switching power control, smooth output control and economic optimization operation can be realized respectively, the operation control mode can be switched manually or automatically according to the real-time operation status [10].

Reactive power coordination control

Control the system voltage to ensure that the system voltage operates in a safe area, and ensure the safety and reliability of the system operation.

The intermittent power supply (such as photovoltaic, wind power generation, etc.) in VPP system will cause voltage fluctuation at the grid connection point. The system analyses and calculates the data of new energy power generation unit through data acquisition, and then executes the command through output control to adjust / switch the equipment with reactive power regulation ability, including inverter, so as to make the output voltage and reactive power within the qualified range, so as to ensure the purpose of power quality. The reactive power control strategy determines the voltage control of the parallel node and the reactive power control of the inverter.

Emergency response

In case of an accident, the adjustable resources in the VPP area can be controlled urgently to respond to the demand of frequency and voltage regulation of large power grid. According to its own operation state, by adjusting the adjustable resources, the DERs in the VPP is used as the power source to transmit power to the superior power grid. At the same time, the actual energy contribution of the system is counted, and the adjustable margin and the duration of emergency support are estimated.

Demand response

According to the instructions given by the superior power grid in advance, the VPP can respond to the invitation. The instructions issued in advance include specific indicators (demand response capacity, price, revenue, system operation trend, etc.) and invitation time range. In the specified time range, by adjusting the adjustable resources (flexible load, energy storage system, river water source system, etc.), the required operation indexes and related real-time operation trend are obtained.

Monitoring information

(1) The data management module collects the monitoring information after cleaning, mainly the operating parameters of VPP equipment, taking air conditioning as an example, including the power of the refrigerator, the temperature of the chilled water inlet and outlet water, etc;

(2) Organize and backup the monitoring information;

(3) The display method includes the platform interface display and web display of VPP. The operators of VPP can perform relevant operations in the display interface to realize information interaction;

(4) Meanwhile, the monitoring information is transmitted to the user, including web, app and other forms, so that users can grasp the operation of the equipment.

Operation and maintenance repair

Operation and maintenance control module includes business functions such as fault repair, equipment inspection, energy use application, planned maintenance and defect management. In view of the characteristics of various business types and processing processes, the operation and maintenance management module provide public service functions such as application approval, work order processing, stop supply plan / event, process configuration management, and flexibly supports business differentiation processing, Meet the business management needs of different users. In addition, the status monitoring and early warning of each business processing process are realized through alarm service, and the equipment operation monitoring and on-site repair processing are realized in mobile terminal.

5. Communication Architecture

This platform supports DL/T 698.45-2017 protocol (for China electric power industry), DL/T 645, CJ/T, MODBUS and other protocols [11]. The communication architecture is shown in Figure 7.

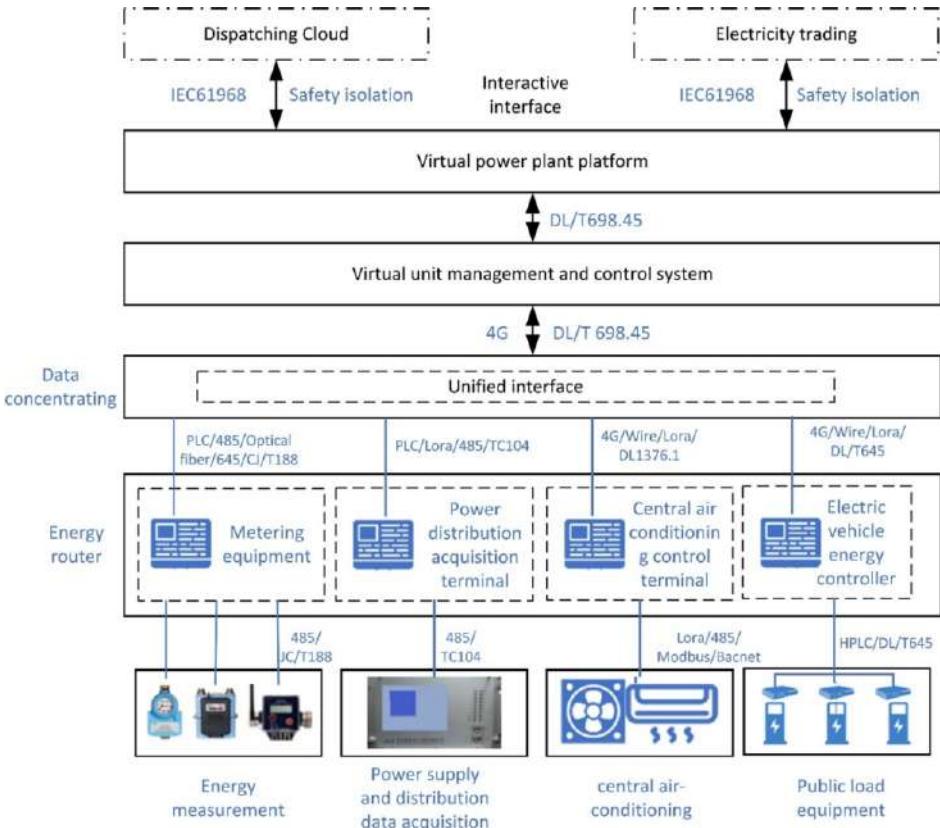


Figure 7. Communication architecture of VPP.

6. Security Architecture

The VPP operation management platform is deployed in the external network (not in utility's network). According to the general network security requirements of Chinese national standards and electric power industry standards, the security protection scheme focuses on strengthening the protection design of terminal and virtual unit management and control system security, network boundary security and system security, so as to ensure the safe, reliable and stable operation of the system to the maximum extent.

In the aspect of system security protection, we focus on the host operating system security reinforcement, baseline compliance verification, deployment of anti-virus and anti-malicious code program; In the aspect of network boundary security protection, we focus on communication data encryption and network traffic security monitoring; In the aspect of security protection of virtual unit management and control system, we focus on Ontology security and top-down access security; In terms of terminal and local network security protection, we focus on the identity authentication of energy consuming equipment and lightweight encryption of control strategy issued by virtual unit management and control system.

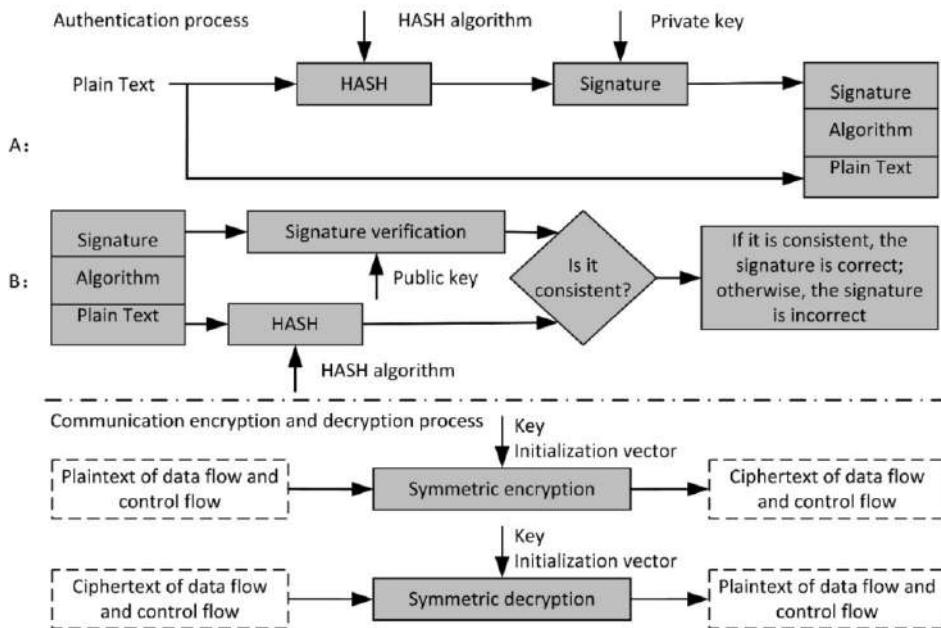


Figure 8. Communication encryption and identity authentication.

Message encryption and identity authentication are very important for the security of VPP platform. As shown in Figure 8, identity authentication is completed based on digital signature and signature verification: A generates fixed length (32 bytes) data for a piece of plaintext data through hash algorithm, and obtains the signature value by signing the fixed length data through private key; B receives the signature value and the plaintext data, uses A's public key to verify the signature value, and performs the same hash operation on the plaintext data, and compares whether the verified data and the hash value are consistent. If they are consistent, the identity of the other party is verified.

The encryption and decryption of communication message is based on symmetric cipher algorithm: the key and initialization vector IV are used for data flow and control flow, and the ciphertext data of data flow and control flow are obtained through symmetric encryption (API interface). Using the same key and initialization vector IV, the original data stream and control stream can be obtained by symmetric decryption (API interface).

The requirements for system data encryption include: the key business data of the system is encrypted and stored; The key business data of the system are transmitted between the systems by encryption and HTTPS; The database connection user has the minimum authority, and the database user password is changed regularly; The database connection information is not exposed to the application system, and the database user password in the application system configuration file must be encrypted; SQL statements are not allowed to be exposed to the client to prevent SQL injection.

7. Case Study

We built a virtual power plant pilot project. The project has 36 distributed photovoltaic power stations, 10 energy storage power stations and 22 controllable loads. On this VPP

software platform, the functions of demand response, energy monitoring, energy analysis, energy optimization and distributed generation trading simulation are realized.

Compared with the traditional demand response software platform, this platform can aggregate a large number of controllable loads which cannot be directly controlled before, with high automation level and fast response; Compared with the same type of foreign VPP platform, the foreign VPP platform focuses on optimizing the generation plan and power consumption plan of distributed generation, while this platform focuses on the control of controllable load. In addition, due to the different rules of China's power market, the functional design of the power trading module of this platform is different from that of the same type of foreign platforms.

8. Conclusion

Through the design and implementation of the virtual power plant platform for demand response, we demonstrated the availability and compatibility of virtual power plant technology in China's power system, and accumulated engineering experience. This shows that the scalability, high availability, fast development and high security of the virtual power plant platform in China's power system can be achieved by using microservice architecture, Internet of Things technology, and communication and security technology complying with Chinese standards, so as to meet the requirements of demand response business. Practice has proved that virtual power plant is a beneficial and powerful technical tool to implement demand response.

In order to realize the practicability and commercialization of this platform, our follow-up work and research are as follows:

- Based on the requirements of microservice architecture, various types of distributed resources are modeled;

- Prepare to use domain driven design method, unify requirement analysis and design programming, and design advanced application function module with Python language;

- Research the big data mining technology included association analysis and cluster analysis and so on, and apply it to the software;

- research the power transaction algorithm using blockchain technology to realize peer-to-peer transaction, and uses it to compile the transaction mechanism and bidding platform module of VPP platform;

- research the distributed optimization algorithm based on leader-follower consistency of Multi-Agent System, and the distributed communication, data collection, information exchange and state coordination among agent units are implemented.

References

- [1] Ai Q. Virtual power plants: the ultimate configuration of the energy internet (Science Press). 2018. p11.
- [2] Yao JG, Yang SC, Wang K. Power grid dispatching control in the interactive environment of "Source-Network-Load" (China Electric Power Press). 2019. p28.
- [3] Gao CW, Li QY, Li HX, Zhai HB. Methodology and operation mechanism of demand response resources integration based on load aggregator. Automation of Electric Power Systems. 2013, **37** (17).
- [4] Han T, Jiang JD. Source-Network-Load-Storage coordinated control system based on distribution IOT cloud platform and virtual power plant. 5th International Conference on New Energy and Future Energy System. 2020, Xi'an, China.

- [5] Jiang HY, Tan ZF and Hu QH. Action analysis of virtual power plants on energy saving and emission controlling of power industry. *Electric Power*. 2010, 43 (6).
- [6] Huang D, Jiang GD, Feng RQ. Smart park integrated management cloud platform architecture based on microservice governance framework. 4th International Conference on New Energy and Future Energy System, 2019, Macau, China.
- [7] Xie CS, Zhong PY. Application system of generic virtual power plant based on service-oriented architecture. *East China Electric Power*. 2013, 41(8).
- [8] Mao T, Guo XX, Xie PP, Zhou J, Zhou BR, Han S. Virtual power plant platforms and their applications in practice: a brief review. 2nd IEEE Sustainable Power and Energy Conference, 2020, Chengdu, China.
- [9] Wang Y, Ai X, Tan ZF, Yan L. Interactive dispatch modes and Bidding strategy of multiple virtual power plants based on demand response and game theory. *IEEE Transactions on Smart Grid*. 2016, 7 (1).
- [10] Li X Y, Zhao DM. Multi-level optimal configuration of virtual power plant considering dynamic characteristics of adjustable resources. *Automation of Electric Power Systems*. 2020, 44(13):17-24.
- [11] Manuel W, Sven L, Albert Z. Experience report: first steps towards a microservice architecture for virtual power plants in the energy sector. *Microservices*. 2019, 2:1-2:10.

Distributed Autonomous Economic Control Strategy for Microgrid Considering Event Triggering Mechanism

Yulong XIONG^{a1}, Shihong MIAO^b, Weichen YANG^b, Zhiwei LIU^b

^a *China-EU Institute for Clean and Renewable Energy, Huazhong University of Science and Technology, Wuhan 430074, China*

^b *School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*

Abstract. To fully adapt to the distributed access of renewable energy, microgrid technology has been developed rapidly. Aiming at the coordination and efficient regulation of distributed resources in microgrid, this paper proposes a distributed autonomous economic control strategy for microgrid considering event triggering mechanism. First, a distributed autonomous economic control architecture is built to provide a distributed operation architecture for optimal regulation of the microgrid. Secondly, a distributed secondary control strategy based on the consensus control theory is established to realize the economic allocation of active power as well as safe and stable operation of the microgrid. On this basis, an event trigger protocol based on the consensus error of the control variables is constructed, which is conducive to reduce redundant communication. The stability of the event trigger protocol is deduced by means of Lyapunov function analysis. The simulation analysis based on the equivalent microgrid verifies that the proposed control strategy can reduce redundant communication and acquire fair distribution of reactive power and active power among DGs, realizing distributed, economical and safe operation of microgrid.

Keywords. consensus control, distributed control, autonomy economical control, event trigger

1. Introduction

With the continuous penetration of distributed generation (DG) such as distributed renewable energy and energy storage equipment [1], the structure and operating characteristics of the microgrid are becoming more and more complex [2], and the controllable data of the system is also showing explosive growth. In this context, the disadvantages of traditional centralized control such as low communication efficiency and poor robustness have become increasingly prominent [3]. Therefore, in order to effectively cope with the challenges brought by the large-scale access of distributed power sources, it is urgent to conduct in-depth research on microgrid control technology [4].

¹Corresponding author: Yulong Xiong, China-EU Institute for Clean and Renewable Energy, Huazhong University of Science and Technology, Wuhan 430074, China; E-mail: 450261823@qq.com.

Adopting distributed control mode, distributed control strategy realizes information exchange between multiple control units through limited communication links and completes optimized control independently within each control unit [5], and then achieve the convergence and optimization of global control strategy based on consensus algorithm [6]. Compared with centralized control strategy, distributed control can realize the decomposition of communication and calculation pressure of control center, which leads to better real-time performance and robustness thus can better adapt to the plug and play of distributed control resources. It can also effectively respond to the demand for flexible adjustment of the system topology of the microgrid [7].

At present, there has been a considerable amount of research in the field of distributed control of microgrid. Olati-Saber [8] proposed a basic theoretical framework for the analysis of consensus algorithms in multi-converter networks, which lays an algorithm foundation for the research of consensus control in various fields. Hug. [9] built a distributed multi-converter system of distribution network, which aggregates controllable units such as distributed generations and electric vehicles, thereby dividing the system into multiple independent converters that operate autonomously. This study provides the basis of the operation architecture for the distributed collaborative optimization of the microgrid. Li [10] proposed an microgrid control strategy based on a hierarchical consensus algorithm. With each single power source as the control object, a three-level control model of the microgrid was established based on droop control and a consensus algorithm, which realized the distribution control of microgrid. Based on the topology of the controlled key nodes, Le et al. [11] aggregates the distributed generations and reactive power compensation devices in the microgrid into active and reactive power control groups respectively, which helps to extract the hidden loading status of each subgrid. Sun [12] proposed an event-triggered model predictive control strategy in a dynamic way, which is able to stabilize the system without the constraints of stability-related terminal. However, the consensus control strategy in researches above fail to comprehensively consider the economic characteristics of distributed control resources. Meanwhile, the continuous communication process of these researches will also cause continuous occupation and waste of communication bandwidth.

In view of the above problems, this paper proposes a distributed autonomous economical control strategy for microgrid considering event triggering mechanism. In Section II, the distributed collaborative control theory is first built to provide the basis of distributed consensus control of the microgrid. In Section III, the consensus control strategies based on traditional droop control is proposed considering the distributed, economical and stable operation of the microgrid. Section IV proposed an event triggering mechanism based on the consensus control strategy in Section III to reduce the redundant communication. Section V provides simulation studies to verify the effectiveness of the proposed control strategy and event triggering mechanism based on an equivalent microgrid. Section VI concludes the paper.

2. Consensus Control Theory

Based on the principle of consensus control, this paper formulates the local control strategy of each converter considering limited information communication and realizes the distributed autonomous control of the microgrid. Specifically speaking, each

converter collects limited state information such as local voltage, current, active power and reactive power, and exchanges state information with neighboring converters to solve the local control variable adjustment strategy. With the iteration of “acquisition-communication-control-acquisition”, the operation state of all converters will finally reach a consistent convergence. The state variable (or consensus variable) of the i -th converter is denoted as φ_i , and the state equation of the user converter is:

$$\varphi_i(k+1) = [\sum_{j \in N_i} w_{ij}(\varphi_j(k) - \varphi_i(k)) + T_i(k)] + S_i \varphi_i(k) \quad (1)$$

where N_i is cluster of converters adjacent to the i -th converter; w_{ij} is the communication weight between the i -th and j -th user converter; T_i is the track controller which can be designed according to user demand; S_i is the self-feedback coefficient.

Based on (1), when the state information of each converter meets the following conditions, the system will achieve consistent stability.

$$\lim_{t \rightarrow \infty} |\varphi_i(k) - \varphi_j(k)| = 0 \quad (2)$$

The Metropolis method is adopted in this paper to calculate w_{ij} :

$$w_{ij} = \begin{cases} 1 / (\max[l_{ii}, l_{jj}] + 1) & j \in N_i \\ 1 - \sum_{k \in N_i} w_{ik} & i = j \\ 0 & \text{other situation} \end{cases} \quad (3)$$

where l_{ii} represents the number of converters which have communication with the i -th converters.

3. Distributed Autonomous Economical Control Strategy for Distribution Network

In order to achieve the distributed autonomous economic control objectives, this paper proposes a two-layered control architecture for the user converters. The primary control adopts a droop control strategy to realize the equal proportion allocation of active power. The secondary control realizes the secondary adjustment of the system frequency, voltage and reactive power:

$$\begin{cases} f_i = f^* - k^P IC_i + \gamma_i \\ V_i = V^* - k^Q Q_i + \lambda_i \end{cases} \quad (4)$$

where f_i and V_i are the frequency and voltage of the i -th converter, respectively; γ_i and λ_i are the secondary control variables of active and reactive power control

strategies, respectively; k^P is the droop control coefficient of incremental cost rate and frequency of microgrid; k^Q is the droop control coefficient of reactive power and voltage of microgrid.

The active power secondary control variable of the i -th converter includes the incremental cost rate secondary control variable (γ_i^{IC}) and frequency secondary control variable (γ_i^f) of the i -th converter's active power output:

$$\gamma_i = \gamma_i^{IC} + \gamma_i^f \quad (5)$$

According to the consensus control theory, the following incremental cost rate secondary control variable is constructed to realize the economical allocation of active power.

$$\begin{cases} \gamma_i^{IC} = (IC_i^* (t) - IC_i) \left(k_p^{IC} + \frac{k_i^{IC}}{s} \right) \\ IC_i^* (t) = IC_i (t) + \left[\sum_{j \in N_i} w_{ij} (IC_j (t) - IC_i (t)) \right] \end{cases} \quad (6)$$

where $IC_i (t)$ is the measured value of incremental cost rate of the i -th converter at time t ; IC_i^* is the reference value of incremental cost rate of the i -th converter updated by the consensus control protocol at time t ; k_p^{IC} and k_i^{IC} are the proportional and integral control coefficients of the PI controller, respectively, N_i is the cluster of converters connected to the i -th converter.

Meanwhile, in order to meet the frequency control objectives of the microgrid, the following frequency secondary control variable with reference value is constructed.

$$\begin{cases} \gamma_i^f = (f_i^* (t) - f_i) \left(k_p^f + \frac{k_i^f}{s} \right) \\ f_i^* (t) = f_i (t) + \left[\sum_{j \in N_i} w_{ij} (f_j (t) - f_i (t)) + b_i (f^* - f_i (t)) \right] \end{cases} \quad (7)$$

where $f_i (t)$ is the measured value of the i -th converter's frequency at time t ; $f_i^* (t)$ is the reference value of the i -th converter's frequency at time t ; k_p^f and k_i^f are the proportional and integral control coefficients of the PI controller, respectively.

Likewise, the voltage secondary control term can also be divided into two parts:

$$\lambda_i = \lambda_i^V + \lambda_i^Q \quad (8)$$

In terms of voltage control, two issues need to be considered. One is to realize the tracking of the PCC node voltage to the reference voltage, and the other is to make the voltage of all user converters converge to a certain average value to prevent the system

voltage from exceeding the safety limit. Thus, the voltage secondary control strategy is constructed as follows.

$$\begin{cases} \lambda_i^V = (V_i^*(t) - V_i) \left(k_p^V + \frac{k_i^V}{s} \right) \\ V_i^*(t) = V_i(t) + \left[\varepsilon_{avg} \sum_{j \in N_i} w_{ij} (V_j(t) - V_i(t)) + b_i (V_{PCC}^* - V_{PCC}(t)) \right] \end{cases} \quad (9)$$

Where $V_i(t)$ is the measured value of the i -th converter's voltage at time t ; $V_i^*(t)$ is the reference value of the i -th converter's voltage at time t ; ε_{avg} is the voltage average convergence factor, used to adjust the allowable space for voltage deviation of each converter; k_p^V and k_i^V are the proportional and integral control coefficients of the PI controller, respectively.

In order to meet the goal of proportional allocation of reactive power in the microgrid, the following reactive power secondary control strategy is constructed.

$$\begin{cases} \lambda_i^Q = (Q_i^*(t) - Q_i) \left(k_p^Q + \frac{k_i^Q}{s} \right) \\ Q_i^*(t) = Q_i(t) + \left[\sum_{j \in N_i} w_{ij} (Q_j(t) - Q_i(t)) \right] \end{cases} \quad (10)$$

where $Q_i(t)$ is the measured value of the i -th user converter's reactive power output at time t ; $Q_i^*(t)$ is the reference value of the i -th converter's reactive power output at time t ; k_p^Q and k_i^Q are the proportional and integral control coefficients of the PI controller, respectively.

4. Event Triggering Mechanism

In order to realize effective communication within the microgrid and save microgrid communication and computing resources, the event trigger mechanism is introduced. Figure 1 and Figure 2 show the information communication architecture and process of event triggering mechanism in the microgrid. ε_i^{event} represents the disturbance to i -th user converter.

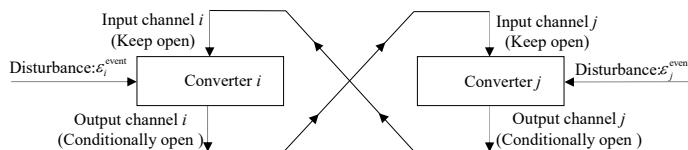


Figure 1. Information communication architecture diagram based on event triggering mechanism.

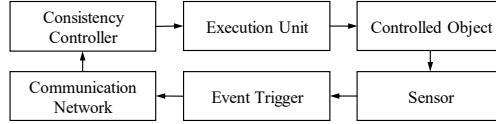


Figure 2: Event triggering process diagram.

This paper constructs the event trigger function of the distributed autonomous control of the microgrid as follows:

$$\left\| e_{x_i}(t) \right\|_2^2 > \eta_i \left\| \sum_{j \in N_i} a_{ij} (x_i(t_{tri,i}) - x_j(t_{tri,j})) \right\|_2^2 \quad (11)$$

In (11), $e_{x_i}(t) = x_i(t) - x_i(t_{tri,i})$. $x_i(t)$ is the state variables of the i -th converter. In this paper, $x_i(t)$ includes $IC_i(t)$, $\omega_i(t)$, $V_i(t)$ and $Q_i(t)$. $e_{x_i}(t)$ is the measurement deviation of $x_i(t)$ at time t . $t_{tri,i}$ is the latest event trigger time. η_i is the proportional control coefficient.

Studying the event trigger function shown in (11), it can be told that the left term of the inequality represents the deviation of the current operating state of the i -th converter from the operating state of the latest trigger time. Meanwhile, the term on the right side of the inequality represents the consistency deviation between the i -th converter and the adjacent j -th converter at the time of the latest event triggering. If the i -th converter cannot maintain the stable state of the previous event trigger when disturbance occurs, the measurement deviation $e_{x_i}(t)$ will increase. When $e_{x_i}(t)$ exceeds the given threshold, which is $\eta_i \left\| \sum_{j \in N_i} a_{ij} (x_i(t_{tri,i}) - x_j(t_{tri,j})) \right\|_2^2$, the i -th converter will not be able to maintain stable operation based only on the system operation information obtained when the last event trigger, and system consistency will be destroyed. At this time, the information communication channel between the i -th converter and its neighbor converters needs to be opened to maintain the consistency and stability of the system.

Based on the given control strategy in Section III. B, the dynamic characteristics of the system can be expressed as (12) and (13).

$$\dot{X}(t) = -cL\hat{X}(t) = -cL(X(t) + E(t)) \quad (12)$$

$$\begin{cases} X(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T \\ \hat{X}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_N(t)]^T \\ E(t) = [e_1(t), e_2(t), \dots, e_N(t)]^T \\ \hat{x}_i(t) = x_i(t_{k,i}) \end{cases} \quad (13)$$

where L represents the Laplace matrix of the system topology diagram; c is the control constant.

To verify the stability of the proposed event trigger function, a Lyapunov function is constructed as follows:

$$V(t) = \frac{1}{2} X^T(t) X(t) \quad (14)$$

Bringing (12) into (14):

$$\dot{V}(t) = -c X^T(t) L \hat{X}(t) = -c (\hat{X}^T(t) - E^T(t)) L \hat{X}(t) = -c \hat{X}^T(t) L \hat{X}(t) + c E^T(t) L \hat{X}(t) \quad (15)$$

Noting that:

$$E^T(t) L \hat{X}(t) \leq \frac{1}{2} E^T(t) L E(t) + \frac{1}{2} \hat{X}^T(t) L \hat{X}(t) \quad (16)$$

Equation (17) can be transformed into:

$$\begin{aligned} \dot{V}(t) &\leq -c \hat{X}^T(t) L \hat{X}(t) + \frac{c}{2} (E^T(t) L E(t) + \hat{X}^T(t) L \hat{X}(t)) \\ &= -\frac{c}{2} \hat{X}^T(t) L \hat{X}(t) + \frac{c}{2} E^T(t) L E(t) \end{aligned} \quad (17)$$

Bringing (12) into (17):

$$\dot{V}(t) \leq -\frac{c}{2} (1 - \lambda_n^2 \eta_{\max}) \hat{X}^T(t) L \hat{X}(t) \quad (18)$$

where $\eta_{\max} = \max \{\eta_i, i = 1, \dots, N\}$ and λ_n is the maximum eigenvalue of the Laplace matrix L . According to (18), it can be told that when control coefficient c and η_{\max} meet the condition that $c > 0$ and $0 < \eta_{\max} < \frac{1}{\lambda_n^2}$, $\dot{V}(t) \leq 0$. Under this condition, formulating system communication and control strategies based on the event-triggered function shown in (11) will ensure the consistency and stability referring to LaSalle invariance principle.

5. Simulation Analysis

In this section, the case studies based on a test microgrid system are carried out. The simulation results of load changes and converter switching are shown and analyzed, which verifies the effectiveness of the proposed control strategy and event trigger mechanism. All simulations are carried out via PSCAD/EMTDC.

5.1. Simulation System

In order to verify the effectiveness of the control strategy proposed in this paper, a simulation model of the AC microgrid as shown in Figure 3 was built. The microgrid is constructed by 3 converters those are coupled by several lines. Converter 1 and Converter 2 are connected to the superior level grid node through public connection bus (i.e. the PCC bus), while Converter 3 is coupled with Converter 2 and has weak connection to the superior node. The incremental cost rate of each converter is simplified to $IC_i = 2\alpha_i P_i + \beta_i$. The specific value of the cost parameters of the converters are shown in Table 1. Parameters of the coupling lines among different converters are shown in Table 2. Main parameters such as PI coefficients are shown in Table 3. The rated voltage of the system is 0.38 (kV), while the rated frequency is set to be 50 (Hz). The public connection bus of the system is equipped with loads, of which the active load is resistive load and the reactive load is inductive load. The initial load of the system is $80+j30$ (kV•A). The voltage control goal of the system is to realize the tracking of the rated voltage of the PCC bus. The frequency control goals of the converters are to restore the frequencies of themselves to 50 Hz.

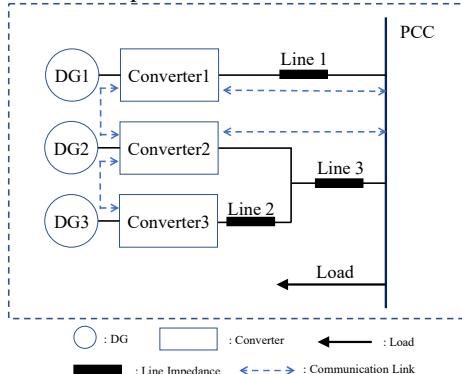


Figure 3. Structure diagram of the simulation system.

Table 1. Cost factors of different DGs.

DG	$(\$/\text{kW}^{-2})$	$(\$/\text{kW}^{-1})$
DG1	0.13	0.059
DG2	0.095	0.0085
DG3	0.082	0.0402

Table 2. Parameters of coupling lines.

Line	Impedance(Ω)
Line1	$0.15+j0.2$
Line2	$0.04+j0.1$
Line3	$0.05+j0.075$

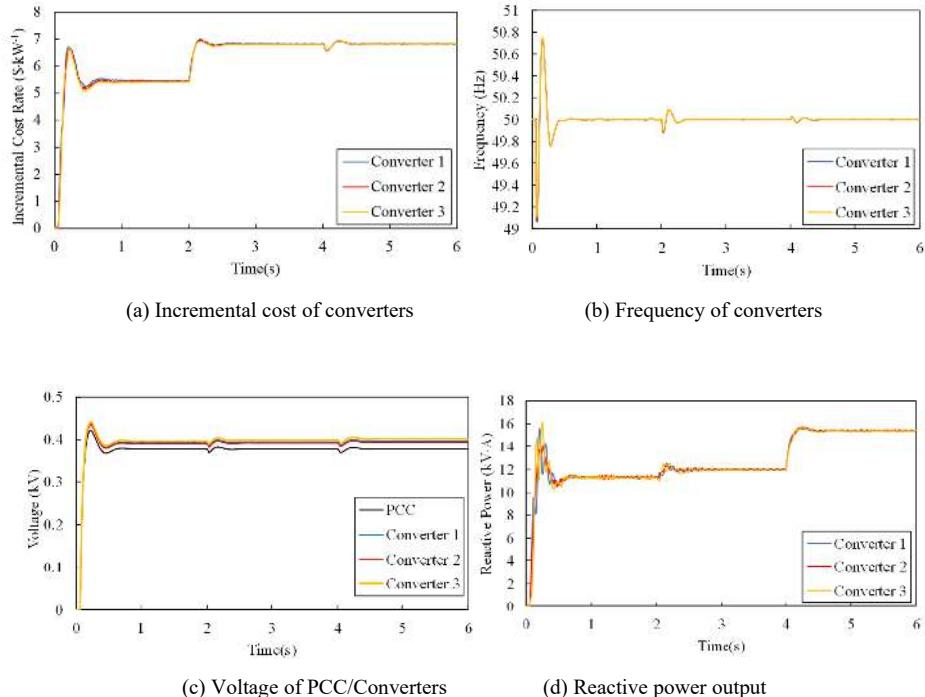
Table 3. Main control parameters.

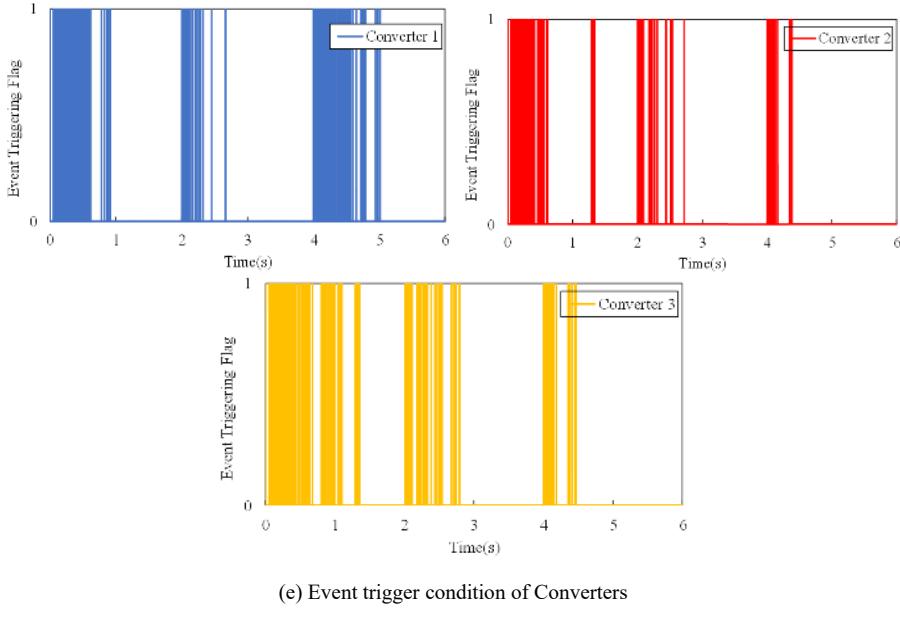
Symbol	value
k^P	0.01
k^Q	0.001
C_{avg}	0.2
$k_p^{\omega} / k_i^{\omega}$	0.5/20
k_p^U / k_i^U	0.1/20
k_p^Q / k_i^Q	0.001/1

5.2. Simulation Results

In order to verify the effectiveness of the discrete and consistent distributed economical control strategy and event trigger mechanism proposed in this paper, this paper considers the microgrid load changes and DG switching, and simulates and analyzes the built microgrid system.

5.2.1. Simulation Results of Load Change





(e) Event trigger condition of Converters

Figure 4. Simulation results of system load change.

Figure 4 shows the simulation results in the case of system load changes. The initial active load of the system is 80kW and the initial reactive load is 30 kVar, respectively. At 2s, the active load is increased to 100kW. Furthermore, the reactive load is increased to 40 kVar at 4s.

It can be seen from Figure 4(a)-(d) that based on the consensus control strategy proposed in this paper, the distributed autonomous economic control of the distribution network user agent can be effectively realized.

Specifically speaking, when the system starts from the shutdown state, it can reach the stable operating state after a 0.5s transient process. During the transient process, the control variables did not experience obvious oscillations. Although the voltage of the PCC node and the frequency of each user agent experienced one and three times of overshoot, respectively, the overshoots are relatively small and the system quickly returned to the normal operating range after overshooting. After entering the steady operation state, the incremental cost rate of all user agents are consistent, which effectively realized the economic allocation of active power among the agents (refer to Figure 5.(a)). The operating frequency of each user agent is consistent with the frequency of the superior grid, which maintains at 50Hz (refer to Figure 5.(b)). Meanwhile, refer to Figure 5.(c), the voltage of PCC is maintained at 0.38kV and the voltage of each user agent is lower than 1.05p.u. of the rated voltage while the voltage gaps among different user agents are small, too. Besides, the reactive power outputs of 3 user agents are basically the same, which realizes the equal allocation of reactive power among the agents. After the distribution network encounters active and reactive load changes in 2s and 4s, the system resumes consistent and stable operation after a short and smooth transient process. The increase of the active and reactive load of the system is shared by each user agent in proportion. Thus, the control strategy proposed in this paper can effectively realize the designed distributed control goal under power fluctuation scenarios, which is conducive to the realization of stable, distributed and economical control of the distribution network. After the microgrid encounters active

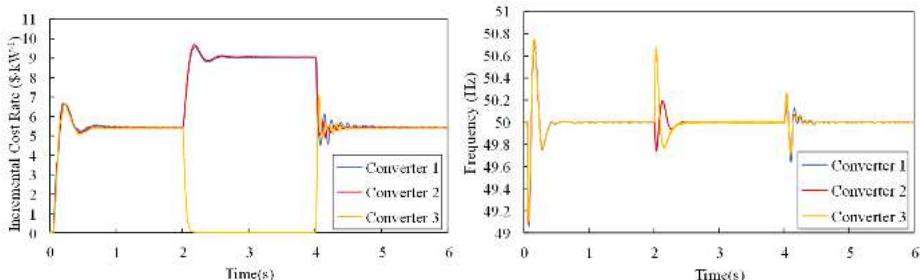
and reactive load changes in 2s and 4s, the system resumes consistent and stable operation after a short and smooth transient process. The increase of the active and reactive load of the system is shared by each converter in proportion. Thus, the control strategy proposed in this paper can effectively realize the designed distributed control goal under power fluctuation scenarios, which is conducive to the realization of stable, distributed and economical control of the microgrid.

According to Figure 4.(e), during the startup and power fluctuation transient process, the event trigger flag of each converter is turned on and remains triggered during system transients. After the system reaches a steady state, the event trigger flag of each converter is gradually turned off. In the whole 6s simulation time, the output channel opening time of Converter 1, 2 and 3 is 22.50%, 19.83%, and 28% respectively. It can be seen that the proposed event trigger mechanism can accurately trigger the information output of the converter when the system experience power fluctuation, and effectively reduces unnecessary information communication thus saves the communication resources of the system.

5.2.2. Simulation Results of Converter Switching

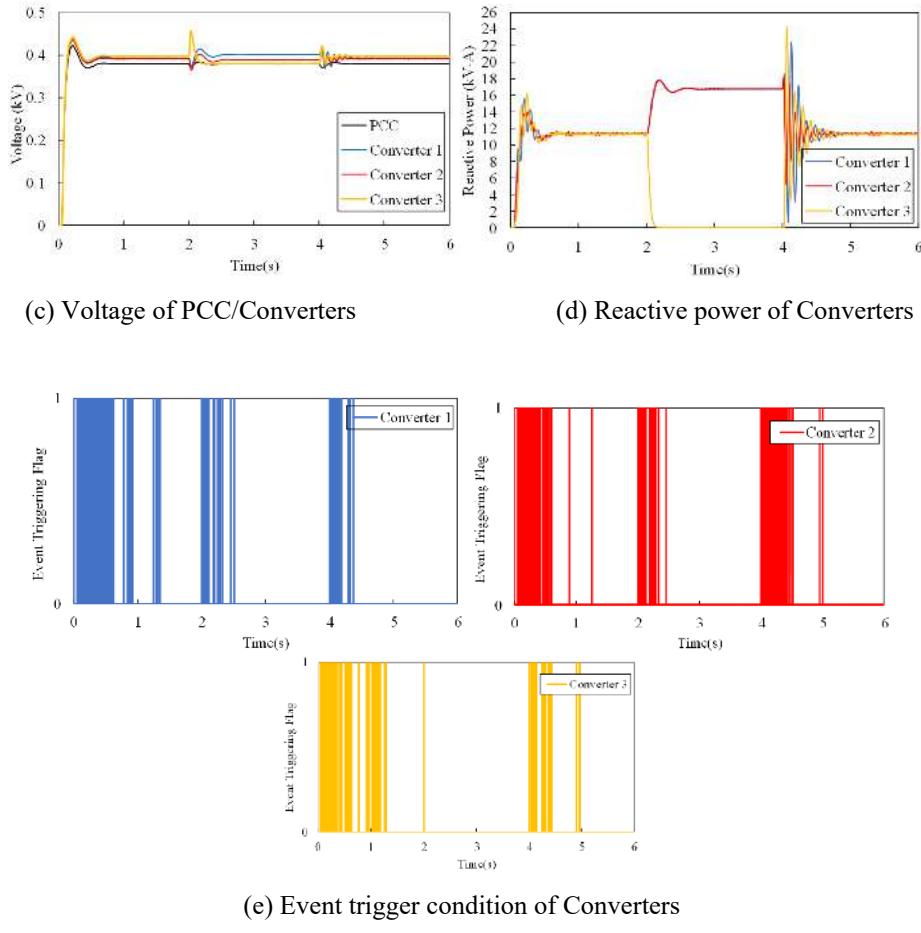
To verify the adaptability of the proposed control strategy and event trigger mechanism to plug and play applications, the simulations of converter switching are carried out. As shown in Figure 3, Converter 3 disconnects from the system at 2s, and reconnects to the system at 4s. In the process of converter switching, the system communication weight matrix is updated based on (4).

It can be seen from Figure 5(a)-(d) that, after Converter 3 exits the operation, the active and reactive power output of Converter 1 and Converter 2 rises rapidly, satisfying the load demand as well as restoring the voltage at PCC and operating frequency to the reference level. In the transient and steady state of the system, adjusted by the active and reactive power secondary control variables, the active and reactive power output of Converter 1 and Converter 2 can maintain the operation principle of equal incremental cost rate allocation of active power and equal proportion allocation of reactive power, thus fulfill the economical operation requirements of the system. After Converter 3 is put back into operation again, the incremental cost rate and reactive power output of each user converge after oscillation, and the system frequency and voltage of PCC can also be restored to the reference value. Therefore, the proposed control strategy can fully adapt to the plug-and-play application requirements of the microgrid, and realize the distributed autonomous economical control of the microgrid.



(a) Incremental cost of Converters

(b) Frequency of Converters



(e) Event trigger condition of Converters

Figure 5. Simulation results of converter switching.

As shown in Figure 5(e), it is the information communication situation of the system before and after the converter 3 is cut out and put in. It can be seen that Converter 3 no longer participates in information communication after switching out. Starting from the moment Converter 3 is cut out, event trigger flags are Converter 1 and Converter 2 are turned on and remain triggered before the system reaches a steady state. After Converter 1 and Converter 2 achieve consistency and stability, the event trigger flags are turned off, while Converter 1 and Converter 2 maintain stable operation based on the latest communication information. Then, after Converter 3 is reconnected to the system, communication is triggered again among the adjacent user converters and the event trigger flags of all converters are turned on. After the system re-enters the steady state, the communication is closed accordingly. The event trigger flags of all converters are turned off and all converters remain uniformly stable operation based on the status information of adjacent converters obtained at the last trigger. Therefore, the event trigger principle based on the consistency deviation proposed in this paper is fully adapted to the "plug and play" application situations.

An example of table styling. It is recommended to add cross references to tables, i.e., please, check Table 1. The style should be switched to Normal.

6. Conclusion

This paper proposes a distributed autonomous economic control strategy for distribution networks that considers event triggering. Regarding the coordinated control of the distribution network, the distributed consensus control strategy is proposed aiming at the economical allocation of active power, the equal proportional allocation of reactive power and the tracking of system frequency and voltage of PCC to the referenced value. On this basis, an event trigger mechanism based on the consistency deviation of control variables is established, which effectively realizes the necessary and efficient communication between adjacent user agents of the distribution network. The simulation studies verify the effectiveness of the effectiveness of the proposed control strategy and event triggering mechanism under different operation situations. Several conclusions can be drawn as follows.

1) The distributed economic control strategy of the distribution network proposed in this paper can effectively maintain the track of system frequency and voltage of PCC to the reference value. It can also realize the economic allocation of active power and proportional allocation of reactive power among user agents. Its good dynamic performance for different scenarios such as power fluctuations and agents switching makes it possess superior practicability.

2) The event triggering mechanism can effectively judge whether the disturbance will lead to changes in the operating state of the system, thereby ensuring the realization of necessary communication and effectively avoiding redundant information communication. Based on the proposed event triggering mechanism, system communication and computing resources can be significantly saved.

3) Applying the distributed autonomous economic control strategy proposed in this article, the solution of the control strategy and the communication of state information can be decomposed to each user agent. Then, under the condition of limited information communication, the proposed control strategy can effectively alleviate the shortcomings of high communication requirements and central information processing pressure of traditional centralized control strategies, thereby promoting the distributed autonomous economic operation of the distribution network.

Since DC/AC hybrid microgrid can better utilize the distributed renewable generation resource, next step of research could be the economical control strategy of DC/AC hybrid microgrid.

References

- [1] H Qu, X Li, L Yang, Y Huang, M Wang, et al. Multi-objective distribution network dynamic reconfiguration and DG control considering time variation of load and DG [J]. High Voltage Engineering, 2019, 45(3):873-881.
- [2] S Li, B Zhang, L Peng, T Mei, X Li, et al. Dynamic island partitioning strategy for active distribution network considering adjustment ability of flexible load [J]. High Voltage Engineering, 2020, 45 (6):1835-1842.
- [3] Q Xu, J Xiao, P Wang, et al. A decentralized control strategy for economic operation of autonomous AC, DC, and Hybrid AC/DC Microgrids [J]. IEEE Transactions on Energy Conversion, 2017, 32(4):1345-1355.

- [4] Z Y Lv, Z J Wu, X B Dou, et al. A distributed droop control scheme for islanded DC microgrid considering operation costs [J]. Proceedings of the CSEE, 2016, 36(4): 900-910.
- [5] H Sun, Z Shen, W Zhou, Q Sun, M Sun, et al. Multi-objective congestion dispatch of active distribution network based on source-load coordination [J]. Automation of Electric Power Systems, 2017, 41(16):88-95.
- [6] H Xin, D Gan, C Dai, H Li, N Li. Virtual power plant-based distributed control strategy for multiple distributed generators [J]. IET Control Theory and Applications, 2013, 7:90-98.
- [7] N Moshtagh, M Michael, A Jadbabaie, K Daniilidis. Vision-based, distributed control laws for motion coordination of nonholonomic robots [J]. IEEE Transactions on Robotics, 2009, 25(4):851-860.
- [8] R Olati-Saber, J A Fax, R M Murray. Consensus and cooperation in networked multi-agent systems [J], Proceedings of the IEEE, 2007, 95(1):215-233.
- [9] Hug, G., Kar, S., and Wu, C. Consensus + innovations approach for distributed multiagent coordination in a microgrid. IEEE Transactions on Smart Grid, 2017, 6(4):1893-1903.
- [10] W Li, C Zhou, R Ye, J Wu. Research on active distribution network control strategy based on hierarchical consistency algorithm [J]. Electric Power Information and Communication Technology, 2016, 14(5):69-75.
- [11] P Lin, C Jin, J Xiao, et al. A Distributed control architecture for global system economic operation in autonomous hybrid AC/DC microgrids [J]. IEEE Transactions on Smart Grid, 2018.
- [12] Z Q Sun, C Li, J H Zhang, Y Q Xia. Dynamic event-triggered MPC with shrinking prediction horizon and without terminal constraint [J]. IEEE Transactions on Cybernetics, 2021 (Early Access).

Analysis of Current Production Status and Key Existing Problems of Tower-Shaped Solar Molten Salt Storage Tank

Hang XI, Qiong WU¹, Xiaojun XIE, Ruigang ZHANG, Bo YANG, Qian LIU

Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China

Abstract. In this paper, the development and prospect of tower-shaped solar thermal power generation technology are briefly introduced, and the importance of production quality of molten salt storage tank in tower thermal power storage system is proposed. The production technology and construction process of molten salt storage tank are described in detail, and the key technology and multiple problems affecting quality are analysed. Aiming at the problem of fillet weld deformation, this paper proposes a new anti-deformation tooling and welding operation technology. At last, this paper presents a construction technology method and a solution to improve the welding quality of molten salt storage tank, which can effectively solve the problem that the bottom plate of molten salt storage tank is out of standard due to welding.

Keywords. tower-shaped solar thermal power generation, construction technology method, salt storage tank

1. Introduction

On September 13, 2016, the National Energy Administration issued 'the Notice on the Construction of Solar Thermal Power Generation Demonstration Projects' (Guoneng Xinneng [2016] No. 223) document, confirmed that the first batch of solar thermal power generation demonstration projects totalled 20, with a total installed capacity of 1.349 million kilowatts. They are distributed in Qinghai Province, Gansu Province, Hebei Province, Inner Mongolia Autonomous Region and Xinjiang Autonomous Region [1]. At present, the domestic construction projects include Dunhuang 100MW molten salt tower solar thermal power generation project, Delingha 50MW molten salt solar thermal power generation project, Hami 50MW molten salt solar thermal power generation project, Qinghai Gonghe 50MW molten salt solar thermal power generation project. Most tower-shaped solar thermal power generation projects are in the experimental stage, and the operating experience is still being explored and improved. To vigorously develop non-fossil energy and gradually replace fossil energy is the primary energy development strategy to ensure China's energy security, achieving sustainable economic development, and reduce pollutants and greenhouse gas emissions [2-3]. The solar radiation is concentrated on a heat absorber installed on a tower of a certain height through heliostats, to collect the heat energy, and steam is generated to

¹Corresponding author: Qiong Wu, Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China; E-mail: wuqiong.0702@163.com.

drive the steam turbine to generate electricity. This method is the solar tower thermal power generation technology. Tower molten salt solar thermal power station is mainly composed of five parts: concentrating system, heat absorption system, heat storage system, steam generation system, and power generation system. The heat storage system is particularly important because the stability of the system is the key link unit of continuous power generation that can be achieved by solar thermal power generation. A delicate design for molten salt volume and wall thickness of tanks in the thermal storage system of CSP plants has been presented in [4]. The strategy can relieve the thermal stress during storage tank preheating and reduce its risk and failure rate. In [5] mainly analyzes the technology and system structure of solar thermal power generation, points out the existing problems and solutions of solar thermal power generation equipment. In [6] presented a method of optimizing the molten salt consumption which can significantly reduce the consumption of molten salt. At present, domestic and foreign countries pay more attention to the production quality and manufacturing process of the equipment in the heat storage system, and even some owners will hire an equipment supervision company to control the heat storage system-manufacturing quality strictly to eliminate hidden production quality problems. In 2016, the Xinyue sandbank power station, the world's largest installed molten salt tower power station, experienced a small-scale molten salt tank leakage accident, which caused the power station to be directly offline and out of service, excluding the planned maintenance period as described by SolarReserve. The power station was unable to generate electricity for at least two months due to the accident. Compared with the economic loss caused by repairing the thermal storage tank, the greater loss lies in the loss of electricity sales revenue caused by the shutdown of the power station. According to the statistics at that time, the monthly generating capacity of the power station reached 30514MWh in September of that year, and the PPA contract price of the project was \$0.135 /kWh. According to this rough estimate, the loss of electricity sales revenue caused by the one-month shutdown was about \$4 million. It can be seen that if the quality problem of molten salt storage tank in the heat storage system is ignored, it will bring significant impact and huge loss to the operating enterprises [7].

2. Introduction to the Production Technology of Molten Salt Storage Tank

Generally, there are 1 low temperature molten salt storage tank (abbreviated as " low temperature tank ") and 1 high temperature molten salt storage tank (abbreviated as "high temperature tank") in the molten salt storage tank of tower solar thermal power generation project. One low temperature tank is used for storing molten salt with a temperature below 400°C, and one independent high temperature tank is used for storing molten salt with a temperature above 400°C. The liquid level of the molten salt storage tank will rise or fall with the operation of the heat absorber and steam turbine unit every day. The design of the molten salt storage tank is suitable for the circulating mode of liquid level operation and the thermal stress changes. Molten salt heat storage medium is adopted in the generator set. The heat storage medium is NaNO₃, KNO₃ binary molten salt and the molten salt heat storage system is equipped [8].

2.1. Molten Salt Storage Tank Manufacturing and Installation Welding

The molten salt storage tank can be divided into three parts: tank bottom, tank wall and tank top. The main manufacturing process contains ten parts which is shown in Figure 1:

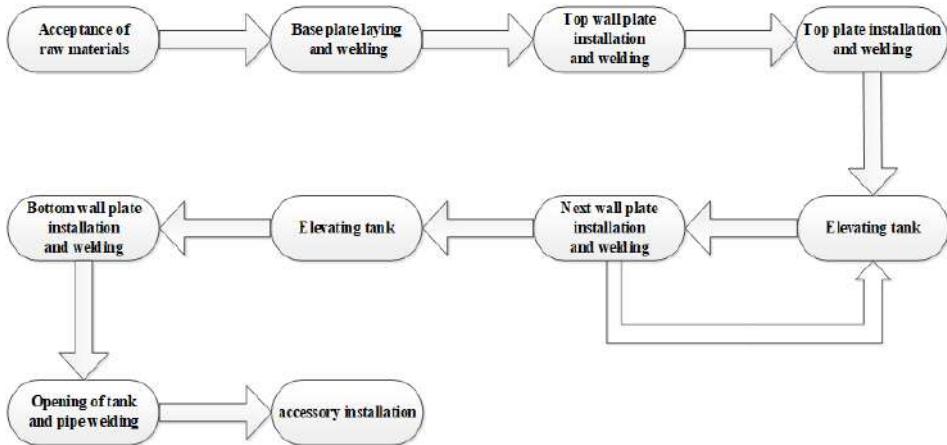


Figure 1. Flowchart of molten salt storage tank manufacturing process.

At present, the main construction parameters of the tower molten salt storage tank are shown in Table 1:

Table 1. Construction parameters of tower type molten salt storage tank.

Project	Interval	Angle
Edge plate joint	Inner 12 ~ 14mm; outer 8 ~ 10mm	55°
Shrinkage joint of middle plate butt joint	6mm	60°
Large Angle seam	2 ~ 3mm	Inside and outside 50°
Circumferential weld of tank wall	1±1mm	outside 60°
Vertical weld of tank wall	2mm	55°

Based on the current domestic construction technology, bottom plate welding is divided into edge plate welding and medium plate welding. The welding of the medium plate is formed by argon arc welding, manual welding with welding rod and manual welding to fill the cover surface. When laying the middle plate, it shall be laid from the center to the periphery according to the layout drawing, and the edge plate shall be laid according to the installation position line. The welding shall be carried out in the order of first short and then long. The primary seam bead shall be welded by subsection back

welding method. When welding the middle plate of tank bottom plate, it shall be carried out in the order as shown in Figure 2:

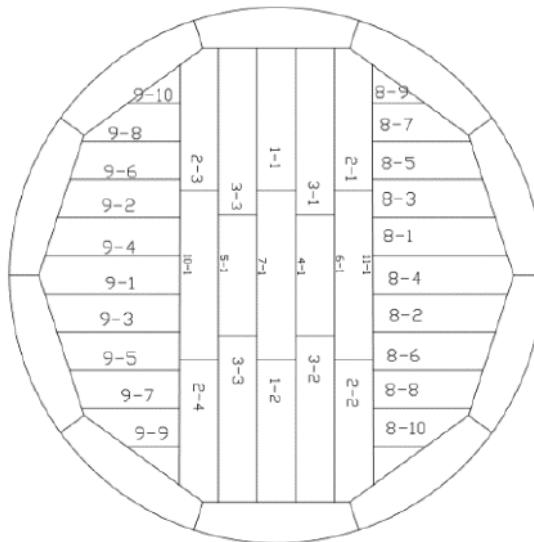


Figure 2. Welding number drawing of tank bottom plate.

Follow the welding sequence and direction shown in the figure above to perform segmented reverse jump welding. The welds in the same group shall be welded at the same time, with the same segment length, the same number of welding layers, the same welding current and welding speed, and the welding joints between layers shall be staggered. In addition, the T-mouth area of 200mm between the middle plate shall be reserved for welding, which shall be completed by manual welding after the completion of long and short seam welding. When welding tank bottom plate, anti-deformation measures should be taken. When welding short seam, the back plate should be made at both ends of the weld 150mm away from the edge (the length of the back plate is not less than 600mm, the width of the back plate is not less than 150mm). When welding long joints use a gantry plate to fix the No. 20 channel steel on both sides of the weld, 200mm away from the weld and clamp it tightly.

The flip-flop method is adopted to install the wall of the molten salt storage tank, that is, the top ring wall and the vault are installed first, then the wall is installed and welded downward in turn, and finally the bottom ring wall and the large fillet weld are installed. According to this installation method, the vertical and circular joints of the panel are welded by argon arc, and then a temporary operating platform is set up inside for root cleaning and welding. The inner and outer sides of the vertical seam are welded by argon electric joint. The outer side of the ring seam is welded by argon electric joint and manual arc welding. The inner side is welded by manual arc welding.

After the top ring panel group is finished, longitudinal seam welding is carried out first. The welding method uses argon electric vertical welding, welding the big groove first, then welding the small groove, and then using the inner operating platform to clean the weld root. Root cleaning can be done by grinder, but the carburizing layer and debris in the planer groove are polished clean by angular grinder.

After the 300mm welding of the outer edge of the bottom ring panel and the bottom plate edge plate is completed, the large fillet weld shall be welded. Arc welding is used

for welding, and the groove is fully penetrated. Therefore, it is necessary to pay attention to prevent large welding deformation due to stress release during welding, which will affect the product quality of molten salt storage tank.

2.2. Welding Method for Various Parts of High and Low Temperature Molten Salt Storage Tank

Because the core work of molten salt tank manufacturing is welding steel material, the welding method and welding material requirements for the welding process of steel materials are particularly important. The main welding methods and characteristics, one is electrode arc welding (SMAW) method, which is characterized by flexible process, strong adaptability, small heat affected zone, simple equipment, easy operation, but poor working conditions, low productivity. The second method is gas metal arc welding (GMAW), which is characterized by high productivity, easy to realize the automation of welding process, flexible operation, but poor wind resistance. The third is to choose submerged arc welding (SAW), which is characterized by high production efficiency and good working conditions, but complex equipment, easy to oxidize metal is difficult to be welded, welding parts are easy to be limited by space position. Fourth, gas shielded tungsten arc welding (GTAW) is characterized by stable arc, high welding quality, easy all-position welding and a wide range of welding materials, but with poor wind resistance and low efficiency [9-10]. SMAW, GMAW, SAW, GTAW are adopted for the steel materials of molten salt storage tanks. During process welding, the temperature between layers should be controlled below 100°C. The specific welding methods and welding materials are shown in Table 2 and 3.

Table 2. Welding method and welding material for low temperature molten salt storage tank.

Number	Welding parts	Welding method	Welding materials
1	Bottom plate	GTAW+SMAW	E5015(J507)
2	Vertical joints on the wall	GTAW+SAW	E5015(J507)
3	Circular joints on the wall	GTAW+SMAW	welding stick /J507
4	Big fillet weld	GMAW+ SMAW	J507+H10Mn2
5	Vault of tank	SMAW	J507

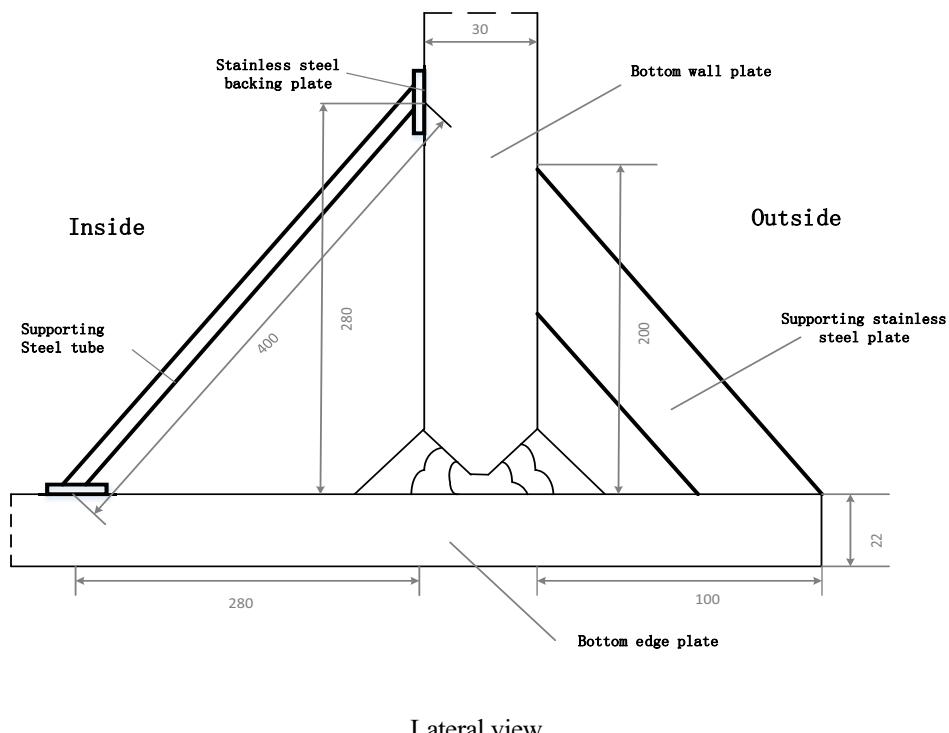
Table 3. Welding method and welding material for high temperature molten salt storage tank.

Number	Welding parts	Welding method	Welding materials
1	Bottom plate	GTAW+SMAW	E347-16
2	Vertical joints on the wall	GTAW+SAW	E347-16
3	Circular joints on the wall	GTAW+SMAW+ SAW	welding stick /E347-16
4	Big fillet weld	GMAW+ SMAW	E347-16
5	Vault of tank	SMAW	E347-16

3. Analysis of Key Technology and Multiple Problems in Molten Salt Storage Tank Production

3.1. Key Welding Process

In the production welding of molten salt storage tank, the fillet welding technology between the bottom plate of high temperature tank and the bottom ring wall plate is particularly important. Due to the large difference between the bottom ring wall thickness and the bottom plate edge thickness, and the groove is full penetration, the fillet weld deformation is also obvious in the welding process. In order to prevent large welding deformation, corresponding measures need to be taken in the welding process. The inner side of the wall plate and the edge plate of the bottom plate should be made of steel pipe, and the stainless steel plate should be welded on both sides of the steel pipe for isolation. The outer side of the wall plate and the edge plate of the bottom plate should be made of stainless steel plate for oblique support, so as to effectively prevent the phenomenon of the edge plate of the bottom plate and the bottom plate from bulging. Welding anti-deformation tooling is shown in Figure 3.



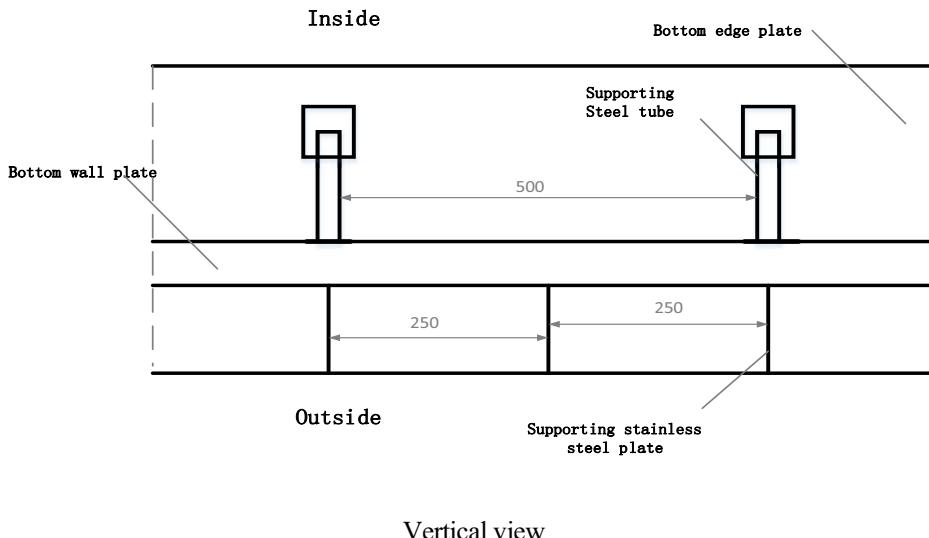


Figure 3. Welding anti-deformation tooling drawing.

Spot welding between the bottom plate edge plate and the backing plate, and the spot welding weld should be as long as possible to prevent the right side of the bottom plate edge plate from warping. At the same time, the heat input, inter-layer temperature and welding sequence should be well controlled during welding. And the local climate temperature should be taken into account when welding the fillet weld, and the welding operation should be carried out in the high-temperature period as far as possible.

The welding personnel should be distributed symmetrically in the circumferential direction of the large fillet weld, and perform welding along the same welding direction at the same time; The welding is separately backwelded in the same direction from the inside and outside of the tank. The inner and outer fillet welds are welded alternately. The inner fillet welds are welded first, and then the outer fillet welds are welded. When welding, the beginning end shall adopt the backward arc method, and the end shall fill the arc pit. Multi-layer and multi-pass welding shall be used for inter-layer welding. The indirect ends of the layers shall be staggered at least 50mm, and the welding speed, welding angle and welding current shall be consistent. The specific welding work should be carried out by welding electrode arc welding from the inside of the tank. The bottom is carried out by welding electrode with small current, fast speed and no swing. The inner side is welded for three layers, and the back side is cleaned, polished and cleaned, and then the penetration detection is carried out. The next procedure can be welded after passing the inspection and detection.

In-process penetration testing shall be carried out after the welding of the primary layer of the large fillet weld between the bottom plate and the wall plate. Permeability inspection shall be carried out on the inside and outside of the fillet welds after all welding. If there are any defects, corresponding schemes shall be formulated to eliminate them. After the tank flushing test is completed, the fillet weld shall be re-inspected in the same manner. It should be noted that the solder joint shall be polished smoothly after the removal of the anti-deformation tooling, and the tank shall be avoided from damage during the removal process.

3.2. Multiple Problems and Treatment in the Welding Process of Storage Tank

After the circular weld welding between the bottom plate edge plate and the middle plate and after the welding of the bottom plate and the wall edge plate is completed. The welding stress is too large to cause the local deformation of the bottom plate and out of tolerance, due to the improper process control in the welding process or the large temperature difference between day and night in the welding environment, which is difficult to control the welding temperature between layers. Serious floor deformation will not meet the requirements of 7.3.2 of GB 50128-2014. Therefore, reasonable rectification measures and effective rectification schemes are put forward to better control the production quality of products.

Firstly, a pressure plate is used to shape the large part of the local deformation of the edge plate of the bottom plate, so that the welding stress is dispersed. By increasing the upper load of the bottom plate, the welding stress is gradually eliminated with time and the increase of the upper load, and the local deformation of the edge plate of the bottom plate is further solved.

After the floor fabrication tooling is completely dismantled, Hawker energy dissipation method can be used to eliminate stress, so that the deformation caused by the floor stress concentration can be controlled within the allowable range of the drawings or standards. The basic principle of Hawker welding stress relief technology is that the fusion welding method is generally adopted in the welding of metal structural parts. In the filling process of metal, excess height, pits and various welding defects are left in the joint parts, resulting in serious stress concentration and a certain amount of welding residual stress. In most cases, residual tensile stress is bad for the fatigue strength of welded structures. So after welding using Hawker energy promote impact tools more than twenty thousand times per second frequency along the direction of the weld position of the weld toe, to produce larger compression deformation, make the weld toe area produces smooth geometry transition, greatly reduces the toe place more than high and depressions caused by stress concentration, adjust the welding residual stress field, eliminate the welding stress. A certain amount of residual compressive stress is generated near the toe, and the material at the toe can be strengthened. Using Hawker energy dissipation method can simultaneously improve several factors affecting the fatigue performance of welds, such as weld toe geometry, residual stress, micro cracks and slag, surface strengthening and other defects. Therefore, this method can greatly improve the fatigue strength and fatigue life of the welded joint, eliminate the welding residual stress, reduce the welding deformation, etc., and can effectively solve the problem that the bottom plate deformation is out of tolerance due to the welding residual stress is not eliminated.

4. Conclusion

To sum up, this paper puts forward the importance of welding quality in molten salt tank manufacturing because the quality problem of molten salt storage tank causes the influence of power generation. Then, the production process and production technology requirements of molten salt storage tanks are introduced in detail. Finally, by analyzing typical quality problems and key process content, this paper proposes the use of increased anti-deformation tooling and Hawker energy dissipation methods to solve the problem of excessive welding stress caused by local deformation of the edge plate of the

bottom plate, while avoiding storage quality defects caused by improper tank welding and other reasons. At present, tower power stations are not mature enough in key technologies, which still need continuous development and exploration. The cost of power generation is also expensive, and a complete and mature system of design, construction, operation and maintenance has not been formed. The problem solution and welding process measures proposed in this paper also hope to play a certain guiding role in the subsequent production work of solar thermal storage tank.

Acknowledgements

This paper was supported by the “Wind-photovoltaic-water coordinating optimization technology research” of science and technology project of Huaneng Group Headquarters (HNKJ18-H35).

References

- [1] Solar Thermal Alliance, Operation record of three solar thermal power stations put into operation in China [EB/OL]. <https://solar.in-en.com/html/solar-2342724.shtml>, 2019-06-19.
- [2] Wang P. Study on design technology of molten salt storage tank for solar thermal power generation. *Qinghai Electric Power*, 2018, 37(3): 37-40 (in Chinese).
- [3] Gao XX. Structure design and performance study of molten salt tank [D]. *Xian: Northwest University*, 2018 (in Chinese).
- [4] Han W, Cui KP, Zhao XH, Francisco M. Design for CSP plants' energy storage system and research on preheating strategy with tanks. *Huandian Technology*, 2020, 42(4): 1674-1951 (in Chinese).
- [5] Li HW. Research on the current situation of solar thermal power generation technology and the existing problems of key equipment. *Electronic Test*, 2020(01):131-132 (in Chinese).
- [6] Sun YH. Design optimization of molten salt quantity in thermal storage system of concentrated solar power plant and its engineering applicatio. *Construction & Design for Engineering*, 2020, (22):101-103 (in Chinese).
- [7] CSPPLAZA. Analysis of seven causes of molten salt leakage accident in Xinyue sandbank Power Station [EB/OL]. <http://guangfu.bjx.com.cn/news/2016/1207/794644.shtml>, 2016-12-07.
- [8] Yuan F, Li MJ, Ma Z, et al. Experimental study on thermal performance of high-temperature molten salt cascaded latent heat thermal energy storage system. *International Journal of Heat and Mass Transfer*, 2018 (118):997-1011 (in Chinese).
- [9] Wang YS, Chu TQ, Wang Z, et al. Experimental study on heat dissipation of molten salt heat storage tank. *Journal of Shenyang Institute of Engineering: Natural Science*, 2019, 15(1):54-58 (in Chinese).
- [10] Yue S, Li M. Solar thermal power energy storage technology and system analysis. *Applied Energy Technology*, 2019(7):54-56 (in Chinese).

Application Status and Research of Instrumentation Device in Solar Thermal Power Generation Industry

Xiaojun XIE, Hang XI, Qiong WU¹, Ruigang ZHANG, Bo YANG, Zengbo LIU

Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China

Abstract. The solar thermal power generation system adopts a dual-axis timely tracking instrument device, which realizes that the sunlight and the central axis of the heliostat instrument device are kept parallel, and greatly improves the utilization efficiency of the light source and the power generation efficiency. At the same time, the study of instrumentation selection in the solar thermal power generation industry cannot be ignored, which can guarantee the normal operation and basic work quality of solar thermal power projects. Therefore, based on instrumentation devices in the solar thermal power generation industry, this article explores the drawbacks of instrumentation devices in the application, and puts forward several research ideas for the drawbacks. Finally, by taking the tower-type solar thermal power generation instrument device as an example.

Keywords. solar thermal power generation system, instrumentation selection, instrument device

1. Introduction

At present, with the trend of global economic development, the content of traditional energy reserves has been decreasing year by year, coupled with the increasing environmental pollution problem, the research and application of new green energy becomes very critical. From December 16 to 18, 2020, the Central Work and Economic Conference clearly pointed out that in order to achieve carbon peak and carbon neutral in the future, accelerating the adjustment and optimization of industrial structure and energy structure, promote coal consumption to peak as soon as possible, vigorously develop new energy, accelerate the construction of the national energy use right and carbon emission right trading market, and improve the double control of energy consumption [1-3]. New energy power generation undoubtedly ushered in an excellent opportunity for development. Solar energy is an inexhaustible source of energy, the use of solar thermal

¹Corresponding author: Qiong Wu, Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China; E-mail: wuqiong.0702@163.com.

power can avoid the reliance on coal, oil, natural gas and other non-renewable energy. In 2019, the installed capacity of grid-connected solar power generation was 204.68 million kilowatts, up 17.2 percent from 2018. By the end of 2019, Qinghai Gonghe 50MW molten salt tower solar thermal power generation project (Figure 1), Golmud 50MW molten salt tower type solar thermal power generation project (Figure 2), Hami 50MW molten salt tower type solar thermal power generation project (Figure 3) and Dunhuang 50MW molten salt linear fresnel solar thermal power generation project (Figure 4) have all been successfully connected to the grid [4].



Figure 1. Qinghai Gonghe 50MW molten salt tower solar thermal power generation project.

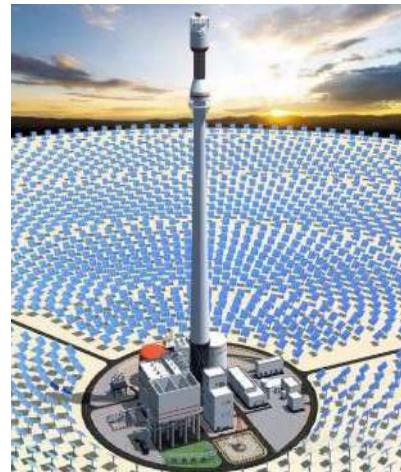


Figure 2. Golmud 50MW molten salt tower solar thermal power generation project.



Figure 3. Hami 50MW molten salt tower solar thermal power generation project.



Figure 4. Dunhuang 50MW molten salt linear fresnel solar thermal power generation.

All of the above indicates that the solar thermal power generation industry, as the foundation of the national power competition in the future, has broad prospects for development. However, there are still some typical quality problems in the current manufacturing and operation stage of the solar thermal power generation industry. Mainly includes the reflectivity of the mirror is unqualified, electric push rod tracking function day by day is failed, steel structure galvanizing treatment is unqualified, the

quality of storage tank steel plate is unqualified, welding quality of molten salt storage tank is unqualified, the high dynamic accuracy of the high pressure rotor of the steam turbine is unqualified, system instrumentation accuracy is unqualified, etc. The above-mentioned typical quality problems are caused by many factors, including design factors, manufacturing factors, installation factors and operation and maintenance factors. The utilization efficiency and generation efficiency of the light source in the solar thermal power generation system cannot be separated from the instrument device. So, it is necessary to study the selection and application of the instrument device. In addition, it is reported that Sun Jiong from Ningbo Ozaki has designed and manufactured the first production line of "armored thermocouple cable and armored platinum resistance cable" in China [5]. The production line successfully shortens the gap between domestic and foreign thermal instrument manufacturing field, and then makes an important contribution to the development of innovative high quality equipment localization. This will also promote the high level of solar thermal instrument manufacturing, and to help us selection and application.

2. Disadvantages of Instrumentation in the Application of Solar Thermal Power Generation Industry

2.1. Poor Application Performance of Instrumentation Device

The selection and application of instrumentation devices in the solar thermal power generation industry play an important role. At present, the design of the instrumentation device is customized in advance according to the needs. The entire design process is completely in accordance with the rules and regulations of the scientific instruction manual to ensure the stability of the instrumentation device [6]. Since the users of the instrumentation device do not pay more attention to the stability of the instrument in the daily use, the poor arrangement of the instrument use position greatly limits the stability of the instrumentation device. In addition, the operator's inattention to the instrument and insufficient understanding of the technology of the instrument device result in the poor stability of the application of the instrument device in the solar thermal power generation industry.

2.2. Insufficient Maintenance of Instrumentation Device

Any instrument device can not be separated from the daily maintenance especially in the solar thermal power generation industry that requires high accuracy. If you do not pay enough attention to the daily maintenance of the instrumentation device and ignore some common small problems. For example, the thermal power generation industry causes long-term chronic corrosion of thermal equipment due to water vapor problem. In a short time, there will be no accident problem in general, but in the long run, it will cause equipment abnormalities. However, without daily maintenance, it is impossible to detect the thermal equipment and reflect the stability of the thermal equipment in real time for the detection of online detection instruments. Therefore, the application of instrument devices in solar thermal power generation projects needs to pay special attention to routine maintenance [7-8].

2.3. Lack of Advancement and Creativity

At present, the manufacturing technology and application function of the instruments used in the solar thermal power generation industry are not advanced enough in our country. Few instrumentation device are developed and applied customized according to the application objects, and there are problems such as the gap between instrument accuracy and foreign products. The above problems seriously restrict the promotion and application of high quality instrumentation products in the solar thermal power generation industry.

3. Application Ideas of Instrumentation Device in the Solar Thermal Power Generation Industry

3.1. Improve the Stability of the Instrumentation Device

In order to improve the stability of the instrumentation, the solar thermal power generation industry needs to adopt a targeted strategy to use different instrumentation device for different power generation projects, while ensuring that the technical standard and specifications of the instrumentation device are the same. In addition, the standard operation and use position of the instrumentation device can also improve the use of stability. In this way, the stability of the maintenance and management of the instrumentation device are ensured, and then the work efficiency of the project construction is improved.

3.2. Pay Attention to the Daily Maintenance of Instrumentation Device

The solar thermal power generation industry would be focused on the importance of online instrumentation applications. The first is to increase the professional team for instrumentation maintenance; the second is to appropriately increase the scientific research funding for instrumentation for the development of high-precision, high-efficiency instrumentation and procurement of parts and replacement and maintenance of parts. In addition, the solar thermal power generation industry needs to strengthen the supervision and management of the instrumentation device, and timely prevent the abnormal application of the online instrumentation device caused by the lack of responsibility of the maintenance personnel and the inadequate maintenance. In terms of instrumentation maintenance, both companies and individuals need to regularly participate in internal communication work, strengthen communication and exchanges between various departments of solar thermal power generation, and improve the application level of instrumentation device [9-10].

3.3. Regularly Calibrate and Overhaul Instrumentation Device

The use accuracy of the instrument directly affects the tracking accuracy, which in turn leads to the reduction of the utilization efficiency and power generation efficiency of the light source of the solar thermal power generation station. It is very necessary to calibrate and overhaul the accuracy of the instrument on a regular basis. Although this work will increase part of the investment, it is very important for the accurate measurement and

normal use of instruments and meters. It can not only ensure the normal operation of the system, but also avoid unnecessary power loss and increase power generation revenue.

4. Application of Instrumentation in the Solar Thermal Power Generation Industry-Taking Tower Solar Thermal Power Instrumentation Device as An Example

4.1. The Structure of Tower Solar Thermal Power Generation Station and Instrumentation

Tower-type solar thermal power generation system is also called centralized solar thermal power generation system. The basic form of the tower-type solar thermal power generation system is to use a group of heliostats that independently track the sun to gather sunlight on the receiver on the top of the fixed tower to generate high temperature, heat the working fluid to generate superheated steam or high-temperature gas, and drive the steam turbine or a gas turbine generator to convert solar energy into electrical energy [11]. Tower-type solar thermal power generation is one of the applications of many projects, and its system includes heliostat array, heat collection system, heat storage system, power generation system and control system. The schematic diagram of the tower-type solar thermal power generation system is shown in the figure 5:

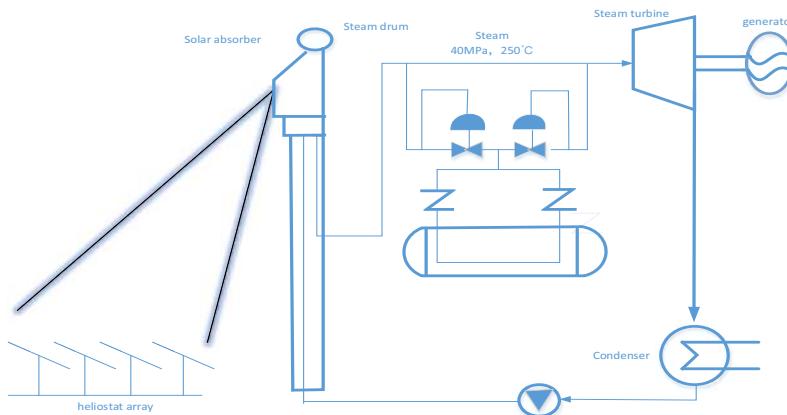


Figure 5. Schematic diagram of tower-type solar thermal power station.

In order to ensure that the system can operate efficiently and accurately, it is necessary to configure relevant instrumentation devices to improve the overall efficiency of the system and the efficiency of power generation. The instrumentation devices involved include heliostats, heat-absorbing heat accumulators, etc.

4.2. Configuration of Tower-Type Solar Thermal Power Generation Instrument

The tower-type solar thermal power station uses the solar energy collection mirror field instrument control system, and under the configuration of wireless instrument and network technology, it can effectively save a large number of control cables and debugging functions, and maximize the overall efficiency of the engineering

infrastructure [12]. The configuration of the tower-type solar thermal power generation instrument process system is as follows: through the placement of the heat-absorbing heat accumulator with high-purity graphite, and the placement of the matching heliostat instrument under each tower to achieve the concentration of sunlight reflection, and converge the concentrated solar beam to the inside of the opening. Among them, the mirror field instrument can use wireless instrument or network technology to control the service in the application configuration [13]. The tower-type solar thermal power generation devices system includes the following instrument control devices. One is that the heat-absorbing and heat-storing instrument device realizes the treatment of the temperature of the medium. The heat-absorbing and heat-storage instrument device needs to work in a suitable temperature range. Once the temperature is too high, it will restrict the safe operation of the system, such as causing the breakdown of the medium layer and the damage of the instrument device. Once the temperature is too low, it will not be able to provide the best power generation technical support for the turbo-generator at the appropriate temperature. The second is the configuration of the steam temperature control instrument for the heat-absorbing accumulator. Because the temperature of the internal heat storage medium will gradually decrease with the steam process (the transfer of heat energy), resulting in a decrease in the temperature difference between the medium and the feed water, it is necessary to control the flow of the feed water through the adjustment function. Among them, the feed water flow rate and the heat storage medium should maintain a proportional relationship, so as to meet the operation requirements of the steam turbine. The third is the configuration of the heliostat control instrument. This configuration is related to the complexity of investment and technology, and is designed reasonably according to actual requirements.

4.3. Network Device Selection of Tower-Type Solar Thermal Power Generation Devices

According to the characteristics of local monitoring equipment, the heliostat instrument of tower-type solar thermal power generation instrument is configured into corresponding wireless multi-combination signal converter, or the heliostat local controller is used [14]. Specifically, the application of local controller of heliostat instrument focuses on the reasonable calculation and state monitoring of heliostat driving motor position instructions. If the heliostat instrument has no local control device, the design of the system needs to be completed under the mirror field control server. In addition, a wireless multi-function node is arranged at the heat absorption and accumulator on the top of the solar tower, and communication is realized between the corresponding heliostat instrument and the transmitter instrument. The power supply mode of wireless instrument device usually adopts the charging mode of local power supply. Since the driving mechanism needs power supply support, only the power cord can be connected, so the corresponding parameters (refresh power, transmit power) are determined according to the actual situation [15]. Wireless transmitter in the data refresh speed can reach 1 second, but also according to the actual situation, reasonable setting of the refresh rate. For example, the wireless transmitter of the heliostat instrument device can collect the angle position in real time, and obtain the altitude angle and azimuth angle of the sun with the help of the meteorological observation station, so as to compare the data with the actual altitude angle and azimuth of the sun, so as to reduce the errors in the long-term operation of the heliostat instrument.

5. Conclusion

This paper mainly takes the tower solar thermal power generation instrument device as an example, from the system introduction, instrument device configuration, device selection of three aspects of the instrument device configuration and network configuration of the comprehensive application. A new application idea is put forward for the application of instruments in the solar thermal power generation industry. In view of the application requirements of this instrument system in solar thermal power generation industry, the selection of heliostat instrument and device is analyzed and studied.

Acknowledgements

This paper was supported by the “Wind-photovoltaic-water coordinating optimization technology research” of science and technology project of Huaneng Group Headquarters (HNKJ18-H35).

References

- [1] Yuan WD. Current development and prospect of solar-thermal power generation in china and abroad. *Power & Energy*, 2015(4):487-490.
- [2] Wei ZK. Current status and perspective for solar power tower technology. *Encyclopedia Forum*, 2019(11):528-529.
- [3] Huang W. Running state detection of heliostat for tower power station [D].*Nanjing Normal University*,2016.
- [4] Solarbe. In 2019, the installed capacity of global solar thermal power generation increased to 6451MW, China contributed 52.41% to the new capacity [EB/OL]. https://www.sohu.com/a/365430945_418320, 2020-1-8.
- [5] Association for Science and Technology Innovation. Ningbo Aoqi: Drive China's Thermal Instrument Manufacturing Technology to the Forefront of the World [EB/OL]. <http://www.csplaza.com/article-11697-1.html>, 2018-3-14.
- [6] Yuan H, Li F, Wang CX. Reason analysis on measuring error of on-line chemical instrument in thermal power plant and improvement measures. *Shanxi Electric Power*, 2019(5):49-51.
- [7] Zhang YJ. Common problems and countermeasures in quality control of power engineering construction projects. *Science and Wealth*, 2020(06):61-61.
- [8] Wang K, Ding L, Zang PW, Sun DK. Analysis on installation of molten salt absorber and selection of related equipment for tower solar thermal power station. *Technology and Market*, 2020(7):53-54, 58.
- [9] Zhao CY. Discussion on the importance of measurement and detection of instruments and meters. *Technology and Information*, 2018, (027):121-123.
- [10]Li PC. Optimization and experimental investigation of cascade Rankine cycle-based solar thermal power generation system [D]. University of Science and Technology of China, 2016.
- [11]Song LN. Design and implementation of heliostat control system [D].University of Electronic Science and Technology, 2016.
- [12]An RC. Design and selection of electric heater for molten salt storage tanks in solar thermal power plants. *China Resources Comprehensive Utilization*, 2019, 037(011):59-62.
- [13]Wang YS. Research on design of molten salt heat storage tank [D]. Shenyang Institute of Engineering, 2019.
- [14]Li CY, Hu LG, XYU Tong-she. Wireless instrument and network configuration of tower type csp power plant. *Electric Power Survey & Design*, 2015(2):78-80.
- [15]Xu FB. Research on structure selection and mechanical property of heliostat supporting structure of dual-aixs tracking CSP system [D]. Beijing: Beijing Jiaotong University, 2012.

Thermal System Simulation Study of Wide-Load Out-Of-Stock Technical Transformation to Reduce SCR Inlet Flue Gas Temperature of 300MW Subcritical Boiler

Tao QIN, Jun RONG¹, Guang YANG, Yankai WANG, Yi HAN, Bin CAI, Yingli YU
Inner Mongolia Electric Power Science and Research Institute, Hohhot 010020, Inner Mongolia, China

Abstract. During the operation of a 300MW subcritical boiler of a power plant, there is a low temperature of the SCR inlet flue gas under medium and low load conditions. In order to effectively solve the problem of low SCR inlet temperature under low load conditions, and improve the adaptability of the coal type. Three kinds of wide load denitrification technology reform schemes are proposed. With the boiler thermal system simulation software BESS, the thermal calculations of the three transformation schemes were carried out. The results show that: the Scheme C is the optimal solution. After the transformation, the temperature of the SCR inlet flue gas increased by 21°C under the ultra-low load condition, and the exhaust gas temperature increased by about 7°C. At the same time, the possible impacts of the reform of the Scheme C and the key issues that need to be paid attention to during the transformation process are evaluated and discussed.

Keywords. 300MW subcritical boiler, SCR inlet flue gas temperature, wide-load denitrification technology transformation, thermal system simulation

1. Introduction

At present, China's coal-fired boilers generally use the SCR process for flue gas denitrification treatment. For coal-fired boilers that use SCR denitrification, the SCR inlet flue gas temperature will decrease as the unit load decreases. The denitrification system catalyst requires a flue gas temperature of 300~400°C. When some power station boilers have low load, the SCR system is forced to withdraw from operation due to the inlet flue gas temperature not meeting the operation requirements, this will cause a significant increase in the NOx emissions of the unit [1]. According to statistics, about 30% of coal-fired boilers in China have problems in the low-load conditions that the SCR system cannot be put into operation due to insufficient smoke temperature.

This paper selects a 300MW subcritical boiler in Inner Mongolia. Aiming at the low temperature of SCR inlet flue gas under low load, several kinds of wide-load denitrification

¹Corresponding author: Jun Rong, Inner Mongolia Electric Power Science and Research Institute, Hohhot 010020, Inner Mongolia, China; E-mail: rongjun123@126.com.

transformation schemes are proposed without affecting the combustion of the furnace, the steam parameters and the stable operation of the medium and high load SCR system. The thermal calculation software BESS is used to calculate the thermal calculation of several schemes, analyze and compare the effects of different schemes, and determine the optimal scheme [2,3].

2. Equipment Introduction

The 300MW boiler is a model HG-1025/17.5-YM11 boiler designed and manufactured by Harbin Boiler Co., Ltd. according to the technology of American ABB-CE company. The boiler is a natural circulation single furnace, one intermediate reheating, balanced ventilation, single steam boiler. The boiler adopts a full steel frame suspension structure, a tight-closed Π -type arrangement, and a solid-state slag discharge. The milling system is a positive pressure direct-blowing pulverizing system equipped with five HP863 medium-speed coal pulverizers produced by Shanghai Heavy Machinery Plant. Two three-point warehouse-type rotary air preheaters are arranged in the lower part of the rear shaft flue. The SCR denitrator reactor allows the flue gas inlet temperature to be between 300°C and 400°C, the schematic layout of the boiler is shown in Figure 1.

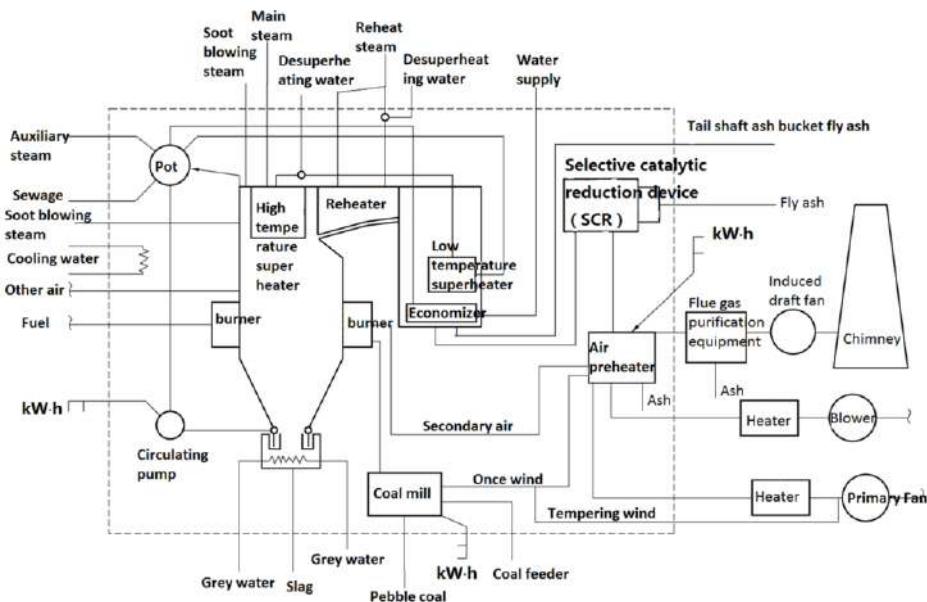


Figure 1. The schematic layout of the boiler.

3. Thermal Calculation Check and Wide Load Denitration Technology Transformation Plan

3.1. Thermal Calculation Check

In order to ensure the reliability of the thermal calculation results, the thermal calculation of the thermal system simulation model is checked according to the actual operation parameters of the boiler, so as to ensure that the check calculation results are in good agreement with the actual operation data of the boiler, and the system model and parameters constructed by the thermal calculation are reliable and effective.

The field test data of 299.3MW, 209.7MW and 75MW are selected for thermal calculation and verification. The coal quality data in the test and the actual operation parameters of the boiler are shown in Table 1 and Table 2.

Table 1. Actual fuel characteristics.

Item	Load conditions		
	299.3MW	209.7MW	75MW
Car (%)	45.32	46.97	48.82
Har (%)	2.34	2.47	2.62
Oar (%)	9.23	9.88	8.84
Nar (%)	0.51	0.53	0.52
Sar (%)	0.59	0.61	0.54
Aar (%)	19.01	16.34	17.07
Mar (%)	23	23.2	21.6
Vdaf (%)	41.41	39.98	37.87
Q _{net,ar} (kJ/kg)	16310	17040	17457

Table 2. Actual operating parameters of the boiler.

Item	Load conditions		
	299.3MW	209.7MW	75MW
Main steam flow (t/h)	918.8	606.8	211.4
Main steam pressure (MPa)	16.30	15.03	11.2
Main steam temperature (°C)	540	541	514
Reheat steam inlet pressure (MPa)	3.56	2.37	0.84
Reheat steam inlet temperature (°C)	321.6	300.8	254
Reheat steam outlet pressure (MPa)	3.36	2.25	0.83
Reheat steam outlet temperature (°C)	543	540	483
Boiler feed water temperature (°C)	270	247	196.6
Boiler feed water pressure (MPa)	17.8	15.9	11.6
Boiler feed water flow (t/h)	913.1	619.8	228.0

Under the load of 299.3MW, 209.7MW, and 75MW, the calibration calculations are carried out respectively, and the calculation results are compared with the actual operating data as shown in Table 3. Among them, the error between the main steam temperature and the actual operating temperature is less than 3°C, the error between the reheated steam temperature and the actual operating data is less than 3°C, and the deviation between the economizer outlet temperature and the actual operating data is controlled within 5°C, the deviation between the exhaust gas temperature and the actual operating data is controlled within 5°C. The unit operating data is measured by on-site sensors, and there will be certain measurement errors. Therefore, the error of the thermal calibration calculation is within an acceptable range, the calculated model and system parameters of the thermal system are reliable and effective. It is worth noting that in the thermal calculation results under the 75MW working condition, there is a certain degree

of under-temperature phenomenon in the outlet temperature of the superheated steam and the outlet temperature of the reheated steam. The inlet flue gas temperature of SCR is significantly lower than the allowable temperature of the denitrification device reactor, and it is in good agreement with the low-load stability test data of the unit, which once again proves the reliability of the established thermal system model.

Table 3. Comparison of calculation results and actual operating data under 299.3MW, 209.7MW, 75MW load.

Item	Actual operating load								
	299.3MW			209.7MW			75MW		
	Operating parameters	Simulation results	Error	Operating parameters	Simulation results	Error	Operating parameters	Simulation results	Error
Superheated steam outlet temperature (°C)	540	543	3	541	543	2	514	511	-3
Reheated steam inlet temperature (°C)	321	321	0	300	300	0	254	254	0
Reheated steam outlet temperature (°C)	543	544	1	540	537	-3	483	481	-2
Working fluid temperature at the inlet of the economizer (°C)	269	270	1	246	247	1	233	230	-3
Inlet flue gas temperature of SCR (°C)	386	384	-2	350	345	-5	276	279	3
Exhaust temperature (°C)	142	142	0	131	126	-5	114	112	-2

3.2. Technical Transformation Scheme of Wide Load Denitration

With the decrease of the boiler load, the flue gas temperature at the outlet of the economizer will also decrease. When the load drops to a certain level, the flue gas temperature at the outlet of the economizer will be lower than 300°C, which can not meet the reaction temperature of denitrification, resulting in excessive NOx emission, catalyst deactivation, increased ammonia escape, etc. The wide load denitrification technology is to increase the flue gas temperature at the outlet of economizer at low load to ensure the chemical reaction temperature of catalyst, so as to achieve the minimum technical output, stable denitrification at full load and all time, and meet the requirements of clean emission [4].

In this paper, three feasible schemes have been selected after actual field tests and preliminary calculations. In scheme A, the bypass at the water side of economizer is adopted, and the regulating valve and connecting pipe are set in front of the header at the inlet of economizer to bypass part of the water supply, so as to reduce the heat absorption of water supply in economizer, and finally achieve the purpose of increasing the flue gas temperature at the outlet of economizer. Scheme B adopts the hierarchical economizer arrangement, which divides the economizer arranged in the boiler tail flue into two parts. The first part is still arranged in the boiler tail flue, with a heat transfer area of about 4030.91m². The second part is arranged behind the SCR reactor, with a heat transfer area of about 2015.45m². The economizer arranged in the boiler tail flue is called the primary economizer, the economizer in the lower part of SCR reactor can be called secondary economizer. Scheme B needs to add new supports and connecting pipes, change the flue after the existing SCR reactor, reinforce the flue, and check the denitrification steel frame and foundation load. In addition, new soot blowing equipment,

auxiliary structures and instruments are needed. In scheme C, the flue gas bypass modification is adopted, and a hole is opened at the flue of the rear steering room to extract part of the high temperature flue gas to the SCR interface. Under low load, the flue gas temperature at the SCR inlet can reach the minimum ammonia injection temperature requirement by extracting the higher temperature flue gas and mixing the flue gas from the economizer outlet. Scheme C needs to increase the flue gas baffle, some steel structures and supports and hangers, and the bypass flue needs to be equipped with expansion joint, electric shut-off baffle and regulating baffle for flue gas flow and temperature regulation. Schemes A to C are shown in Figures 2 to 4.

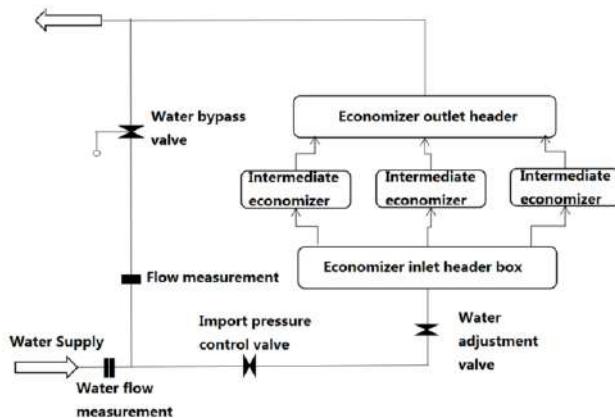


Figure 2. Scheme A, the water side economizer bypass arrangement.

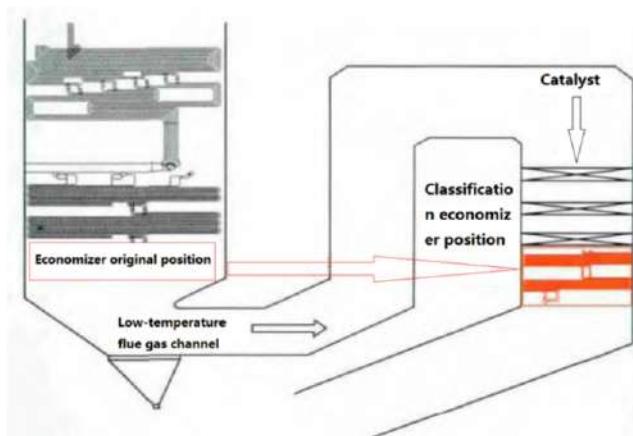


Figure 3. Scheme B, Classification economizer transformation.

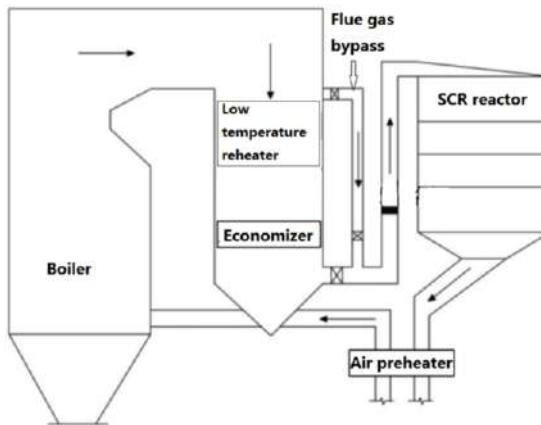


Figure 4. Program C, flue gas bypass transformation.

4. Calculation Results and Comparison of Wide-Load Denitrification Technology Transformation Schemes

Under the 25% rated load condition, the calculation data of the thermal system after the three transformation schemes and before the transformation are compared and analyzed, as shown in Table 4.

Through calculation and analysis, after scheme A water bypass transformation, under 25% rated load condition, in order to make SCR inlet flue gas temperature reach the allowable temperature of denitrification device reactor (above 300°C), the water side bypass flow should reach more than 79% of the total flow. At this time, the flue gas temperature at SCR inlet rises to 301.15°C, the flue gas temperature of boiler rises from 112°C to 120°C, the boiler efficiency decreases from 93.72% to 93.29%, and the working medium temperature at economizer outlet rises to 302°C, reaching the boiling point temperature. After using scheme B to retrofit the staged economizer, under the condition of 25% rated load, the flue gas temperature at SCR inlet is 302.618°C, the flue gas temperature of boiler is 112.832°C, and the boiler efficiency is 93.715%. After retrofitting the flue gas bypass with scheme C, under the condition of 25% rated load, when the bypass flue gas flow is greater than 19%, the flue gas temperature at SCR inlet can be higher than 300°C, which can meet the temperature requirements of denitrification device reactor. At this time, the flue gas temperature at SCR inlet is 300.145°C, the flue gas temperature of boiler is 120.054°C, and the boiler efficiency is 93.3186%.

It can be seen from Table 4 that after adopting the three schemes for transformation, there is no significant difference between the superheated steam outlet temperature and the reheated steam outlet temperature before the transformation, so the steam quality will not be affected. The changes of flue gas temperature and thermal efficiency of three kinds of wide load denitrification transformation schemes and before transformation are shown in Figure 5. After the transformation of scheme A and scheme C, the exhaust gas temperature increased by about 8°C compared with that before the transformation, and the boiler efficiency was lower by about 0.4%. Scheme B exhaust gas temperature and boiler efficiency are basically the same as before the transformation. The results show

Table 4. Comparison of calculation data between the three transformation schemes and the thermal system before transformation.

Item	Before transformation	Plan A	Plan B	Plan C
Working fluid temperature at the outlet of the final reheat器 (°C)	481.03	483.35	480.91	482.36
Working fluid temperature at the outlet of the final superheater (°C)	511.52	514.13	511.56	509.99
SCR inlet flue gas temperature (°C)	279.76	301.15	302.62	300.15
exhaust temperature (°C)	112.73	120.42	112.83	120.05
Boiler efficiency (%)	93.72	93.29	93.72	93.32
Working fluid temperature at the outlet of the economizer (°C)	230.77	310.03	230.67	226.11

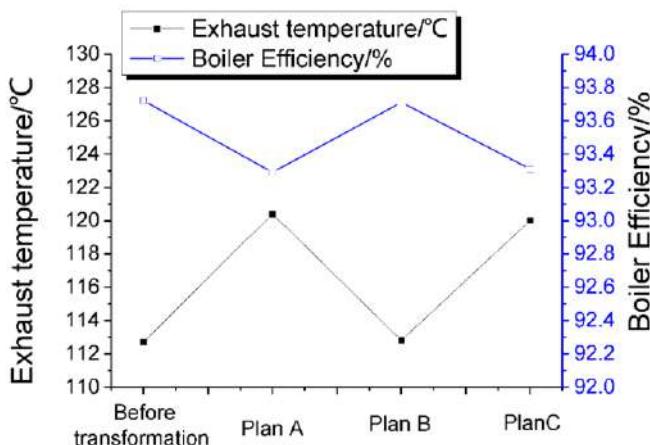


Figure 5. Changes of flue gas temperature and boiler thermal efficiency before and after renovation.

that the effects of the three transformation schemes on the thermal efficiency of the unit are all within an acceptable range, among which option B has almost no effect on the thermal efficiency of the unit.

Considering the safety of the unit, the working pressure of the economizer is 11.6MPa under 25% rated load conditions. After calculation and analysis, the saturation temperature of water under this pressure condition is 310.5°C. Scheme B and Scheme C economizer outlet working fluid temperature is lower than before the transformation, and compared with saturation temperature has a large margin, there is no problem of working fluid boiling in the economizer. The bypass flow of scheme A needs to account for more than 79% of the total flow to be able to increase the temperature of the flue gas at the inlet of the SCR to above 300°C. At this time, the working fluid temperature at the outlet of the economizer is 310.026°C, which is very close to the saturation temperature of the water under this pressure condition. There is not enough under-temperature margin, which can easily cause the working fluid in the economizer to boil, threatening the safe and stable operation of the unit. When the bypass flow is reduced, the temperature of the working fluid in the economizer will decrease, and the temperature of the flue gas at the inlet of the SCR will also decrease. Therefore, plan A cannot meet the requirements of stable operation of the denitrification system under ultra-low load conditions.

In order to ensure the stable operation of the unit under high load conditions, it is necessary to calculate the thermal system under high load conditions for Scheme B. In this paper, 299.3MW operating conditions are selected for calculation, and the results are shown in Table 5.

Table 5. 299.3MW load thermal system calculation results.

Calculation Results	Flue gas outlet temperature	Working fluid outlet temperature
Partition screen superheater (°C)	1163.38	435.682
Rear screen superheater (°C)	1048.71	515.921
Screen reheat (°C)	924.078	446.527
Final reheat (°C)	800.182	542.575
Final superheat (°C)	698.343	546.984
Horizontal low temperature superheater (°C)	461.604	404.706
Primary economizer (°C)	412.652	287.745
Secondary economizer (°C)	396.23	293.699
Rotary air preheater (°C)	144.211	374.381
SCR inlet flue gas temperature (°C)	412.65	-

After adopting plan B for transformation, the SCR reactor is arranged between the primary economizer and the secondary economizer. Under high load conditions, the flue gas temperature at the inlet of the SCR is about 412°C, which is higher than the maximum operating temperature of the reactor of the denitrification unit, which affects the stable operation of the denitrification unit under high load conditions. Therefore, Option B cannot meet the requirement of stable operation of the denitrification system under high load conditions.

In summary, scheme C is the optimal scheme for the flexibility transformation of wide-load denitrification. After the transformation, due to the adjustability of the bypass flue transformation, the boiler efficiency of the unit under high load conditions is not affected, and the boiler efficiency under medium and low load conditions is slightly reduced, up to about 0.4%.

5. Program Evaluation and Precautions

5.1. Case Evaluation

After the equipment modification according to plan C, the high-temperature flue gas bypass modification during the operation of the unit can be adjusted according to the load situation, so it is only necessary to evaluate the tail flue gas flow rate and the pulverizing system under low load conditions.

(1) Assessment of the impact on the operation of the pulverizing system. After calculation, under 25% rated load conditions, the flue gas temperature at the inlet of the air preheater is increased, and the temperature of the hot air at the outlet of the air preheater is increased by 19°C, which is beneficial to increase the drying output of the coal mill.

(2) Estimation of the change of flue gas flow rate. After the transformation, under 25% rated load conditions, 17% of the flue gas needs to be extracted to the bypass flue. At this time, the flue gas flow rate in the main flue is reduced by about 0.2m/s. As the flue gas flow rate is slightly reduced, the flue may be more prone to fouling. Technical measures to prevent fouling of the flue should be considered during the implementation of specific renovations.

(3) Economic evaluation of deep peak shaving. Through wide-load denitrification technology transformation, the unit can stably participate in deep peak shaving. According to the "Notice of North China Energy Regulatory Bureau on Carrying out Simulation Operation of Inner Mongolia Peak Shaving Auxiliary Service Market", assuming that the average daily peak shaving load of the unit is reduced to 35% of the rated load, and the average daily peak shaving time is 4h, based on the Bohai Rim thermal coal price of 570 yuan/ton, the unit's daily peak-shaving power generation cost increases by 909.7 yuan, due to the lack of power generation by peak shaving, the actual profit was reduced by about 840 yuan, and the grid compensation cost was about 21,000 yuan. A rough estimate is that the annual peak shaving income is about 60 million yuan.

5.2. Matters Needing Attention in the Actual Transformation Process

(1) The reliability of the flue gas baffle. The flue gas bypass modification program has higher requirements on the reliability of the flue gas damper. If the sealing performance of the flue gas baffle is poor, or the flue gas baffle cannot be closed after being opened, some high-temperature flue gas leaks from the bypass flue under high load conditions and directly enters the SCR device. There is a risk that the flue gas temperature will be higher than the maximum allowable temperature of the catalyst, which may cause fatal damage to the catalyst. If it is not operated at low load for a long time, that is, the baffle door is in a normally closed state, it may cause dust accumulation, jamming and unable to open. In addition, due to the short flue at the SCR interface, the system is prone to problems with the uniformity of flue temperature mixing, and it is easy to appear that the average flue temperature is appropriate, but the local temperature is higher or lower, which will adversely affect the life of the denitrification catalyst. Therefore, in the actual implementation of the transformation, the reliability of the flue gas damper plays a vital role in the success of the transformation.

(2) The problem of flue gas temperature deviation. After the boiler economizer is installed with a flue gas bypass, the extracted high-temperature flue gas and the low-temperature flue gas flowing through the economizer may be insufficiently mixed, resulting in uneven flue gas flow and temperature fields, and large flue temperature deviations. As a result, the local flue gas temperature in the low-load denitrification zone of the unit is lower than 300°C, which deviates from the optimal reaction temperature of the denitrification catalyst, reduces the catalyst activity, reduces the denitrification efficiency, and increases the ammonia escape rate. Therefore, in the process of reforming the flue gas bypass scheme, attention should be paid to the uniformity of the flue gas field. A deflector can be installed in the flue to make the high-temperature flue gas and the low-temperature flue gas fully mixed.

(3) The pressure difference of the air preheater increases. If the problems of flue gas flow field and temperature field are not fully considered during the transformation, a large deviation of the flue gas temperature may occur after the transformation, which leads to an increase in the ammonia escape rate of the denitrification system. Unreacted

liquid ammonia reacts with SO₃ and water generated during the combustion of the boiler to form NH₄HSO₄. NH₄HSO₄ has a dew point of 147°C, which is concentrated on the surface of an object in the form of liquid or dispersed in the flue gas in the form of droplets. Liquid NH₄HSO₄ is a very corrosive and viscous substance, which easily adheres to fly ash in the flue gas. As the temperature of the cold end of the air preheater is lower than 147°C, the liquid NH₄HSO₄ will adhere to the heating surface of the air preheater, which will not only cause corrosion to the cold end components, but also further adhere to the fly ash, which can not be removed by soot blowing through the air preheater. As a result, the pressure difference of the air preheater gradually increases, which may eventually cause the unit to have to be shut down for flushing or furnace shutdown on one side, which seriously threatens the safe and stable operation of the unit. In order to avoid the problem of increasing the pressure difference of the air preheater, the problem of flue gas mixing should be fully considered during the transformation process [5].

6. Conclusions

In order to solve the problem of low SCR inlet flue gas temperature during the operation of a 300MW unit, the unit has deep peak regulation capability and can stably put the SCR denitrification device into operation under ultra-low load conditions. Three solutions are proposed, and the thermal system calculation and comparison analysis are carried out on the three solutions respectively. Finally, the program C flue gas bypass modification program is selected for wide-load denitrification modification. After the transformation, under ultra-low load conditions, the inlet flue gas temperature of the SCR can be increased to over 300°C through bypass flue adjustment, and at the same time, the safety and economy of the unit under high load conditions will not be affected. In addition, after the transformation of Plan C, the possible impact of the daily production and operation of the unit and the economic benefits of participating in deep peak shaving were carefully evaluated, and the key issues that need to be paid attention to during the transformation were discussed.

Acknowledgements

This research was financially supported by the Inner Mongolia Electric Science Research Institute 2020 youth science and technology project NO. 510241200018.

References

- [1] X W Zhou, Q J Hua, A Wang, W Xu, H Chen. Study on optimization and transformation technology of SCR system in deep peak-shaving boiler for 1000MW units [J]. Power System Engineering, 2020, 36 (1): 57-60.
- [2] L L Liu, Y Chen, X F Li, M Kuang, Z H Wang, J H Zhou and K F Cen. Thermodynamic calculation of various heating surface adjustments of a 600 MW boiler: Used to reduce the exhaust gas temperature [J]. Energy Engineering, 2015, (5): 57-62.
- [3] L L Liu, X F Li, Y Chen, M Kuang, Z H Wang, J H Zhou, Cen K F Cen. Thermodynamic calculation for the tail heating surface optimization used to reduce the exhaust gas temperature of a 600MW sub-critical boiler [J]. Boiler Technology, 2017, 48 (1): 1-5.

- [4] Q Miao. Progress and prospects of denitrification technology [J]. Clean Coal Technology, 2017, 23 (2): 12-19.
- [5] X P Zheng, S X Li, J Guo, Z L Liu, D H Yuan. Analysis and treatment of air pre-heater clogging in CFB boiler [J]. Inner Mongolia Electric Power, 2019, 37 (5):29-32.

Research on Identification of Power Grid Weakness Based on Bayesian Inference

Jiang HU^{a1}, Wei LI^b, Wenxia LIU^a, Xianggang HE^a, Yu ZHANG^a

^a Grid Planning & Research Center, Guizhou Power Grid Co., Ltd., Guiyang, Guizhou, 550003, China

^b Guiyang Power Supply Bureau, Guizhou Power Grid Co., Ltd., Guiyang, Guizhou, 550003, China

Abstract. With the gradual reform and development of the power grid, it is of great significance to study how to effectively identify and evaluate the weak links of the power grid for the actual planning, construction, and operation of the power grid. This paper analyzed the power grid's historical component data and real-time operation state parameters. We established a weak link identification model based on Bayesian reasoning. Firstly, we constructed the node branch Bayesian network according to the network topology relationship. The power transmission distribution factor is modified according to the historical operation load of the grid components, and the conditional probability table is calculated based on the grid structure; finally, we used the maximum possible explanation algorithm in the Bayesian network. The weakness degree of all components in the network is calculated, and the maximum probability weak link sequence is obtained. The correctness and effectiveness of the proposed method are verified by IEEE 39 bus simulation and regional power grid data.

Keywords. power grid weakness, bayesian inference, transmission distribution factor

1. Introduction

With the increase of the scale and structure complexity of large power grid, the fault risk coefficient and safety index are further improved. The Texas blackout in February 2021 in the United States triggered a heated discussion on power grid security issues [1]. Quickly identifying the weak links in the power grid has become an important problem in operation planning. In the existing research, the identification and estimation of weak links in the power grid mainly focus on nodes and branches. The main methods include grid structure index, state reliability index, probability, and data-driven.

At present, most of the research methods based on grid structure index are oriented to line vulnerability estimation. In reference [2], the electrical efficiency is calculated based on the grid admittance matrix to evaluate the vulnerability according to the most influential power path transmission rule. However, the model assumption fails to fit the actual grid transmission situation. In reference [3], the power supply reliability index is calculated through grid structure analysis. Identify the weak links in the power grid; Literature [4,5] proposes using the maximum power flow to analyze the topology of

¹Corresponding author: Jiang Hu, Grid Planning & Research Center, Guizhou Power Grid Co., Ltd., Guiyang, Guizhou, 550003, China; E-mail: Hujiang9409@163.com.

complex networks and calculate the vulnerability of transmission lines. Although it can make up for the above problems to a certain extent, it still fails to simulate the actual power grid operation. Based on this, literature [6] evaluates the vulnerability of transmission networks from time series fault propagation based on the cascaded Fault Graph (CFG) of fault chain theory. About [7], starting from the idea of network construction, this paper analyzes the weak links existing in the actual regional power grid. Still, the universality and adaptability of this method are difficult to meet the demand. According to the reliability indexes such as transmission margin, power transmission distribution factor (PTDF), and electrical betweenness, the existing research uses power flow calculation, static safety check, and short-circuit current analysis to analyze and evaluate, this kind of method is more universal. Reference [8] calculates the value at risk and average excess loss based on the comprehensive outage probability model, The weak link of a thorough evaluation of line fault; In reference [9], a condition evaluation index system is established based on the principal component analysis method to identify the weak links in power grid operation from two aspects of weight and margin space; In reference [10], the power transmission distribution factor is used to define the network efficiency, which reflects the impact of faults on the system structure and the degree of state influence; Reference [11] establishes the contribution of each component to the system under typical operation mode based on reliability index. However, with the further increase of network complexity, the above methods' computing power and network adaptability will be restricted. With the breakthrough of artificial intelligence algorithms in recent years, the method based on data-driven and probability analysis has been put forward. The paper [12,13] proposes an improved PageRank algorithm to analyze and calculate the correlation and scalability of the network to realize the rapid identification of weak lines; Reference [14] Based on the high-dimensional random matrix theory, analyzes and calculates the node fluctuation after the network receives the disturbance, and realizes the mining and processing of the system voltage data.

According to the existing research, it shows that there are two main difficulties in the identification of power grid weak links. One is that it can match the complex structure and data scale of a large power grid, and the other is that it can be applied to the actual power grid. To meet these two points, we proposed a weak link identification model based on Bayesian reasoning for the first time, which reduces the modeling error of traditional algorithm through the data-driven method. This method combines actual historical data to find out the weak links of the actual power grid. First, the historical operating status is integrated into the real-time operating status of the power grid, and the distribution probability table of node branches is established. And then realize the correlation analysis between offline data and real-time information. We constructed a Bayesian network model based on different failure modes, and the reasoning algorithm solved the vulnerability of all fulcrums in the network. Finally, we got a complete set of weak link sequences. It can meet the requirements of high efficiency, accuracy, and adaptability of weak link identification in an actual power grid.

2. Mathematical Model of Bayesian Inference

Bayesian network reasoning is a kind of correlation probability prediction method based on the given state distribution space, which uses a knowledge map to realize the reasoning of the given state results. Common Bayesian reasoning includes posterior

probability problem, maximum a posterior hypothesis (MPH), and maximum possible explanation (MPE). In this paper, the mathematical model of weak link identification and evaluation is established based on MPE.

2.1. Bayesian Network

Bayesian network is a kind of directed acyclic probability graph model, which can deal with many uncertain factors in the power grid. Assuming that there are random variables in the network, the causal relationship is determined by the directed pointer between nodes. Each node is relatively independent and has a conditional probability distribution, then the joint probability distribution of the network is as follows:

$$\begin{aligned} P(X_1, \dots, X_n) &= P(X_1)P(X_2|X_1)\dots P(X_n|X_1, X_2, \dots, X_{n-1}) \\ &= \prod_{i=1}^n P(X_i|X_1, X_2, \dots, X_{i-1}) \end{aligned} \quad (1)$$

In this paper, based on the historical eigenvalues of branches and nodes of power grid, the prior probability distributions of sub-level and sub-level of Bayesian network are established respectively, and the top-down relationship is directed correlation.

$P(X_m|X_n)$ is the conditional probability value, which maps the network association relationship, and then uses the network topology relationship to establish the corresponding directed pointer graph. The conditional probability distribution is established by calculating the structure association degree. Considering the power flow distribution in power grid operation, the Bayesian network model is constructed, as shown in Figure 1.

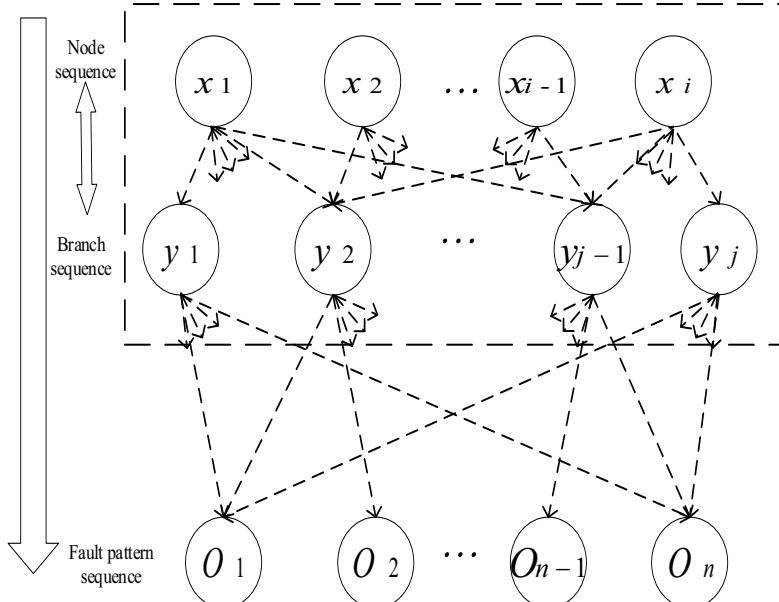


Figure 1. Bayesian network structure chart.

In the network, the nodes and lines are coupled with each other. The fault events are used to attack the node line sequence network, and the probability distribution of the network is obtained.

2.2. Maximum Probable Explanation Algorithm

In Bayesian networks, Finding the maximum a posteriori probability hypothesis \bar{h} with known evidence set E , The explanation of the maximum probability of evidence sets makes all variables in the network consistent with the corresponding state set, that is, the maximum probable explanation (MPE).

$$\begin{aligned} \arg \max_h P(H|E=e) &= \arg \max_h \frac{P(H|E=e)}{\sum_h P(H|E=e)} \\ &= \arg \max_h P(H, E=e) \end{aligned} \quad (2)$$

In this paper, the weakness degree of all branches and nodes is taken as a hypothetical variable to form a state set after a fault event. After reasoning, the weakest rate set with maximum probability is obtained, and then the Viterbi algorithm is used to solve the problem. The fault event O_k is introduced in turn to calculate the maximum probability of all single branch node paths $(x_i - y_j)$ in the network, and finally the weak link is obtained through the optimal path backtracking.

- Initialize the probability value $\delta_1(i) = \pi b_i(o_1)$ of node branch weak state sequence and the parameter $\Psi_1(i) = 0$. And then it passes by $t = 2, 3, \dots, T$.

$$\begin{aligned} \delta_t(i) &= \max_{1 \leq j \leq N} [\delta_{t-1}(j) a_{ji}] b_i(o_t) \\ \Psi_t(i) &= \arg \max_{1 \leq j \leq N} [\delta_{t-1}(j) a_{ji}] \end{aligned} \quad (3)$$

$\delta_t(i)$ is the maximum probability in the fault sequence with state I at time t ; N is the number of states.

- Solving the maximum state probability value P^* .

$$\begin{aligned} P^* &= \max_{1 \leq i \leq N} \delta_t(i) \\ i^* &= \arg \max_{1 \leq j \leq N} [\delta_t(i)] \end{aligned} \quad (4)$$

i^* is the weak sequence represented by the corresponding probability value.

- The optimal path is derived, that is, the weak sequence of branch x and node y with the maximum probability.

$$I^* = (x_1^*, x_2^*, \dots, x_i^*, y_1^*, y_2^*, \dots, y_j^*) \quad (5)$$

3. Weak Link Identification and Evaluation Model

3.1. Initial Probability Distribution Based on Improved PTDF

The power transmission distribution factor (PTDF) represents the change of power transmission from source node to load node in power flow. When the injection power ΔP_s of the power node in the power grid changes, the power change ΔP_{ij} in the corresponding power load branch is the unit power injected by the nodes at both ends and the unit power value of the reference node remains unchanged, and the power change of the branch is calculated.

$$\Delta P_{ij} = \frac{X_{ig} - X_{il} - X_{jg} + X_{jl}}{X_{ij}} \Delta P_s = \lambda_{ij} \Delta P_s \quad (6)$$

In the above formula, X_{ig} X_{il} X_{jg} X_{jl} represents the impedance values from branch nodes i and j to power load nodes s and t respectively. λ_{ij} is the power transmission distribution factor of unit power between power load nodes in branch, and its value represents the contribution of branch to power flow transmission.

Considering that the maximum transmission capacity of different lines is inconsistent, in order to ensure a high degree of matching with the power grid operation, this paper will use the historical load data to sample and extract, and improve the initial node branch probability distribution based on PTDF.

- Firstly, the sequence of sub nodes and intermediate nodes in Bayesian network is divided according to the topological relationship of power grid, and the equipment load rate is calculated by the load of historical lines and stations.

$$\begin{aligned} \lambda_l &= \frac{Q_{\Sigma l}}{P_l^* \times T_l} \times 100\% \\ \lambda_t &= \frac{Q_{\Sigma t}}{S_t \times T_t} \times 100\% \end{aligned} \quad (7)$$

In equation (7), the line load rate λ_l is the ratio of the product of the annual total power transmitted by the line $Q_{\Sigma l}$, the economic power transmitted by the line P_l^* and the annual utilization hours T_l ; The load rate of the station is λ_t , which is the ratio of the annual power transmission $Q_{\Sigma t}$ of the main transformer to the product of the rated capacity S_t and the annual utilization hours T_t of the station.

- Furthermore, it is determined that the network structure characteristics include single power line rate of substation, single variable rate of substation, connection rate of intermediate line, connection rate of intermediate line between stations, etc. the initial

probability distribution $P(x_i)$ and $P(y_j)$ of nodes and branches are calculated respectively through formula (8) and formula (9).

$$P(x_i) = \lambda_i * \theta_i \quad (8)$$

$$P(y_j) = \lambda_j * \lambda_{ij} * \theta_j \quad (9)$$

- The conditional probability distribution of node branch is mainly determined by the carrying capacity of the branch.

$$P(y_j|x_i) = \frac{X_{ij}}{2 \times \sum_n X_{in}} \times 100\% \quad (10)$$

$$\sum_n X_{in}$$

In equation (10), $\sum_n X_{in}$ is the sum of all branch impedances associated with the node. The heavier the proportion is, the higher the vulnerability of the node is associated with the branch.

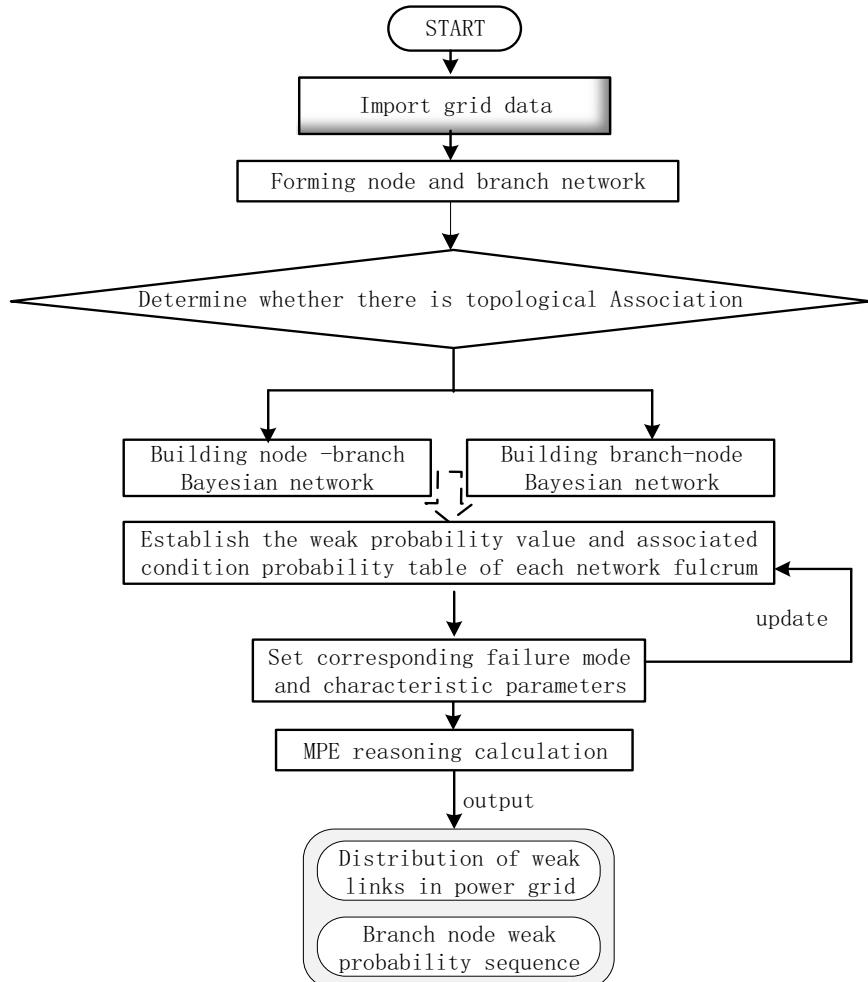
3.2. Fault Mode and Feature Extraction

According to the power system security and stability guidelines and related standards, in order to meet the operation reliability and planning directivity of power grid, this paper proposes to establish the node bottom layer based on N-1 fault mode for the middle layer nodes. At the same time, according to the security risk of power grid operation, it includes complex fault conditions such as switch refusing to operate, in particular, For the double circuit lines with high risk coefficient in the actual power grid operation, the fault mode based on N-2 is established. Considering that the frequency of cascading failure events in actual operation is low, and the alarm information is uncertain, the probability of protection switch refusing to operate and maloperation is low, the corresponding occurrence probability is set in reference [15] for 220kV and above voltage level AC system relay protection equipment and its operation, for example, the initial probability value of switch and protection refusing to operate fault mode is set as 0.01%, The initial probability of malfunction mode is set as 0.06%, and the conditional probability is set according to the attributes of the upper nodes. Table 1 lists the corresponding situation of conditional probability distribution of underlying failure modes in Bayesian network. $P(O_i)$ is the probability of different failure modes. $P(O_i|y_j)$ is the probability of O_i failure in the existing bearing capacity of node y_j .

Table 1. Failure mode probability distribution.

Failure mode	Failure probability	Conditional probability
N-1	$P(O_1)$	$P(O_1 x_i) P(O_1 y_j)$
N-2	$P(O_2)$	$P(O_2 x_i)$
Considering switch and protection failure mode	$P(O_3)$	$P(O_3 x_i) P(O_3 y_j)$
Considering the Malfunction Mode of switch and protection	$P(O_4)$	$P(O_4 x_i) P(O_4 y_j)$
Considering alarm information false alarm and missing alarm	$P(O_5)$	$P(O_5 x_i) P(O_5 y_j)$

3.3. Identification Algorithm and Probability Estimation

**Figure 2.** Bayesian reasoning flow chart.

By constructing the Bayesian network and the distribution table of conditional probability between the connected nodes, the maximum possible interpretation MPE algorithm is used for reasoning to further identify the weak nodes in the network. When

the corresponding failure mode is set, the bifurcation attack will be launched to any fulcrum in the original node branch network. The original network will extend to the sub-nodes for failure events. When multiple failures or cascading failures occur, the fulcrum in the network forms a corresponding logical relationship with the failure events, and the distribution probability in the Bayesian network is updated based on the failure event probability. The vulnerability of each fulcrum is calculated by reasoning algorithm. The process of weak link identification and evaluation inference model is shown in Figure 2.

4. Simulations

This paper mainly uses two different examples to analyze. One is the traditional IEEE39 node system, and the other is the actual regional grid example. The former is to better compare the existing methods, And the latter is to verify the feasibility and accuracy of the virtual grid. The simulation platform includes MATLAB and open2000. Open2000 is a power grid dispatching operation software mainly used for actual power grid production and dispatching.

4.1. IEEE 39 Bus Simulation Example

In this paper, the IEEE39 bus system is selected for simulation, and the vulnerability of each node branch in the network is tested through different failure modes. The topology is shown in Figure 3 and the results are shown in Table 2.

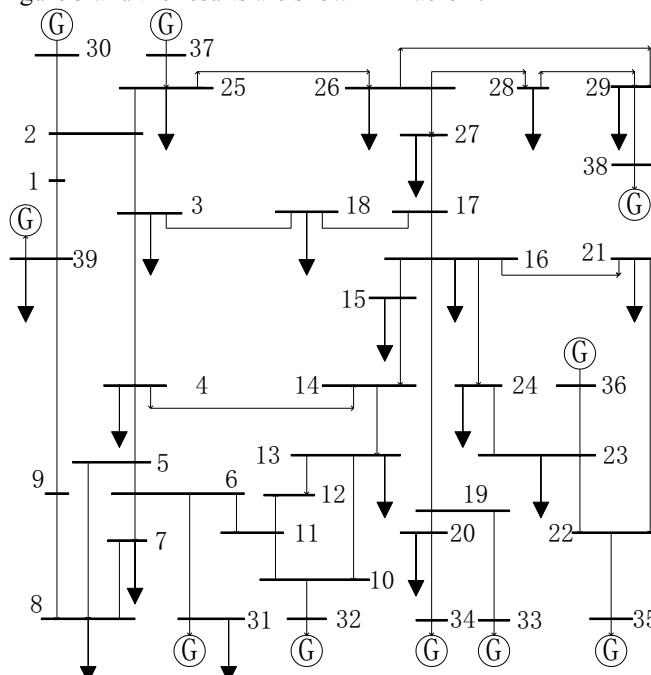


Figure 3. IEEE39 node network topology.

Table 2. Vulnerability results of different failure modes.

Branch vulnerability \ Failure mode	$\overline{\pi}_{y13-14}$	$\overline{\pi}_{y14-15}$	$\overline{\pi}_{y16-17}$
O_{y13}	0.670	0.601	0.498
O_{y14}	0.141	0.721	0.364
O_{y16}	0.141	0.546	0.697
$O_{y13} O_{y14}$	0.658	0.711	0.364
$O_{y13} O_{y16}$	0.671	0.597	0.697
$O_{y14} O_{y16}$	0.141	0.698	0.746

In Table 2, we list six failure modes, including single node failure (O_{y13} O_{y14} O_{y16}) and double node failure ($O_{y13} O_{y14}$ $O_{y13} O_{y16}$ $O_{y14} O_{y16}$). When different faults occur, the electrical distance of the same branch vulnerability $\overline{\pi}_{y13-14}$, $\overline{\pi}_{y14-15}$ and $\overline{\pi}_{y16-17}$ is greatly affected by the attack point. When the fault scope affects the branch, the branch vulnerability is higher, and its value is less different from that of the single fault mode of the branch. Therefore, the vulnerability solved by this method mainly depends on the network structure and operation state, it is less affected by the type of fault.

According to the Bayesian MPE reasoning results, the branches with the highest weakness are shown in Table 3, and the key line identification results based on the maximum flow system in reference [4] are selected as the comparative analysis. From the comparison in Table 3, we can see that both methods can effectively identify weak lines, mainly including the important power transmission lines 16-17, 15-16, 2-3, etc. These lines are the important connecting lines between the unit and load, which restrict the transmission capacity of the unit.

Table 3. Ranking of branch vulnerability in IEEE39.

Order	Proposed method	Reference [4]	Branch vulnerability
1	16-17	16-17	0.721
2	15-16	2-3	0.697
3	2-3	14-15	0.670
4	16-19	15-16	0.597
5	14-15	16-19	0.593
6	1-2	4-14	0.532

Compared with the literature [4], this paper calculates that the 1-2, 2-25 lines with higher weakness are the transmission lines of units, which bear large power transmission in the system structure and restrict the stable operation of generator units 30 and 37. Therefore, compared with the current grid structure index research methods, this paper pays more attention to the operation distribution of the actual power flow, which is more critical in the actual power grid planning; In addition, line 13-14 is calculated as the key weak branch by Bayesian inference algorithm. Combined with simulation analysis, it can be seen that when the line fails, the transmission of 32 units will be greatly limited, and the transmission pressure of adjacent 4-5 lines will increase, so this method can better identify the weak lines in the power grid. The Bayesian inference model proposed in this paper is essentially a data-driven and probability analysis method, which combines the power grid operation state and topology structure to establish the model. It is less dependent on the physical model and can effectively cover all the nodes and branches in

the network. The model has strong expansibility and can better adapt to the planning and operation needs of the actual power grid.

4.2. A Simulation Example of a Regional Power Grid

In order to fit the practical application, this paper identifies the weak links of the primary grid in a certain area based on the method proposed in this paper. The regional grid is included in the Bayesian model, and the initial probability distribution of the model is calculated based on the historical utilization efficiency of substations and lines. All components in the model are calculated by the ergodic method. According to the MPE algorithm, we figured the weakness of each node in the Bayesian network.

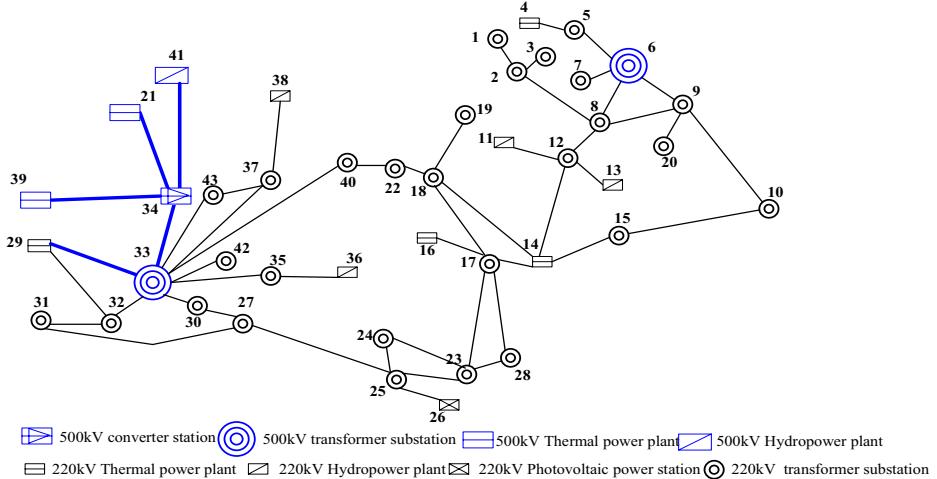


Figure 4. Regional power grid topology.

According to the further analysis of the calculation results, lines 6-8 with the highest vulnerability are the necessary network access of 500kV substation 6 in this area, and the load rate is high. Once the fault occurs, the 500kV power transmission channel and surrounding lines in this area will face the risk of overload. Line 1-2 is an essential connecting channel between the regional power grid and the external power grid, and node 1 transmits a large amount of new energy power to the region. Lines 22-40, 18-22, 23-24 and 24-25 are the lines that are easy to cause neck jamming in the regional power grid. Lines 17-23, 16-17 and 14-18 are the outgoing channels of large units 14 and 16. In case of failure, we can not guarantee the power plant's output, and the system will lose many loads. Therefore, the importance of these lines is determined by the structure and operation.

Figure 4 shows a simplified topology of 220kV and above power grid in a certain area. The power grid in this area belongs to the site with rich power sources, connecting multiple essential transmission channels. When the weak links in the power grid are not strengthened in time, or the grid structure is not improved, it will lead to severe machine and load shedding. Therefore, it is of great significance for the safe and stable operation of the whole power grid to identify the weak links that need to be solved in the actual process. In this paper, considering that the primary type of power supply in this area is thermal power and the winter load is large, the system data under dry and large mode is

selected for analysis and calculation. According to the calculation results, the top ten weak lines are selected and listed in Table 4.

Table 4. Ranking of branch vulnerability in regional power grid.

Order	Proposed method	Branch vulnerability
1	16-17	0.721
2	15-16	0.697
3	2-3	0.670
4	16-19	0.597
5	14-15	0.593
6	1-2	0.532
7	13-14	0.512
8	3-4	0.482
9	16-24	0.441
10	2-25	0.417

According to the further analysis of the calculation results, lines 6-8 with the highest vulnerability are the necessary network access of 500kV substation 6 in this area, and the load rate is high. Once the fault occurs, the 500kV power transmission channel and surrounding lines in this area will face the risk of overload, Line 1-2 is an essential connecting channel between the regional power grid and the external power grid, and node 1 transmits a large amount of new energy power to the region, Lines 22-40, 18-22, 23-24 and 24-25 are the lines that are easy to cause neck jamming in the regional power grid, Lines 17-23, 16-17 and 14-18 are the outgoing channels of large units 14 and 16. In case of failure, we can not guarantee the power plant's output, and the system will lose many loads. Therefore, the importance of these lines is determined by the structure and operation.

5. Conclusions

This paper proposed a Bayesian inference model to identify the weak links in the power grid. We fused the historical load of components in the power grid with the actual operation data. The distribution probability table considering the power transmission distribution factor is constructed by using the data-driven method, and the node branch Bayesian network is established through the power flow topology and grid structure. According to different failure modes, the network is attacked by a bifurcation attack, and the distribution probability table is updated. Finally, we calculated the vulnerability of all branches by the MPE algorithm. The simulation results verify the effectiveness and adaptability of the method.

With the large-scale new energy access to the power grid in the future, it will further aggravate the limited power flow in the local power grid. It will affect the fluctuation of voltage, current, and frequency of the power grid to a certain extent. Identifying the weak link of the power grid by combining the node voltage data and operation telemetry information will be the research focus of the follow-up work.

Acknowledgements

This research was financially supported by the Guizhou Power Grid Co. under Grant of GZKJXM20210369.

References

- [1] X M An, H D Sun, X H Zhang, et al. Analysis and lessons of texas power outage, in Proceedings of the CSEE. 2021, 41(10):3407-3415.
- [2] G Chen, Z Y Dong, D J Hill, et al. An improved model for structural vulnerability analysis of power networks [J]. *Physica A*, 2009, 388: 4259-4266.
- [3] J K Fang, C Su, Z Chen. Power system structural vulnerability assessment based on an improved maximum flow approach [J]. *IEEE Transactions on Smart Grid*, 2018, 9(2):777-785.
- [4] C Y Tao. Weak link identification of electricity gas integrated energy system based on timed Petri net [J]. *Electrical Measurement & Instrumentation*, 2021, 53(10):39-48.
- [5] X G Wei, JB Zhao, T Huang. A novel cascading faults graph-based transmission network vulnerability assessment method [J]. *IEEE Transactions on Power Systems*, 2018, 33(3):2995-3000.
- [6] L X Hu, W M Zhen, L Wang, et al. Power supply reliability assessment based on distribution network topology matrix [J]. *Power System and Clean Energy*, 2019, 35(2): 44-51.
- [7] S Zhao, W Qian, B R Cao. Analysis of weak links in 220 kV and above power grid [J]. *Yunnan Electric Power*, 2017, 49:98-101.
- [8] Q Yu, X Luan, Q He, et al. Evaluation of power grid weak links on the basis of comprehensive fault probability model [J]. *Science Technology and Engineering*, 2016, 16(30): 92-97.
- [9] E Bompard, E Pons, D Wu. Extended topological metrics for the analysis of power grid vulnerability [J]. *IEEE Systems Journal*, 2012, 6(3): 481-487.
- [10] R W Wang, Y G Zhang, Y Yang, et al. Vulnerable line identification of complex power grid based on electrical betweenness [J]. *Power Systems Protection and Control*, 2014, 42(20):1-6.
- [11] Y Zhen, K G Xie, H B Li, et al. Application of the weak part identification based on reliability tracking method to reliability improvement of provincial power network [J]. *Electrical Measurement & Instrumentation*, 2015, 52(6):118-123.
- [12] Ma Z Y, Liu F, Shen C, et al. rapid identification of vulnerable lines in power grid using modified pagerank algorithm-Part I: Theoretical foundation [J]. *Proceedings of the CSEE*, 2016, 36(23):6363-6370.
- [13] Z Y Ma, F Liu, C Shen, et al. rapid identification of vulnerable lines in power grid using modified pagerank algorithm-Part II: Factors affecting identification results [J]. *Proceedings of the CSEE*, 2017, 37(1):36-44.
- [14] B Wang, J L Wang, D C Liu, S Y Chen. Research on evaluating vulnerability of power network based on high-dimensional random matrix theory [J]. *Proceedings of the CSEE*, 2019, 39(6):1682-1691.
- [15] L Zhang, P F Lv, H Shen, et al. Analysis on protective relaying and its operation conditions of SGCC in 220kV and above voltage AV system in 2013 [J]. *Power System Technology*, 2015, 39(4):1153-1159.

Discussion on Integrated Modeling Technology of New Generation Centralized Control Station Equipment Monitoring System

Yi ZHAI¹, Tianpin XU, Qijing YANG, Weihua CAO, Zhichao QI, Chunyao LIU
Nari Group Corporation (State Grid Electric Power Research Institute), Nanjing Jiangsu 210061, China

Abstract. At present, the substation SCD model mainly aims at the modeling of primary equipment and secondary equipment, and the description of auxiliary equipment and substation area model is missing. According to the characteristics of equipment monitoring of the new generation centralized control station, the description of auxiliary equipment model and area model needs to be added to the original IEC61850 standard. In the first mock exam station, the current business requirements of centralized control station are analyzed. A unified model structure of master station is designed. Based on the definition of IEC61850 standard, a method and example of extending auxiliary device model and regional model are given. The proposed auxiliary equipment and area modeling method based on IEC61850 standard can make up for the shortcomings of the original standard, support the equipment monitoring of the new generation centralized control station and realize rich application functions.

Keywords. IEC61850, SCD, auxiliary equipment, substation area model

1. Introduction

At present, the domestic regional dispatching operation mode has been mature and stable, but the operation and management mode for substation monitoring has not been finalized. Different regions have adopted different construction methods due to the different degree of economic and power grid development. At present, there are three main modes of substation monitoring [1]: 1. Substation substation monitoring. 2. The mode of centralized control station adopts the slice monitoring. 3. Centralized control center mode: several operation and maintenance stations with few people on duty are set up

¹Corresponding author: Yi Zhai, Nari Group Corporation (State Grid Electric Power Research Institute), Nanjing Jiangsu 210061, China; E-mail: vitty2005@163.com.

according to the operation radius. One centralized control center can monitor hundreds of substations. At present, this mode can be subdivided into two modes in China: 1. Separation of dispatching and centralized control. 2. Integration of dispatching and centralized control. At present, 41.1% of substations in state grid system adopt traditional mode and centralized control station mode. Primary equipment and signal information monitored by the master station are actually modeled in the substation. The master station belongs to repetitive modeling, and does not follow the principle of "source maintenance, resource sharing".

In 2020, the key tasks of the State Grid Corporation are to implement the equipment ownership system, optimize the operation and maintenance mode of the substation according to local conditions, and implement centralized monitoring of the substation. The State Grid equipment department actively implements the company's work requirements, carries out the optimization of substation operation and maintenance monitoring mode, speeds up the construction of substation centralized control station, and improves the main and auxiliary equipment monitoring and other technical support means. Combined with the overall scheme of the new generation of independent and controllable substation secondary system, fully inheriting the construction experience and achievements of the existing dispatching system and auxiliary control system, adhering to the "problem oriented, demand-oriented and goal oriented", this paper puts forward the idea of a new generation of centralized control station monitoring system with main and auxiliary functions.

In summary, these requirements make the data communication between the main and sub stations break the point-to-point forwarding mode, and adopt the main plant station interaction technology which is more flexible, efficient and compatible with the existing communication system, Through the bottleneck of interaction between the two ends, the main station automation system and substation automation system are closely combined from the dimension of application function to realize interconnection of the automation system of the main station [2-5]. Since the promulgation and implementation of q/gdw679 technical specification for the construction of intelligent substation integrated monitoring system [6], it is proposed that the integrated monitoring system of substation integrates all the subsystems and equipment in the station.

In reference [7-9], seamless communication integrated modeling of main substation and substation and centralized monitoring function design are proposed.

Reference [10] proposed a method to expand the protection information system based on IEC61850 standard, but did not mention how to expand the auxiliary equipment model and substation area model

The purpose of this paper is to discuss the unified modeling of main and auxiliary equipment, and then put forward the idea of how to expand the auxiliary equipment model on the basis of the original SCD standard.

2. Requirement Analysis of Integrated Modeling

2.1. Model Requirements

For the traditional centralized control and monitoring center, due to the new equipment and new technology of the intelligent substation, the data requirements of the main station system for the substation have new changes, which are no longer limited to the traditional four remote information, but need the operation status information of the

substation system and equipment (such as the primary and secondary equipment status monitoring information, equipment account and configuration information, model, graphics, protection recording, data processing, Fault recording, network message, auxiliary control equipment and other file information).

Taking the 500kV Smart Substation in Zhejiang Province as an example, the "regional dispatching information forwarding table" is counted, including 1998 remote signaling and 381 telemetry.

In addition to the above statistics, it also analyzes the specific content of the data sent to the main station by the intelligent substation. The new data are mainly reflected in the following aspects: intelligent terminal, general object-oriented substation event (goose) / sampling value (SV) link signal of merging unit, network switch alarm information, network switch alarm information, and so on at the same time, due to the different business needs of regional and provincial dispatching, many new sending signals are sent in the way of combined point number.

From the above analysis, it can be seen that intelligent substation not only increases the amount of data, but also adds a lot of new types of data. The increase of data is both an opportunity and a challenge for the data analysis application function of main station and substation. On the one hand, multi types of data can provide rich basic data for analysis and calculation applications, Enhance the professional technical level of advanced application function; On the other hand, due to the deployment of a large number of IED in Smart Substation, in the development trend of unattended substation, the operation status of primary and secondary equipment and monitoring system in the station needs convenient remote monitoring and analysis means, but at present most of the data can not be sent, and is stranded in the station, which brings hidden dangers and risks to the reliable monitoring of equipment operation status.

2.2. Business Requirements

In terms of business functions of dispatching master station and substation, through the pilot application and construction of intelligent substation, the integrated monitoring system integrates data acquisition and monitoring (SCADA), equipment condition monitoring, protection substation, auxiliary system monitoring and other subsystems, and the system complexity also increases. At present, the dispatching system or centralized monitoring model is manually modeled by the master station. At present, there is no auxiliary device model in the SCD file of the smart substation, and there is no description of such equipment in the IEC61850 specification. Therefore, an example of the figure 1, it is necessary to extend the IEC61850 standard to increase the description of the auxiliary device.

By standardizing the substation model system, integrating the information model of substation, centralized control station and business middle station, and insisting on one source of substation professional data, the information fusion and connection of automation system and information system can be realized.

The first mock exam is based on the unified model. The integration of the primary equipment model, the two equipment models, the auxiliary equipment model, the equipment ledger, and the physical ID is realized in the substation.

Data connection: with the physical ID and equipment account as the link, the connection and transformation between the equipment model and the business platform model are completed in the centralized control station, so as to realize the integration and

connection between the main and auxiliary equipment monitoring system of the substation and the business platform data.

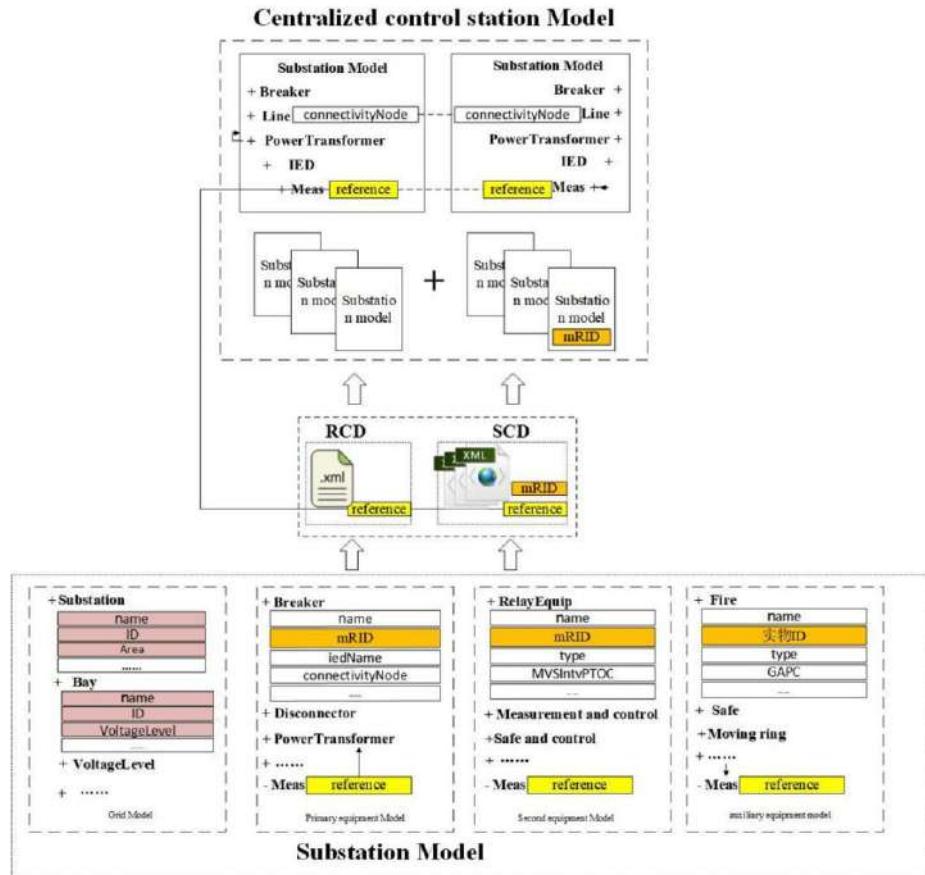


Figure 1. The first mock exam of substation unified model architecture.

3. The Key Technology

3.1. Geographical Area Model

The typical area division of substation is usually based on the principle of relatively independent physical area distribution. The typical area division is as follows: the surrounding wall is divided into five areas, and the wall with gate is divided into two areas according to the left and right wall of gate. According to the voltage level and equipment type, the primary equipment outdoor field can be divided into transformer outdoor field, 500kV outdoor field, 220kV outdoor field or 110kV outdoor field. Indoor primary switch gear, secondary equipment, station equipment, main control equipment, communication equipment and storage battery can be divided into 500kV protection room, 220kV protection room, 35kV high voltage room, station power distribution room, production complex building or main control room according to indoor relatively independent physical area.

If a region is divided into several independent sub regions, the region can be divided into several sub regions. The sub region division is usually based on the principle of relatively independent physical region distribution. For example, the area of production complex building can include main control room, communication room, battery room, production preparation room, corridor and other sub areas. If an area is divided into floors, the identification and naming of each sub area should be able to distinguish the floors.

3.2. Auxiliary Equipment Model

Substation auxiliary equipment model should include fire protection, security, dynamic environment, video monitoring, online inspection and other equipment. IEC61850 describes primary equipment by conducting equipment class, and distinguishes specific equipment by attribute type. As shown in table 1, "CMR" represents air conditioner.

Table 1. Classification table of auxiliary equipment.

Classification	Auxiliary Equipment	Abbreviation
	Micro weather sensor	MCM
	Temperature sensor	TMP
	humidity sensor	HUM
	Water immersion sensor	WIM
	Water level sensor	WLL
	SF6 monitoring sensor	QSF
Moving Ring Equipment		
	Water pump	WPM
	fan	FMR
	air conditioner	CMR
	dehumidifier	DHR
	SF6 fan	SFF
	floodlight	LED
	Lighting control terminal	LCT

The auxiliary control equipment is described by extending the device class, and the specific equipment is distinguished by the English abbreviation.

For the measuring point information of the auxiliary control equipment, it is described by LNode node, and the examples are as follows:

```

<Device desc="Water spray fire extinguishing system" height="" length="" mRID="" name="wsy1" position-x="" position-y="" position-z="" type="WSY" width="">
  <LNode iedName="FFZ0001" IdInst="LD0" lnClass="GGIO" lnInst="1" lnType="GW_FFIRECONTROL_LD0_GGIO" prefix="WSY"/>
</Device>

```

3.3. IEC61850 Extension

This paper extends the substation area model and auxiliary equipment model on the basis of IEC61850, as shown in Figure 2. The green box indicates the content of the extension. This paper extends the class; The hollow triangle represents the inheritance relationship; The solid diamond represents the inclusion relationship.

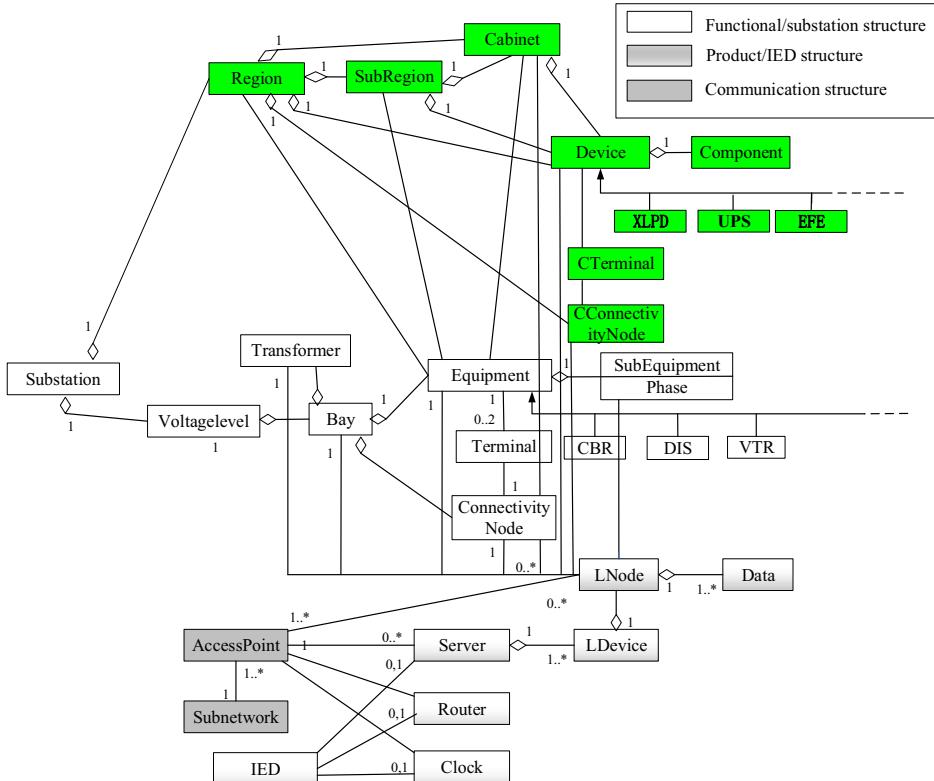


Figure 2. SSD structure diagram of Substation.

Firstly, the region class is added to describe the main area, such as "production complex building", and the extended sub region class is added to describe the sub areas, such as "first floor", "main control room" and so on. Cabinet description is added. Cabinet class, region class and sub region class belong to aggregation. Cabinet class must contain a sub region or a region. Add the device class to describe the auxiliary device.

The device class belongs to the aggregation relationship with the region, sub region and cabinet classes, and must include one of them. It describes the regional relationship of the auxiliary device, and describes the specific type of the auxiliary device through the type attribute. For example, "ups" stands for power supply, "CMR" stands for air conditioning. Examples are as follows:

```
<Region name="region1" desc="Production complex building" position-x="" position-y="" position-z="" mRID="" substationName="shanghailu">
```

```

<SubRegion name="subregion0" desc="1st floor" position-x="" position-y="" position-z="" mRID="" length="" width="" height="">
    <SubRegion name="subregion1" desc="Main control room" position-x="" position-y="" position-z="" mRID="" length="" width="" height="">
        <Device name="cmr1" desc="air conditioner" type="CMR" mRID="" position-x="" position-y="" position-z="" length="" width="" height="">
            <LNode         lnInst="1"         lnClass="xxxx" iedName="A_X0001" ldInst="LD0" prefix="" lnType="" />
        </Device>
    </SubRegion>
</SubRegion>
</Region>

```

4. Conclusion

With the development and progress of technology, the interaction between master station and substation will be closer, and the content of interaction will be further improved. At present, IEC61850 standard lacks the description of substation area model and auxiliary equipment model, and the description of primary equipment in IEC61850 is relatively complete. Based on this idea, this paper puts forward the idea of extension of area model and auxiliary model, so as to be consistent with the standard as far as possible, and the extended part also conforms to the standard idea. This paper is an exploration of the extension of auxiliary model in IEC61850 standard. The development of any new technology is inseparable from the engineering application.

References

- [1] Z L He. Relationship between intergrated automatic control system and power dispatching automation system [J]. Power System Technology, 2003, (03):68-70.
- [2] Y Z Xin, J J Shi, J Y Zhou, et al. technology development trends of smart grid dispatching and control systems [J]. Automation of Electric Power Systems, 2015, 39(1):2-8.
- [3] J G Yao, S C Yang, M H Shan. Reflections on operation supporting system architecture for future interconnected power grid [J]. Automation of Electric Power Systems, 2013, 37(21):52-59.
- [4] A L Chen, F Ye, et al. A design method of database information model using IEC 61850 common data classes [J]. Automation of Electric Power Systems, 2013, 37(13):88-92.
- [5] H D Zhang, H Zhang, et al. Model conversion method from smart substation SCD to control center CIM/E [J]. Automation of Electric Power Systems, 2012, 36(15):91-95.
- [6] C Fan, Y M Ni, R H Dou, et al. Interpretation of relevant specifications of integrated supervision and control systems in smart substations [J]. Automation of Electric Power Systems, 2012, 36(19):1-5.
- [7] D D Xu, W M Mi, Y Deng, et al. Integrated modeling for coordinated sharing of models/images and dispatch master substation [J]. Power System Technology, 2012, 36(11):235-239.
- [8] F L Jin, H Wang, G M Fan, et al. Design of centralized substation monitoring functions for smart grid dispatching and control systems [J]. Automation of Electric Power Systems, 2015, 39(1):241-247.
- [9] J L Li, Y Z Xin, G M Fan, et al. dynamic message coding technology of power systems [j]. automation of electric power systems, 2015, 39(1):14-18.
- [10] S Q Hu, L Li, Z Qi, et al. Research on and application of IEC 61850 modelling and CIM extension for protection relay information management system [J]. Automation of Electric Power Systems, 2016, 40(6):119-125.

Development and Application of Virtual Simulation Teaching Platform for the Transcritical CO₂ Two-Stage Compression Refrigeration System Based on LabVIEW

Yuyao SUN^{a,b}, Jinfeng WANG^{a,b,c}, Jing XIE^{a,b,c,d,l}, Chenlong WU^a

^a College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306, China

^b Shanghai Professional Technology Service Platform on Cold Chain Equipment Performance and Energy Saving Evaluation, Shanghai 201306, China

^c Shanghai Engineering Research Center of Aquatic Product Processing & Preservation, Shanghai 201306, China

^d National Experimental Teaching Demonstration Center for Food Science and Engineering, Shanghai Ocean University, Shanghai 201306, China

Abstract. Based on the principle of transcritical CO₂ two-stage compression refrigeration system, a virtual simulation teaching platform was developed. Two operation modes can be selected which are practice mode and test mode. According to questions and tips about experiments, students can do simulation experiments by open required valves, running the refrigeration system, adjusting pressure control valves and so on. When these operations are completed, students can save the file of the simulation experiment and exit the system. Then, test marks of students can be exported to Excel for teachers to view. In addition, some risks of operational errors can also be avoided, such as equipment damage. Moreover, if there is a computer, the simulation experiments can be done by this virtual simulation teaching platform, which is very convenient. The virtual simulation teaching platform is innovative and practical. The application of this virtual simulation teaching platform can improve the quality of the traditional teaching in the refrigeration field.

Keywords. refrigeration, LabVIEW, virtual simulation teaching platform, CO₂, simulation experiment

1. Introduction

Nowadays, various refrigeration systems are utilized in cold chain. In order to meet the society trend of rapid development, improving the performance of refrigeration systems has become a research focus. There are many refrigerants used in refrigeration systems, such as R22, R134a, R404a, R290 and CO₂ [1-5]. Different refrigerants can make different cooling effects and different environmental impacts. Due to some artificial refrigerants used can make bad phenomena like greenhouse effect, more natural refrigerants put into use [6]. CO₂ is one of excellent natural refrigerants. GWP and ODP

^lCorresponding author: Jing Xie, College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306, China; E-mail: jxie@shou.edu.cn.

of CO_2 are 1 and 0, respectively [7,8]. Moreover, the thermomechanical property of CO_2 is also very good. Transcritical CO_2 two-stage compression refrigeration system is a kind of widely applied refrigeration system which can make higher energy efficiency [9]. In terms of students whose majors are refrigeration, learning the theory of this refrigeration system in detail is beneficial for the development of the refrigeration field in the future.

The application of virtual simulation teaching platform can avoid the large cost of actual experiments, so it is necessary to conduct teaching the theory of transcritical CO_2 two-stage compression refrigeration system by virtual simulation teaching platforms. Moreover, the Ministry of Education proposed that the deep integration of information technology and education laboratory teaching in universities should be taken seriously [10,11]. In order to support this view of the Ministry of Education, many universities actively combine virtual simulation teaching platforms with teaching laboratory courses [12]. LabVIEW is a kind of graphical programming language which can be used to establish virtual simulation experiments [13]. In the refrigeration field, this software is utilized for virtual simulation experiments of various refrigeration systems [14,15]. A virtual simulation teaching platform for the transcritical CO_2 two-stage compression refrigeration system based on LabVIEW was developed by Shanghai Ocean University. Students can intuitively learn the theories of transcritical CO_2 two-stage compression refrigeration system and do relevant simulation experiments by this virtual simulation teaching platform. This virtual simulation teaching platform provides a powerful aid to the education of students whose specialty is refrigeration.

2. Principle of Transcritical CO_2 Two-Stage Compression Refrigeration System

The system schematic diagram of transcritical CO_2 two-stage compression refrigeration system is shown in Figure 1 (a). There are eight main components in this refrigeration system which are compressor in the low-pressure stage, compressor in the high-pressure stage, gas cooler 1, gas cooler 2, intercooler, electronic expansion valve 1, electronic expansion valve 2 and evaporator. The pressure-enthalpy diagram matching this system schematic diagram is shown in Figure 1 (b).

As shown in Figure 1 (b), there are ten processes in one transcritical CO_2 two-stage compression refrigeration cycle which are as follows.

- (1) Point A to point B: CO_2 is compressed for the first time in the compressor in the low-pressure stage.
- (2) Point B to point B': CO_2 is cooled in the gas cooler 1 at constant pressure.
- (3) Point C to point D: CO_2 is compressed for the second time in the compressor in the high-pressure stage.
- (4) Point D to point E: CO_2 with high temperature is cooled in the gas cooler 2 at constant pressure.
- (5) Point E to point F: in electronic expansion valve 1, CO_2 is throttled and the pressure of CO_2 is decreased.
- (6) Point F to point G: the heat exchange process between low-temperature CO_2 and high-temperature CO_2 in the intercooler (the thermal energy of low-temperature CO_2 is raised).
- (7) From point B' and point G to point C: CO_2 from the gas cooler 1 and the intercooler are mixed.

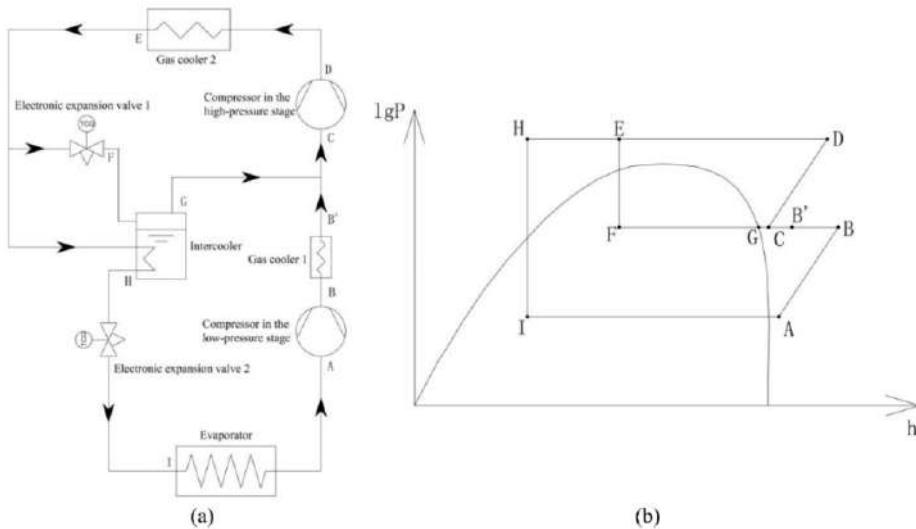


Figure 1. transcritical CO_2 two-stage compression refrigeration system, (a) system schematic diagram, (b) pressure-enthalpy diagram.

(8) Point E to point H: the heat exchange process between high-temperature CO_2 and low-temperature CO_2 in the intercooler (the thermal energy of high-temperature CO_2 is declined).

(9) Point H to point I: in electronic expansion valve 2, CO_2 is throttled and the pressure of CO_2 is decreased.

(10) Point I to point A: CO_2 absorbs the heat and is evaporated in the evaporator at constant pressure.

According to the principle of transcritical CO_2 two-stage compression refrigeration system, the virtual simulation teaching platform for the transcritical CO_2 two-stage compression refrigeration system based on LabVIEW is designed by Shanghai Ocean University. This virtual simulation teaching platform can be used in the teaching of refrigeration-related courses, and students can study knowledge about transcritical CO_2 two-stage compression refrigeration system by this virtual simulation teaching platform.

3. Design of the Virtual Simulation Teaching Platform for the Transcritical CO_2 Two-Stage Compression Refrigeration System

3.1. Basic Information of the Virtual Simulation Teaching Platform for the Transcritical CO_2 Two-Stage Compression Refrigeration System

The virtual simulation teaching platform for the transcritical CO_2 two-stage compression refrigeration system was designed by LabVIEW 2016 with the Windows 10 operation system. Many illustrations like compressor, gas cooler and evaporator were provided by the image gallery of LabVIEW. Some special illustrations could be drawn by AutoCAD and be imported to LabVIEW. Based on these illustrations, clear and intuitive front panels were created. On the basis of this virtual simulation teaching platform, opportunities for students to operate experiments about the transcritical CO_2 two-stage

compression refrigeration system can be increased. In addition, this virtual simulation teaching platform is positive to the improvement of teaching effectiveness.

3.2. Teaching Module Design of the Virtual Simulation Teaching Platform for the Transcritical CO₂ Two-Stage Compression Refrigeration System

The processes of the teaching module of the virtual simulation teaching platform for the transcritical CO₂ two-stage compression refrigeration system is disclosed in Figure 2. First, users should import their own academic numbers (students) or work numbers (teachers) and corresponding passwords. If academic numbers (or work numbers) and corresponding passwords are recognized, the login interface of the software will be turned to the selection interface of the operation modes. There are two operation modes can be chosen which are practice mode and test mode. Contents of experiment operation in these two modes are the same, but there is a time limit in the test mode, and the user's experiment operation is scored by the system automatically. In this software, the time set for each test is 600 s. According to the requirements of teaching, teachers can set up questions and tips in this virtual simulation teaching platform. After the selection of operation modes, users can see these questions and tips set in advance first. When these questions and tips are confirmed, users can press the "OK" button. Then, the transcritical CO₂ two-stage compression refrigeration system will be displayed on the screen. Users must operate the equipment in the transcritical CO₂ two-stage compression refrigeration system based on the questions and the tips set in advance. If there is an error in the operation, a matching tip about how to conquer this problem will be shown at once. User can revise error operation after understanding this tip. However, if the error in the operation is too serious, tips names "Serious error occurred" will appear, and the software will end after the exit button being pressed. This setting simulates the emergency shutdown function of the transcritical CO₂ two-stage compression refrigeration system when a serious problem appears. In reality, this function can be carried out by safety valves. When users finish all operation, the submit button can be pressed. Then, all practice or test is over. Moreover, if the operation time is longer than

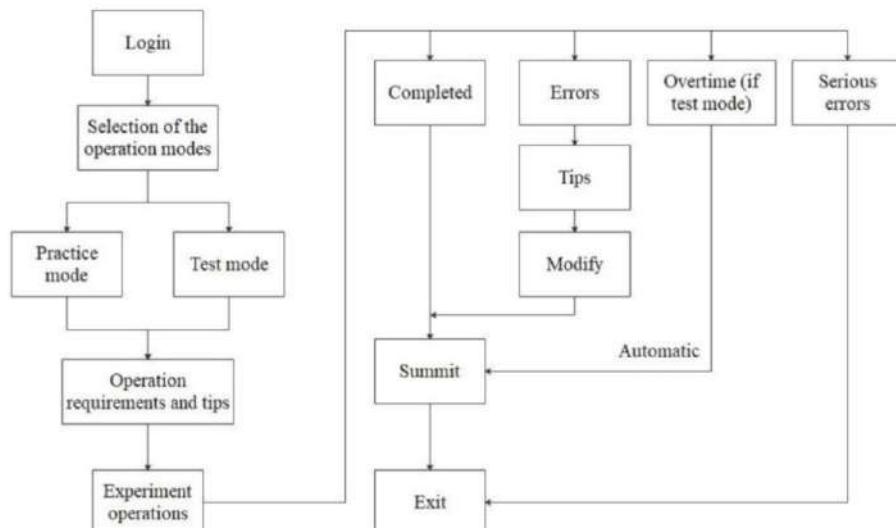


Figure 2. The processes of the teaching module.

600 s in the test mode, test papers will be submitted automatically. Test marks of students can be exported to Excel for teachers to view. Some operation interfaces are shown in Figure 3.

Based on the virtual simulation teaching platform for the transcritical CO₂ two-stage compression refrigeration system, students can pre-learn and revise the experiment operation of the transcritical CO₂ two-stage compression refrigeration system by the practice mode and conduct the test of the transcritical CO₂ two-stage compression refrigeration system by the test mode.

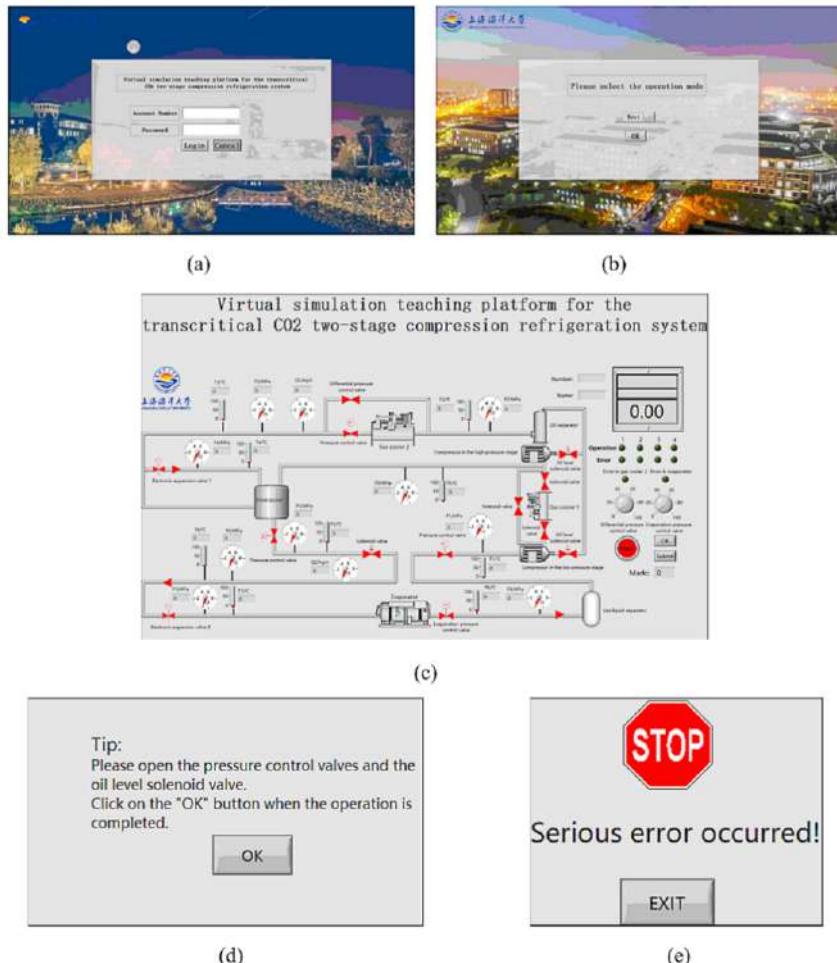


Figure 3. Operation interfaces, (a) login interface, (b) selection interface of the operation modes, (c) experiment operation interface, (d) tip example, (e) serious error report interface.

3.3. Simulation Experiment Design of the Virtual Simulation Teaching Platform for the Transcritical CO₂ Two-Stage Compression Refrigeration System

The interface of the simulation experiment design of the virtual simulation teaching platform for the transcritical CO₂ two-stage compression refrigeration system is shown in Figure 3 (c). The refrigeration system for the experiment operation is designed based on the principle of transcritical CO₂ two-stage compression refrigeration system. Besides eight main components in the transcritical CO₂ two-stage compression refrigeration system, an oil separator, a gas-liquid separator, various regulating valves, temperature sensors, pressure sensors and flow meters are added. The extra equipment is used to control the refrigeration system and observe the parameter changes at each state point in the refrigeration system. The operation conditions of the transcritical CO₂ two-stage compression refrigeration system is controlled by various valves which are on the right of the interface, like the differential pressure control valve and the evaporation pressure control valve. The number of these valves can be added or be subtracted based on the teaching requirements.

If users want to know whether the experiment operation is accurate, they can observe the indicator lights on the interface. There are two rows of indicator lights which represent whether the experiment operation has done and whether it is correctly, respectively. If the experiment operation has done, the indicator lights in the first row will be switched on. If there is a mistake in any experiment operation, the corresponding indicator lights in the second row will be switched on. If all experiment operations have been finished, the main switch can be pressed. Then, the refrigeration system will start running. In addition, if there are deviations in the parameters, users should adjust valves until parameters meet requirements.

The operation time is displayed in the top right corner of the interface. If the time is displayed to 600 s, the experiment operation will be interrupted automatically and the test paper will be submitted automatically. In addition, the test time can be changed by teachers. Furthermore, the test marks of users can be seen from the test mark box. The test mark in the test mark box changes with the variations of experiment operations.

4. Application and Characteristics of the Virtual Simulation Teaching Platform for Transcritical CO₂ Two-Stage Compression Refrigeration System

4.1. Application of the Virtual Simulation Teaching Platform for Transcritical CO₂ Two-Stage Compression Refrigeration System

The experimental data are obtained from the refrigeration system shown in Figure 4. The data is input in the software during the course of software writing. In the process of teaching, users can carry out simulation experiments by the virtual simulation teaching platform.



Figure 4. Transcritical CO₂ two-stage compression refrigeration system.

A simulation experiment is conducted in this paper, which is to adjust the refrigeration system in the operation condition of Table 1. The refrigeration system used for this experiment is transcritical CO₂ two-stage compression refrigeration system with auxiliary gas cooler. The auxiliary gas cooler is the gas cooler 1 in the refrigeration system.

Table 1. Parameters of the simulation experiment.

State point	T (°C)	h (kJ/kg)	s (kJ/(kg · K))	P (MPa)
A	-20.00	440.70	1.98	1.79
B	42.28	482.21	2.00	4.06
B'	13.88	442.04	1.86	4.06
C	10.90	436.76	1.85	4.06
D	78.00	476.05	1.86	9.20
E	32.00	282.92	1.26	9.20
F	5.90	282.92	1.30	4.06
G	5.90	426.75	1.81	4.06
H	4.99	207.08	1.00	9.20
I	-23.00	207.08	1.04	1.79

It is assumed that the operator of the experiment is a student. First, the student should import his academic number and the corresponding password. Next, the practice mode is chosen. Then, the experiment operation interface is displayed on the screen. The processes of this simulation experiment are described as follow.

(1) Open the oil level solenoid valves which are next to compressors, and click on the "OK" button when the operation is finished.

(2) Open all pressure control valves and electronic expansion valves, and click on the "OK" button when the operation is finished.

(3) Open all solenoid valves except for the solenoid valve on the line in parallel with gas cooler 1, and click on the "OK" button when the operation is finished.

(4) Press the main switch to start the refrigeration system. The parameters at each state point will be gradually stable at a pre-set operation condition.

(5) Parameters (temperature, pressure and so on) at each status point should be adjusted by adjusting pressure control valves, such as the differential pressure control valve and the evaporation pressure control valve. PA should be 1.79 MPa and PE should be 9.20 MPa, which represent evaporation pressure and condensation pressure, respectively. Other parameters can be adjusted by the same method. After these adjustments, parameters can meet the requirements of the simulation experiment.

(6) Save the file of the simulation experiment and exit the system.

Based on the above-mentioned operations, a simulation experiment about the transcritical CO_2 two-stage compression refrigeration system is completed, which can be seen from Figure 5. Many other simulation experiments also can be done with this virtual simulation teaching platform. Doing experiments with this virtual simulation teaching platform can reduce the experimental costs and avoid the damage to the equipment due to operator errors. In addition, according to the course schedule, teachers can set up the requirements of simulation experiments and some operation tips in advance in the virtual simulation teaching platform according to the course schedule.

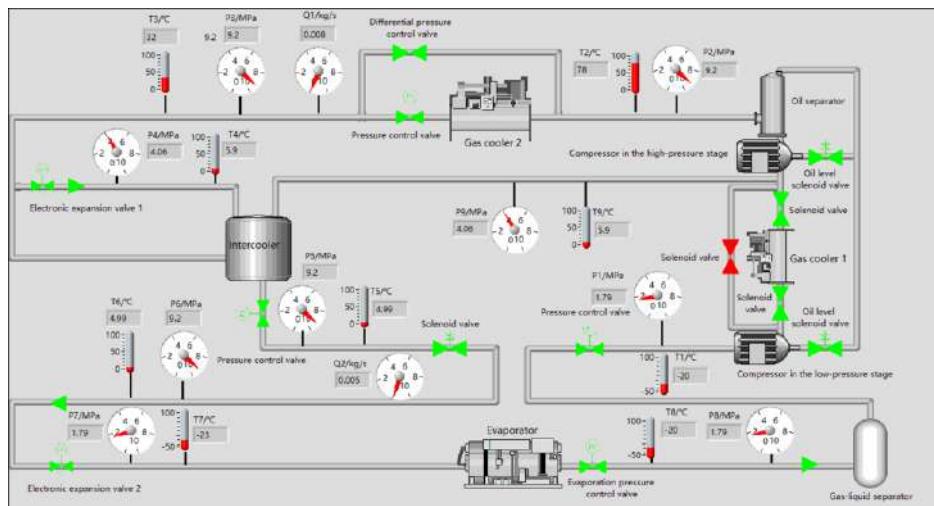


Figure 5. Simulation experiment.

4.2. Application of the Virtual Simulation Teaching Platform for Transcritical CO_2 Two-Stage Compression Refrigeration System

Characteristics of the virtual simulation teaching platform for transcritical CO_2 two-stage compression refrigeration system are as follows.

(1) This virtual simulation teaching platform can be combined with traditional teaching. The transcritical CO_2 two-stage compression refrigeration system can be displayed in the virtual simulation teaching platform intuitively which is helpful for students to better understand the composition and operation of this refrigeration system.

(2) If there is a computer, the simulation experiments can be done by this virtual simulation teaching platform, which is very convenient. In addition, the running memory of this virtual simulation teaching platform is only 5.32 MB.

(3) Users can do simulation experiments by this virtual simulation teaching platform repeatedly. In addition, the costs of doing simulation experiments are much smaller than those of doing real experiments.

(4) Perform experiments by this virtual simulation teaching platform can avoid risks of operational errors, such as equipment damage.

(5) Experiment operation tests can be held by this virtual simulation teaching platform. This method is more efficient than using actual experiment equipment to conduct experiment operation tests.

On the basis of the above-mentioned characteristics of the virtual simulation teaching platform, it is obvious that the virtual simulation teaching platform is both innovative and practical and has a positive impact on teaching in the refrigeration field.

5. Conclusions

A virtual simulation teaching platform for the transcritical CO_2 two-stage compression refrigeration system based on LabVIEW is disclosed in this paper. This virtual simulation teaching platform can be combined with traditional teaching about transcritical CO_2 two-stage compression refrigeration system. In addition, if there is a computer, the simulation experiments can be done by this virtual simulation teaching platform, which is very convenient.

Moreover, students can train more times of simulation experiments by this virtual simulation teaching platform than by real experiments. And, the costs of doing simulation experiments are much lower. Some risks of operational errors can also be avoided, such as equipment damage. In addition, the running memory of the virtual simulation teaching platform is small.

Furthermore, the practice mode and the test mode can be selected in this virtual simulation teaching platform. Tips about experiment operations can also be set up in the software which can help students understand experiment operations about the transcritical CO_2 two-stage compression refrigeration system.

Finally, the virtual simulation teaching platform can be used for the simulation of many studies about the transcritical CO_2 two-stage compression refrigeration system, which is beneficial to the investigation of the transcritical CO_2 two-stage compression refrigeration system in the future.

Nomenclature			
Temperature	T (°C)	Pressure	P (MPa)
Enthalpy	h (kJ/kg)	Entropy	s (kJ/(kg·K))
Subscripts			
State point	A, B, B', C, D, E, F, G, H, I		

Acknowledgements

This research was supported by Science and Technology Innovation Action Plan of Shanghai Science and Technology Commission (19DZ1207503), Public Service Platform Project of Shanghai Science and Technology Commission (20DZ2292200).

References

- [1] C S Choudhari, S N Sapali. Performance investigation of natural refrigerant R290 as a substitute to R22 [J]. *Refrigeration Systems Energy Procedia*, 2017, 109: 346-352.
- [2] Wantha C. Experimental investigation of the influence of thermoelectric subcooler on the performance of R134a refrigeration systems [J]. *Applied Thermal Engineering*, 2020, 180: 115829.
- [3] V Oruç, A G Devecioglu. Experimental investigation on the low-GWP HFC/HFO blends R454A and R454C in a R404A refrigeration system [J]. *International Journal of Refrigeration*, 2021, 128: 242-251.
- [4] T Bai, D Li, H X Xie, G Yan, J L Yu. Experimental research on a Joule-Thomson refrigeration cycle with mixture R170/R290 for -60°C low-temperature freezer [J]. *Applied Thermal Engineering*, 2021, 186: 116476.
- [5] F Giunta, S Sawalha. Techno-economic analysis of heat recovery from supermarket's CO₂ refrigeration systems to district heating networks [J]. *Applied Thermal Engineering*, 2021, 193: 117000.
- [6] K Z Skácanová, M Battesti. Global market and policy trends for CO₂ in refrigeration [J]. *International Journal of Refrigeration*, 2019, 107: 98-104.
- [7] M Z Pan, X Y Bian, Y Zhu, Y C Liang, F L Lu, G Xiao. Thermodynamic analysis of a combined supercritical CO₂ and ejector expansion refrigeration cycle for engine waste heat recovery [J]. *Energy Conversion and Management*, 2020, 224: 113373.
- [8] L Peng, G Q Shu, H Tian. Carbon Dioxide as working fluids in transcritical rankine cycle for diesel engine multiple waste heat recovery in comparison to hydrocarbons [J]. *Journal of Thermal Science*, 2019, 28 (3): 494-504.
- [9] Y L Song, H D Wang, X Yin, F Cao. Review of transcritical CO₂ vapor compression technology in refrigeration and heat pump [J]. *Journal of Refrigeration*, 2021, 42 (02): 1-24.
- [10] Y P Wang, J Yong, W R Wang, B Yue. Construction and exploration of a model virtual simulation experiment project [J]. *Computer Education*, 2019, 9: 38-41.
- [11] H B Liu, J Shen, G S Wang, S Y Liu, Q Guo. Construction of virtual simulation experimental teaching resource platform based on engineering education [J]. *Experimental Technology and Management*, 2019, 36 (12): 19-22+35.
- [12] Y D Chen, G Y Jia, L Dai, J Yang, L P Zhang. Development and application of virtual simulation experimental teaching platform for tunnel construction [J]. *Experimental Technology and Management*, 2021, 38 (06): 151-155+160.
- [13] S Sivarajani, S Velmurugan, K Kathiresan, M Karthik, B Gunapriya, C Gokul, M Suresh. Visualization of virtual environment through labVIEW platform. *Materials Today: Proceedings*, 2021, 45: 2306-2312.
- [14] R Cabello, D Sánchez, R Llopis, L Nebot-Andrés, D Calleja-Anta. Energy evaluation of a low temperature commercial refrigeration plant working with the new low-GWP blend R468A as drop-in of R404A [J]. *International Journal of Refrigeration* 2021, 127: 1-11.
- [15] R B Barta, D Ziviani, Groll E A. Design and commissioning of a modular multi-stage two-evaporator transcritical CO₂ test stand [J]. *International Journal of Refrigeration* 2021, 2: JIJR 5165.

Study on Optimization Path of Energy Consumption Structure of Container Terminal

Ruiyu LI^{a1}, Haifeng FANG^b, Guodong WANG^c, Qingfang MA^c

^a*Environment Prevention and Energy Conservation Center, China Waterborne Transport Research Institute, Beijing 100088, China*

^b*Putian Xiuyu Port Co., Ltd, Putian 351158, China*

^c*Professional Competence Center, Liaoning Port Group Co., Ltd, Dalian 116007, China*

Abstract. This study accounts current energy consumption of various types of equipments in Chinese container terminals through investigating typical terminals; compares and analyzes the clean energy application technologies from the perspectives of technical level, investment cost, and others; on this basis, construct the predictive model of energy consumption structure, and uses scenario analysis to carry out energy consumption predictions under each scenario and analyzes the effect of policy intervention, technological development and other factors. According to the predictive results, this study holds that in order to optimize energy structure of container terminal, container terminals should strongly promote the application of clean energy to port machinery instead of fuel on the basis of the industrial development and cost reduction of high-power and large-capacity power batteries; at the same time, strengthen policy encouragement and guidance are needed.

Keywords. Energy consumption structure, clean energy, carbon reduction, green transformations

1. Introduction

As an important energy-using system of the transportation industry, port industry takes on major responsibilities for energy conservation, carbon reduction and green transformation. According to the data of the International Energy Agency, China's energy consumption in transportation field in 2018 was 467 million tons of standard coal, accounting for 10.2% of the total energy consumption; Direct carbon emissions from transportation field are about 1 billion tons, accounting for about 10.7% of the total carbon emissions. In 2018, China's apparent oil consumption will reach 648 million tons, more than half of which will be used as transportation fuel. Diesel, gasoline, fuel oil and aviation kerosene will account for more than 90% of transportation energy. In the field of transportation energy consumption, waterway energy consumption accounts for about 23% [1].

Waterway transportation consists of port and shipping, among them, the port as an important link has become an important source of energy consumption and greenhouse

¹Corresponding author: Ruiyu Li, Environment Prevention and Energy Conservation Center, China Waterborne Transport Research Institute, Beijing 100088, China; E-mail: liruiyu@wti.ac.cn.

gas emissions. Low-carbon green development of the port will have a huge impact to the port and the city's economic development. According to the estimation of relevant research materials, at present, the total energy consumption of the port directly used for handling cargo production is about 2 million tons of standard coal per year, including about 600000 tons of diesel oil, equivalent to 874000 tons of standard coal, accounting for 43.7%, and direct carbon emissions is 1.9 million tons [2]. The scale and carbon emissions of container terminals are relatively high among all types of ports in China, and the carbon emissions peak of container terminals plays an important role in carbon emissions peak of ports in China. Optimization energy consumption structure is an important way to promote carbon peak [3-5].

2. Analysis of Current Energy Consumption Structure in Container Terminals

2.1. Production Process and Equipment of Container Terminal

The production units of container terminal include cargo handling production, auxiliary production and ancillary production. The cargo handling process of container terminals is shown in Figure 1.

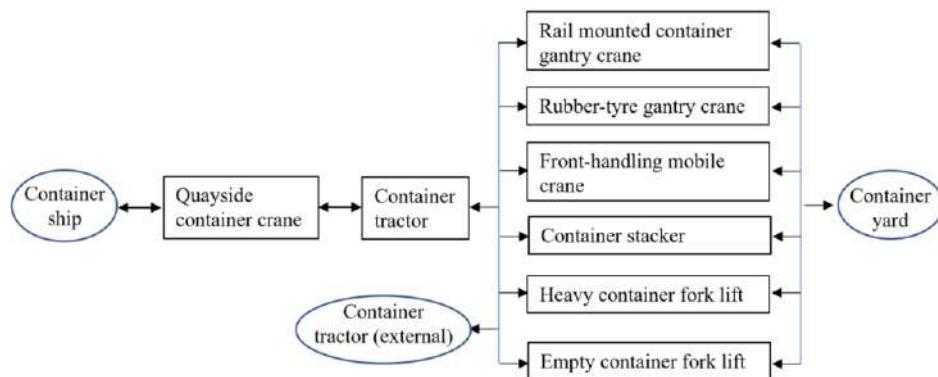


Figure 1. Cargo handling flow chart of container terminals.

Main handling machines of container terminal includes:

- Quayside container crane is crane for handling containers at the frontier of container terminals, hereinafter referred to as QC. QC works by electricity.
 - Container tractor is special tractors for towing container trailers, including inner container tractors for transportation inside the container terminal and outer container tractors for transportation outside the port, hereinafter referred to as CT. CT runs on diesel oil or LNG.
 - Rail mounted gantry crane is the gantry crane for handling containers in container yard that walks supported by steel wheels, hereinafter referred to as RMG. RMG works by electricity.

- Rubber-tyre gantry crane is the gantry crane for handling containers in container yard that walks supported by tires, hereinafter referred to as RTG. RMG works by electricity or diesel oil.
 - front-handling mobile crane is telescopic jib crane for handling containers in container yard that mounted on self-propelled tire chassis, hereinafter referred to as FHM. FHM works by diesel oil.
 - Container stacker is machinery for stacking empty containers in container yard, hereinafter referred to as stacker. Stackers work by diesel oil or LNG.
 - Heavy container fork lift is fork truck for handling and stacking heavy containers in container yard, hereinafter referred to as HFL. HFL works by diesel oil.
 - Empty container fork lift is fork truck for handling and stacking empty containers in container yard, hereinafter referred to as EFL. EFL works by diesel oil or electricity.

The types of energy consumed by equipment of container terminal are shown in Figure 2.

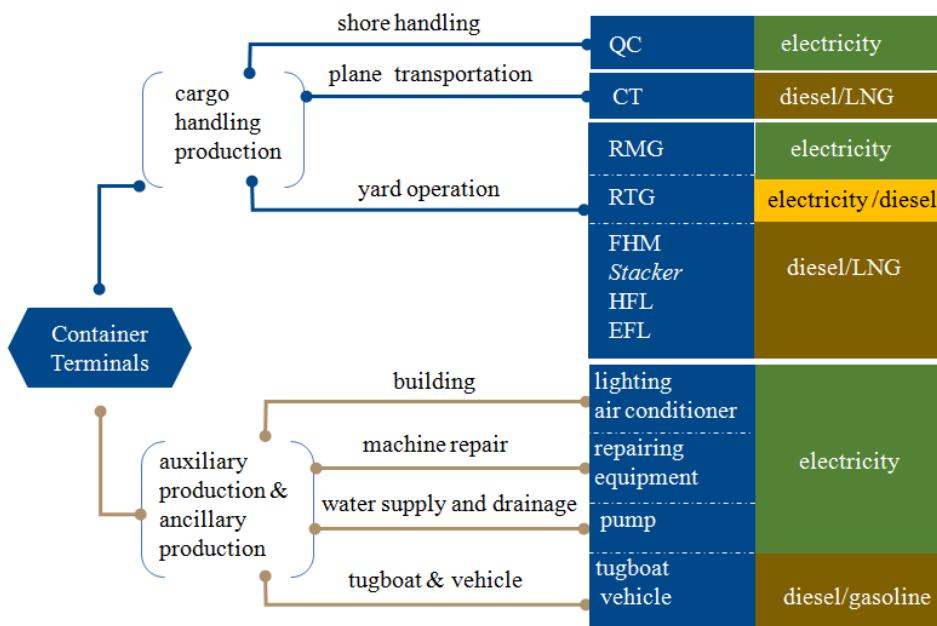


Figure 2. Types of energy consumed by equipment of container terminal.

2.2. Current Situation of Energy Consumption Structure of China's Container Terminal

Based on energy consumption of various types of equipment of China's typical container terminals, as well as the container terminals throughput and other production data published in the 2020 statistical yearbook, this study accounts energy consumption of various types of equipment of China's container terminals in 2020 as follows.

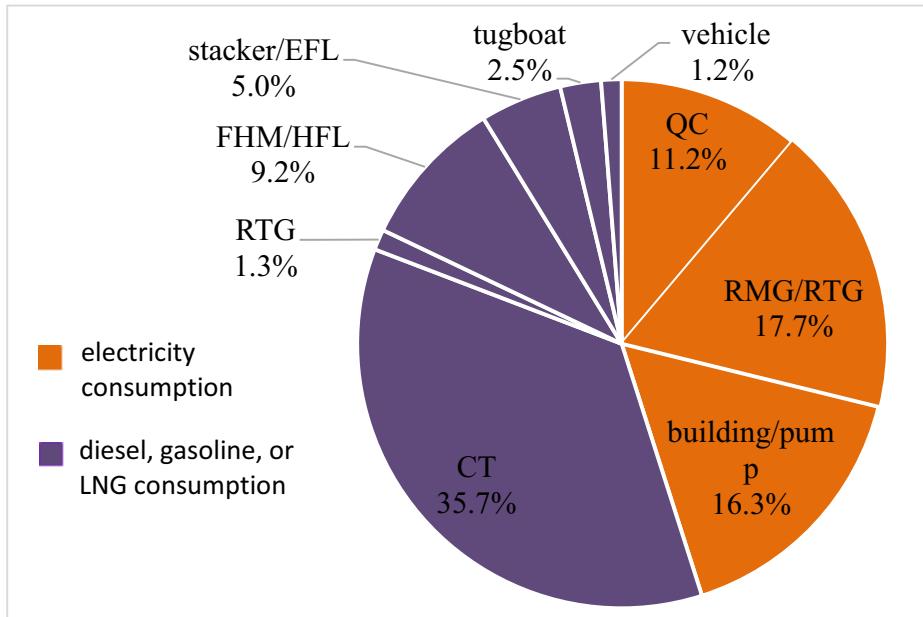


Figure 3. Current energy consumption structure of China's container terminals in 2020.

As can be seen from Figure 3 that in 2020, the total energy consumption of China's container terminals will be 872,000 tons of standard coal, of which the consumption of diesel, gasoline and LNG will be 479,000 tons of standard coal, accounting for about 55% of the total energy consumption, and the direct carbon emissions of China's container terminals will be about 1.033 million tons. In addition, it can be seen from Figure 3 that the consumption of diesel oil, gasoline and LNG of China's container terminals is more than that of electric power. In all of the equipment consuming diesel, gasoline or LNG, container tractor consumes most energy, and energy consumption accounts for about 35.7%, and direct carbon emissions account for about 65.1%.

3. Optimization Measures of Energy Consumption Structure of Container Terminal

Optimization measures of energy consumption structure of container terminal include promoting the application of clean energy such as electric energy and hydrogen energy instead of oil/gas in port machineries. This paper analyzes the clean energy application technologies of different types of port machineries and vehicles in container terminals as follows [6]:

3.1. Container Tractor (Lithium-Powered Batteries)

- *Status of technology application:* Unmanned container tractors are used in Tianjin Port and Xiamen Port Harun Wharf; manned electric container tractors are used in ports such as Wuhan Yangtze Port.

- *Status of technology development:* The maximum speed of electric container tractors is adjustable and the output torque is stable, which meets the basic technical parameters of container trailers. Battery life and charging time are difficult to meet the requirements of 24h continuous operation, affecting the efficiency of terminal operations.
- *Cost-effectiveness:* According to the 10-year life cycle, it is estimated that the annual operating cost of electric container tractors is about 10% higher than that of diesel-driven ones.

3.2. Container Tractor (Hydrogen Fuel Cells)

- *Status of technology application:* In 2019, three hydrogen energy container tractors were officially put into live demonstration operation in Qingdao Port, putting China's first new energy (hydrogen fuel + pure electric energy) heavy trucks into use, known as the first mass use of energy container tractors in the domestic port industry.
- *Status of technology development:* The main performance indicators of container tractors driven by hydrogen fuel, such as power and endurance meet the production needs, but the hydrogen fuel supply conditions are not perfect.
- *Cost-effectiveness:* According to the 10-year life cycle, it is estimated that the annual operating cost of container tractors driven by hydrogen energy is about three times higher than that of diesel-driven ones.

3.3. RTG (Mains+Lithium Batteries / External Power Supplies)

- *Status of technology application:* Domestic container terminal RTG's loading and unloading operations have been driven by mains power, but most terminal RTG transfers are still driven by small diesel engines.
- *Status of technology development:* The large-capacity lithium battery can basically meet the energy demand of wheeled crane transfers.
- *Cost-effectiveness:* As energy storage technology develops, the cost of chemical fuel cells such as lithium batteries has gradually reduced.

3.4. FHM/HFL (Lithium Batteries)

- *Status of technology application:* Electric container reach stackers have been tested or put into use in Ningbo Port, Xuzhou Port and other ports.
- *Status of technology development:* Electric container reach stackers vary greatly in load and have high requirements for battery power density.
- *Cost-effectiveness:* The purchase cost of an electric container reach stacker is much higher than that of a diesel-driven one, but its fuel cost is lower.

3.5. Stacker/EFL (Lithium Batteries) [7]

- *Status of technology application:* Electric empty container handlers have been applied in Shantou CMPort Container Company and Xiamen Port Haitian Terminal.
- *Status of technology development:* The response speed of electric empty container handlers on the move is faster than that of internal combustion engines. The lifting speed is more uniform with unobvious ups and downs. Thus drivers can quickly adapt to and get familiar with the operation. In addition, the noise is significantly smaller

than that of internal combustion engines. But the battery operation capability needs to be further improved.

- *Cost-effectiveness*: Compared with diesel-driven ones, the cost of electric empty container handlers is higher. But the investment cost can be recovered within about 3 years.

4. Scenario Analysis of Optimization Energy Consumption Structure of Container Terminals

4.1. Set Scenarios of Energy Consumption Structure of Container Terminals

This paper takes the clean energy technology application scale of various port machines and vehicles as the influencing factors, and the application scale is influenced by policy intervention, technological development and so on. According to the clean energy technology application scale of various port machines and vehicles, this study takes 2030 as the target year and set three scenarios as follows.

4.1.1. Baseline Scenario

The baseline scenario refers to under a background of no new policies and the development of technologies such as clean energy is slow, the clean energy technology application scale of various port machines and vehicles are not widespread. Under this scenario, the port enterprises phase out fossil fuel equipment in accordance with their ages and working performance by convention, and few clean energy equipment are selected.

4.1.2. Policy Intervention Scenario

Policy intervention scenario means that the government adopts policy constraints, tax incentives, pilot demonstrations and other measures to encourage and guide port enterprises to carry out optimization measures of energy consumption structure, and under which the clean energy technology application scale of various port machines and vehicles will progress. Under this scenario, port companies invest more money in optimization energy consumption structure, accelerate the phase-out of fossil fuel equipments, and more clean energy or low-carbon equipments are selected. In addition, port enterprises are increasingly adopting information technology and intelligent technology to improve the production efficiency of port equipments.

4.1.3. Technological Breakthrough Scenario

The technological breakthrough scenario refers to under a background of clean energy application technology breakthrough, the clean energy technology of port mobile machinery and vehicles develop rapidly. Port enterprises accelerate the adoption of clean energy technologies and the selection of high-efficient low-carbon products for mobile machinery and vehicles on a large scale. And, with the rapid development of intelligent and information technology, the container terminal intelligence has increased significantly.

Based on aforementioned information, the paper sets out proportion of the clean energy technology application scale of various port machines and vehicles respectively

in table 1.

Table 1. Proportion of clean energy technology application scale of various port machines and vehicles under three scenarios.

	<i>Baseline scenario</i>	<i>Policy intervention scenario</i>	<i>Technological breakthrough scenario</i>
CT	40%	60%	90%
RTG	30%	50%	70%
FHM/HFL	30%	50%	70%
Stacker/EFL	40%	60%	90%
Tugboat	20%	40%	60%
Port vehicles	40%	60%	90%

4.2. Forecast Energy Consumption Structure of Container Terminal

This paper forecast the energy consumption structure in China's container terminals under three scenarios as follows [8-9].

4.2.1. Baseline Scenario

The energy consumption structure in China's container terminal in 2030 under baseline scenario is forecast as follow.

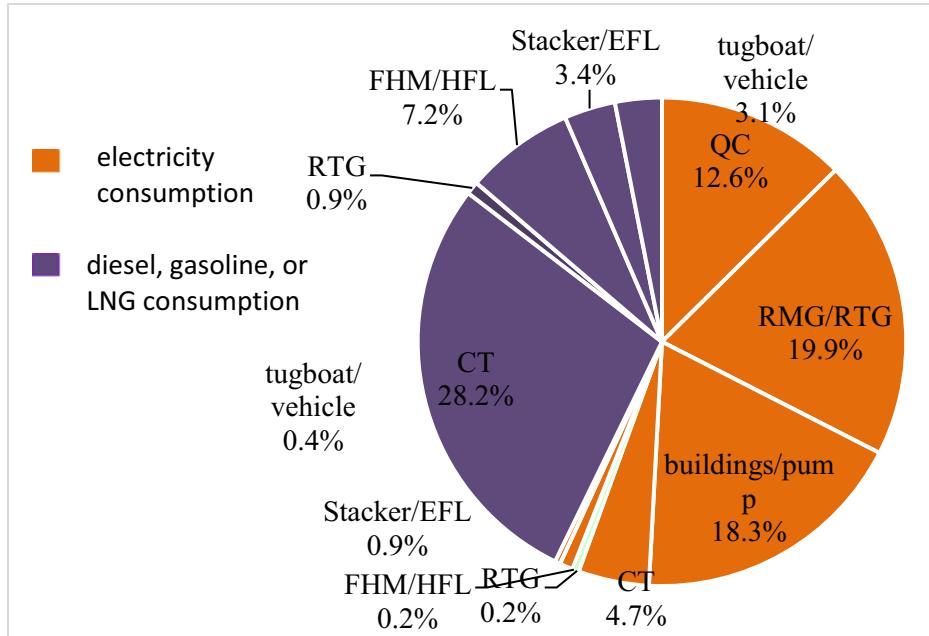


Figure 4. Forecast energy consumption structure in China's container terminal under baseline scenario.

As can be seen from Figure 4 that, in 2030 under the baseline scenario:

- Consumption of diesel, gasoline and LNG will account for about 42.8% of the total energy consumption, which is 12.2% lower than now. In addition, it can be seen from Figure 4 that the consumption of diesel oil, gasoline and LNG of China's container

terminals is less than that of electric power. In all of the equipment consuming diesel, gasoline or LNG, container tractor consumes most energy, and energy consumption accounts for about 28.2%.

- The direct carbon emissions of China's container terminals per throughput will be about 31% lower.

4.2.2. Policy Intervention Scenario

The energy consumption structure in China's container terminal in 2030 under policy intervention scenario is forecast as follow.

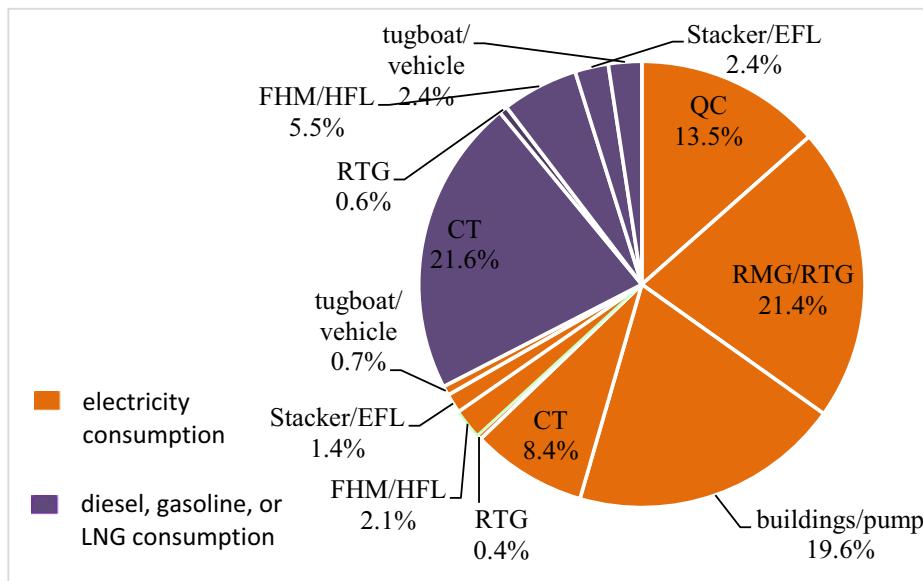


Figure 5. Forecast energy consumption structure in China's container terminal under policy intervention scenario.

As can be seen from Figure 5 that, in 2030 under the policy intervention scenario:

- Consumption of diesel, gasoline and LNG will account for about 32.6% of the total energy consumption, which is 22.4% lower than now. In addition, it can be seen from Figure 5 that the consumption of diesel oil, gasoline and LNG of China's container terminals is less than that of electric power. In all of the equipment consuming diesel, gasoline or LNG, container tractor consumes most energy, and energy consumption accounts for about 21.6%.
- The direct carbon emissions of China's container terminals per throughput will be about 51% lower.

4.2.3. Technological Breakthrough Scenario

The energy consumption structure in China's container terminal in 2030 under

technological breakthrough scenario is forecast as follow.

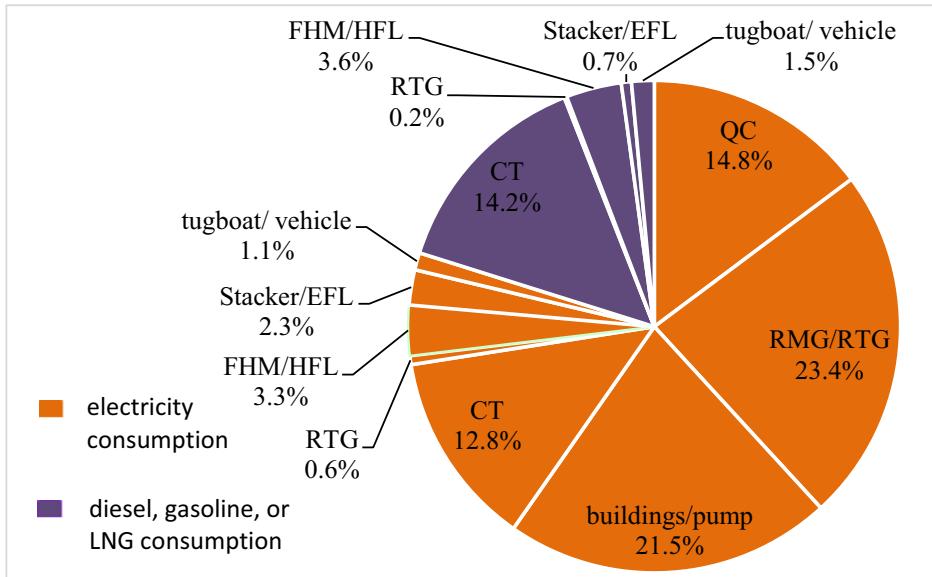


Figure 6. Forecast energy consumption structure in China's container terminal under technological breakthrough scenario

As can be seen from Figure 6 that, in 2030 under the technological breakthrough scenario:

- Consumption of diesel, gasoline and LNG will account for about 20.2% of the total energy consumption, which is 34.8% lower than now. In addition, it can be seen from Figure 5 that the consumption of diesel oil, gasoline and LNG of China's container terminals is less than that of electric power. In all of the equipment consuming diesel, gasoline or LNG, container tractor consumes most energy, and energy consumption accounts for about 14.2%.
- The direct carbon emissions of China's container terminals per throughput will be about 72% lower.

5. Optimization Path of Energy Consumption Structure of Container Terminal

Based on the prediction of the improvement of energy structure and carbon emission intensity of container terminal under three scenarios, this paper holds that, in order to accomplish China's carbon emission peak task in 2030 ahead of schedule, the port industry should focus on technological breakthroughs, and government should introduce a series of policies.

5.1. Promote the Application of Clean Energy Technologies for Port Machinery and Vehicles

At present, the main problem that restricts the large-scale application of electric-energy

transport vehicles and mobile machinery is that energy density and power density of the power battery can't meet the demand of port production. Port production is busy and port machinery and vehicles are used frequently, but the battery last no longer to meet the demand of port production. In addition, the cost of electric machinery and vehicles is higher than that of fuel, so it would place a heavy burden on the container terminal economy if they change so large number of fuel machinery and vehicles. [10-13]

As a clean fuel with high calorific value, hydrogen fuel is considered to be the ultimate solution to the future energy crisis. The application of hydrogen fuel cell is another important direction. Up to now, hydrogen container tractors have been used in Qingdao Port Container Terminal, but it still takes some time for the industrialization of hydrogen fuel cells, and the high price of hydrogen fuel have restricted the universal application. [14]

Therefore, it is necessary to focus on lithium batteries, hydrogen fuel cells and other power cell technologies, and accelerate the research and development of clean energy technologies for port mobile machinery and transport vehicles, so as to achieve breakthroughs in power cell technologies and significant cost reductions, and thus accelerate the process of carbon emission peak at container terminals.

5.2. Use of High-Efficient, Low-Carbon Fuel Machinery and Vehicles

Reducing carbon emissions from internal combustion engines should improve the thermal efficiency of internal combustion engines, and use "high-efficient and low-carbon" internal combustion engines that consume less fuel while producing the same power. With the development of intelligent, lightweight, diversified and other new technologies for internal combustion engines, the application of efficient and intelligent combustion systems for internal combustion engines should be strengthened.

The innovation and research of high compression ratio, active and passive pre-combustion chamber, thin combustion technology, ignition compression ignition technology, high-efficiency supercharger technology, and deep friction reduction technology should be enhanced to achieve further improvement in the thermal efficiency of internal combustion engines [10-12].

5.3. Improving the Productivity of Fuel Machinery and Vehicles

With the development of intelligent, information technology, container terminal can optimize scheduling of machinery and vehicle, and then improve production efficiency. Using technologies such as IoT and intelligent algorithms to initiate real-time positioning of trucks and mobile machinery, and carry out intelligent scheduling, with the aim of reducing no-load operation time of port equipment, and improving equipment load factor, etc. Optimize the driving path of trucks and mobile machinery to improve equipment productivity, thus reducing energy consumption and carbon emissions [15-16].

6. Conclusion

In the world transportation system, the role of container transportation presents the characteristics of a new value system and maintains a rapid development trend. Optimizing the energy structure of container terminals and promoting carbon neutralization in container terminals are of great significance to promote the early

realization of carbon neutralization in the port industry and deal with global climate change. The optimization path of energy structure and the realization of carbon peak and carbon neutralization of China's container terminal will provide global container terminals a template for the transformation and upgrading, and energy conservation and carbon reduction. At present, the port industry is in the historical stage of high-tech development strategy. China's container terminal aims to achieve the carbon peak of container terminal by 2030, which requires to:

- (1) apply new energy technology, and promote the application of clean energy technologies for port machinery and vehicles.
- (2) apply intelligent technology to container terminal production, and improving the productivity of fuel machinery and vehicles.
- (3) promote the improvement of container terminal Automation and electro gasification production process, and use of high-efficient, low-carbon fuel machinery and vehicles.

References

- [1] X J Wu, Research on the development of low-carbon and green container port in China, Master's thesis, Dalian Maritime University, Dalian, China, 2015.
- [2] Q Y Li. Carbon emission status and carbon reduction path analysis of waterway transport in China [J]. Transport Energy Conservation & Environmental Protection, 2021, 17 (2):1-4 (in Chinese).
- [3] Q Y Wang. 2019 China energy statistics [M]. Beijing: IGDP, 2019 (in Chinese).
- [4] H L Wang, J K He. Peaking rule of CO₂ emissions, energy consumption and transport volume in transportation sector [J]. China Population, Resources and Environment, 2018, 28 (2):59-65 (in Chinese).
- [5] H L Wang. Model simulation of China's Low-carbon transformation mechanism and policy [D]. Beijing: Tsinghua University, 2017 (in Chinese).
- [6] D Mccollum, C Yang. Achieving deep reductions in US transport greenhouse gas emissions: scenario analysis and policy implications [J]. Energy Policy, 2009, 37 (12): 5580-5596.
- [7] X D Qi, X J Li, B W Zhang. Cost effectiveness analysis on emission reduction of greenhouse gas in China: taking electric vehicle as example [J]. Technology Economics, 2017, 36 (4):72-78 (in Chinese).
- [8] B J Tang, X Y Li, B Yu, et al. Sustainable development pathway for intercity passenger transport: A case study of China [J]. Applied Energy, 2019, 254: 113632.
- [9] Z Y Yuan, Z Y Li, L P Kang, et al. A review of low-carbon measurements and transition pathway of transport sector in China [J]. Climate Change Research, 2021, 17 (1): 27-35
- [10] D Xu. Analysis of advantages and disadvantages of hydrogen energy application in port, Water transportation in China, 2020(11)-133.
- [11] Z H Feng, X C Wang, H Y Zhang, et al. Path and policy of green transportation development from low carbon perspective [J]. Transport Research, 2019, 5(4): 37-45.
- [12] Z Ahoura, H Shima, F Nicole. High-tech business location, transportation accessibility, and implications for sustainability: Evaluating the differences between high-tech specializations using empirical evidence from U.S. booming regions [J]. Sustainable Cities and Society: 2019,50: 101648.
- [13] W M Wey, J Y Huang. Urban sustainable transportation planning strategies for livable city's quality of life [J]. Habitat International, 2018, 82: 9-27.
- [14] Y Yang, C Wang, W L Liu, et al. Microsimulation of low carbon urban transport policies in beijing [J]. Energy Policy, 2017, 107: 561-572.
- [15] H Hao, Y Geng, X Ou. Estimating CO₂ emissions from water transportation of freight in China [J]. International Journal of Shipping & Transport Logs, 2015, 7 (6): 676-694.
- [16] K Jiang, X Tian, J Xu. China's green economic transition toward 2049 [R]. Washington DC: Brookings Institution, 2020.

Development and Implementation of High Precision Positioning Technology for Electric Vehicle Wireless Charging Coils

Jie LI^a, Kejun QIAN^a, Yafei LI^a, Yi LIU^a, Zifeng LIU^{b 1}, Linlin TAN^b

^aJiangsu Electric Power Company Suzhou Power Supply Company, Suzhou, China

^bSchool of Electrical Engineering, Southeast University, Nanjing, China

Abstract. This paper mainly studies the development and implementation of the positioning technology of the electric vehicle wireless charging coil, so as to accurately detect the position deviation of the receiving coil, so that the electric vehicle wireless charging system can provide electric energy for electric vehicles more efficiently. Based on the positioning method of electric vehicle based on three detection coils, this paper proposes a calculation method to describe the offset degree of coil based on fuzzy mathematics theory. The algorithm is verified by simulation and experiment, and the influence factors of the error accuracy and the source of the error are analyzed. The work done in this paper has a strong practical significance for the efficient realization of electric vehicle wireless energy transmission.

Keywords. electric vehicle wireless charging, three detection coils, fuzzy mathematics theory, coil offset

1. Introduction

According to a report by MarketsandMarkets, a market research and consulting firm, by 2020, wireless charging for electric vehicles has generated a staggering \$13.78 billion in revenue worldwide [1]. At present, it seems that EV charging technology still has many technical bottlenecks: Compared with traditional cars, the charging duration and range of ELECTRIC vehicles are urgent problems to be solved, and there is no mature technical scheme for high-power charging either at the end of the power grid or at the end of electric vehicles. The huge gap with traditional vehicles in the number and scale of gas stations and other infrastructure is the bottleneck that electric vehicles must overcome. Charging facilities need to be built reasonably and effectively to overcome the range anxiety of electric vehicles.[2] Therefore, how to achieve efficient charging is now one of the core issues attracting much attention. Wireless charging of electric vehicles is non-

¹Corresponding author: Zifeng Liu, School of Electrical Engineering, Southeast University, Nanjing, People's Republic of China; E-mail: 1282402375@qq.com.

contact charging, so the charging coil can be buried underground to save space. Another advantage is that it is convenient to fix the transmitting coil underground to prevent unnecessary position shift of the transmitting coil due to various factors. The receiving coil can be fixed on the chassis of the electric vehicle. The electric energy is transmitted through the electromagnetic coupling between the transmitting coil and the receiving coil. In general, the transmitting coil of wireless charging is placed underground, and in some special cases, it can be placed on the ground, provided that strict height calibration is required [3]. The premise of efficient charging is the need to achieve accurate alignment between the transmitting coil and the receiving coil.

The mismatch caused by coil offset will have a significant impact on the transmission efficiency of electric energy, and may lead to the instability of resonant wireless energy transmission system. When the coil mismatch increases, the mutual inductance value between induction coils will decline sharply. The coils need to be nearly perfectly aligned to keep the power transmitted wirelessly within an acceptable range. Generally, the maximum allowable offset of the radio energy transmission system is 10cm, and the power transmission efficiency within the offset range at this position is within the acceptable range. With the increase of the offset, the transmission efficiency will decline sharply. However, some studies show that most drivers cannot accurately stop within the error allowable range. In addition, the average offset of parking will exceed 70cm when the driver has no knowledge of any external device [4]. Therefore, the offset of the coil needs to be detected and corrected to maintain acceptable power transfer efficiency.

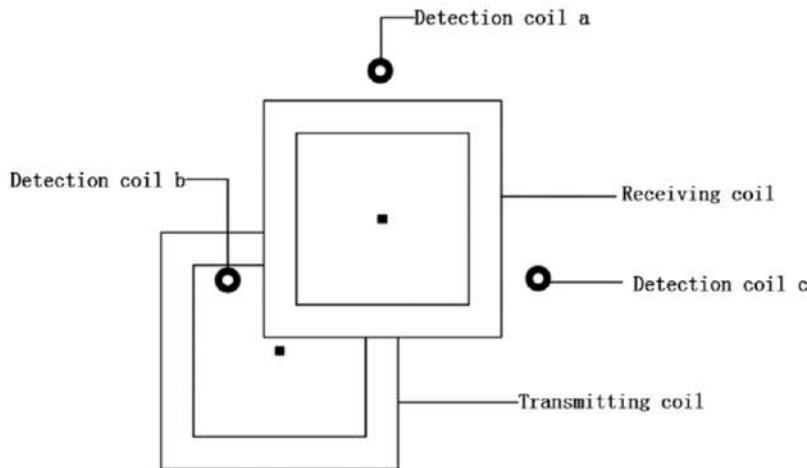


Figure 1. Schematic diagram of balance coil positioning scheme system.

The electric vehicle charging positioning has a lot of the existing research results, China first automobile limited liability company is proposed to increase the four testing coil, to distinguish the ground end and on-board the relative position [5]. Huazhong University of Science and Technology proposed a method of coupling the active excitation of secondary side coil with primary side coil to realize real-time position detection of vehicles by measuring the current variation in primary side coil [6]. Zhejiang University proposed an online wireless charging positioning system for electric vehicles, which made the output voltage on the differential inductor coil by coupling the secondary

power coil after energization, and then obtained the relationship table between the output voltage of the differential inductor coil and the position of the secondary coil through a large number of experiments, so as to determine the position of the vehicle [7]. A position detection scheme for dynamic wireless charging system of electric vehicles based on magnetic positioning technology was designed in Beijing Jiaotong University [8]. Allan et al. from the International Society of Automotive Engineers have proposed a method for detecting the position of charging target devices in parking lot scenarios by evaluating the received signal strength indication of dedicated short-range communication signals and global positioning system data [9].

The main work content of this paper is to investigate the existing methods of electric vehicle wireless charging positioning, introduce the working principle of electric vehicle wireless charging positioning system and expatiate on its research significance. Analyze the electromagnetic model and circuit model of wireless charging coil positioning system, analyze the models of transmitting coil, receiving coil and detecting coil in the wireless charging system, design the parameters of detecting coil, and analyze the LCC-S circuit model of the wireless energy transmission system. The theoretical feasibility of wireless charging positioning method based on three detection coils is verified. Figure 1 shows the schematic diagram of balance coil positioning scheme system.

On the basis of the positioning method of electric vehicle based on three detection coils proposed in this paper, a calculation method based on fuzzy mathematics theory is put forward to describe the offset degree of the coil. The algorithm is verified by simulation, and the simulation results verify the effectiveness of the proposed method. The simulation results show that the smaller the mesh width is, the higher the positioning accuracy is. Then the influence factors of error precision and the source of error are analyzed.

2. Characteristic Analysis of Electric Vehicle Wireless Charging Coil System

2.1. Analysis of Basic Model of Coupling Coil

The energy transmission of the electric vehicle wireless energy transmission system is realized by the electromagnetic coupling between the receiving coil and the transmitting coil. The efficiency of the system's energy transmission is closely related to the coupling coefficient between the coils. In this chapter, the coupling coil model is firstly analyzed, and then the resonant compensation network LCC-S circuit model is analyzed. ANSYS Maxwell electromagnetic simulation software was used for auxiliary calculation.

2.1.1. Analysis of Basic Electromagnetic Model of Rectangular Coil

The mutual inductance between rectangular coils has a great influence on the transmission efficiency of electric vehicle wireless charging system. For the mutual inductance between them, Niemann formula can be used for derivation. The Niemann formula is as follows:

$$M = \frac{\mu_0}{4\pi} \oint p \oint d\vec{a} \frac{dp}{r} \quad (1)$$

Figure. 2 shows the schematic diagram of mutual inductance calculation of rectangular coil:

In the figure, coil 1 represents the transmitting coil, and coil 2 represents the receiving coil. It can be seen that the transmitting coil $\llbracket O \rrbracket_1$ is located at the origin of the spatial coordinates, and the coordinate of the receiving coil O_2 is (p, q, h) . The two sides of the transmitting coil and the receiving coil can be named as ABCD, and the length of the side is represented by m and n . Then the formula (2) can be obtained from Equation (1):

$$M = \frac{\mu_0}{4\pi} \left(\iint \frac{d\vec{A}_1 \cdot d\vec{A}_2}{R_1} + \iint \frac{d\vec{A}_1 \cdot d\vec{C}_2}{R_2} + \iint \frac{d\vec{C}_1 \cdot d\vec{A}_2}{R_3} + \iint \frac{d\vec{C}_1 \cdot d\vec{C}_2}{R_4} \right. \\ \left. + \iint \frac{d\vec{B}_1 \cdot d\vec{B}_2}{R_5} + \iint \frac{d\vec{B}_1 \cdot d\vec{D}_2}{R_6} + \iint \frac{d\vec{D}_1 \cdot d\vec{B}_2}{R_7} + \iint \frac{d\vec{D}_1 \cdot d\vec{D}_2}{R_8} \right) \quad (2)$$

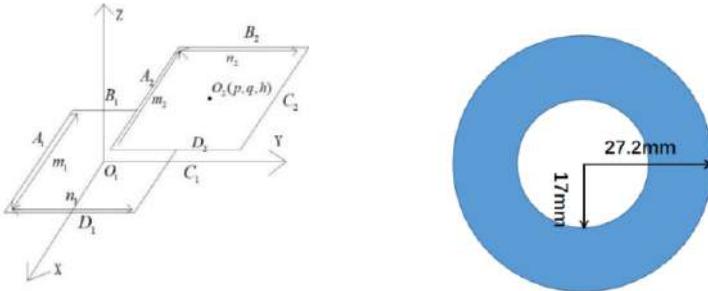


Figure 2. Rectangular coil electromagnetic model **Figure 3.** Rectangular coil electromagnetic model

After obtaining (2), we only need to calculate each term in (3) respectively to obtain the mutual inductance expression between the single-turn rectangular transmitting coil and the receiving coil. However, this expression is very complex and troublesome to calculate, so Maxwell electromagnetic simulation software is used for auxiliary calculation below.

2.1.2. Coil Parameter Design

In this paper, the wireless charging positioning of electric vehicles is obtained by detecting the inductive voltage of the detection coil, which directly determines whether the final positioning can be completed. Therefore, it is very important to design the parameters of the detection coil.

First on detecting coil wire material choice, for wireless transmission system, the size of the coil impedance has a great influence for the transmission efficiency, the actual situation should try to choose low resistance material testing coil, copper wire, generally is the first selection of conductor material, by the theory of electromagnetic field, the copper coil of the resistance by the loss of resistance and high frequency radiation resistance, expression is as follows:

$$R_0 = \sqrt{\frac{\omega\mu_0}{2\sigma}} \frac{l}{2\pi a} = \sqrt{\frac{\omega\mu_0}{2\sigma}} \frac{N_w r}{a} \quad (3)$$

$$R_r = \sqrt{\frac{\mu_0}{\epsilon_0}} \left[\frac{\pi}{12} N_w^2 \left(\frac{\omega r}{c} \right)^4 + \frac{2}{3\pi^2} \left(\frac{\omega h}{c} \right)^2 \right] \quad (4)$$

Where μ_0 is the vacuum permeability, ϵ_0 is the vacuum dielectric constant, σ is the coil conductivity, N_w is the number of coil turns, and h is the coil turn spacing. c is the speed of light in vacuum, l is the length of the wire, and a is the radius of the wire section. For wireless charging systems for electric vehicles, the American Society of Automotive

Engineers (SAE) has published a report on guidelines for wireless charging for electric vehicles SAE-J2954. The working frequency of electric vehicle wireless charging system is 85kHz [10], and the high-frequency radiation resistance in the general coil can be ignored.

Since the coupling coefficient of the circular coil is larger under the same conditions and can better reflect the voltage change after the receiving coil offset, the circular coil is selected as the detection coil for wireless charging positioning [11]. In other conditions are equal, the greater the area of the coil, the coil of the coupling coefficient will lead to greater, but this part of the design for electric vehicle charging positioning system in detecting coil, if is too big, can affect the positioning accuracy, so the coil area is not too big, so the preliminary design for detecting coil is composed of ring, outer ring is 27.2 mm, inner ring is 17 mm. As this experiment is an approximate qualitative analysis, although the number of coil turns will improve the overall accuracy, it will also make the simulation time too long. Therefore, copper ring is chosen to replace the multi-turn copper wire arrangement to simulate the detection coil. The final test coil design diagram and parameter table are shown in the following figure and table 1:

Table 1. detection coil design parameter sheet.

name of parameter	parameter
shape	circular ring
texture	copper
external diameter	27.2mm
inner diameter	17mm
number of turns	10

2.2. Analysis of Electromagnetic Model of Detection Coil

In this paper, a positioning method based on three coils is proposed. Different shapes, sizes and inductance values of coils will affect their induced voltage. In order to reduce the interference of other variables on this study, three detection coils with the same shape, size and inductance values are selected, whose parameters are shown in Section 2.1.2.

One of the three coils, coil A, is placed on the straight line of the central axis of the rectangular coil, and the other two coils, coil B and coil C, are placed around the two adjacent sides of the nearest edge of the rectangle and coil A, and the three coils are placed in an equilateral triangle. The three detection coils are placed outside the receiving coil and at the same plane (same height) with the receiving coil. The magnetic field excited by the transmitting coil at a certain height is approximately circular in the XOY plane. Now, the origin is set at the position coincident with the center of the transmitting coil, and the polar coordinate system is established. The mutual inductance is expressed as:

$$M(r) = C \quad (5)$$

In the above formula, C is a constant, indicating that the magnitude of mutual inductance is only related to the distance r , and the parameter related to the center location of the coils on both sides is a monotone function r , if the center position of the receiving coil is at the point $P(r_1, \theta_1)$, without considering the case of the deflection of the coil Angle.

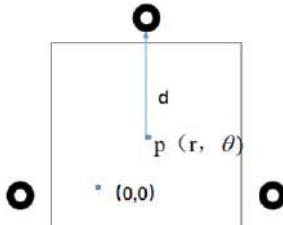


Figure 4. Schematic diagram of positioning system based on three detection coils.

Figure 4 shows the schematic diagram of positioning system based on three detection coils. d is the distance between the three detection coils and the center of the receiving coil, and the three detection coils are arranged in an equilateral triangle. After the detection coil is positioned, the inductive voltage is proportional to the mutual inductance value, so the inductive voltage is only related to the distance between the coordinate plane and the detection coil, and it is a monotone function. If the distance between one point and three detection coils is a constant value in the plane, the positioning can be achieved with this method, and the proof is as follows:

Given a certain distance to any point P in the plane to form a circle with P as the center; An ellipse with P and Q as focal points is constituted by any point on the ellipse (which is not on the line PQ). If M is any point on the ellipse (which is not on the line PQ), then there exists and only exists a symmetric N point on the line PQ , so that the distance between P (Q) and M and N is equal. Therefore, there is one or two points in the plane whose distance to any two known points is constant; In the equilateral triangle plane, if any point M is taken, it can be found that the M point can respectively find three points symmetric about the three sides, and these three points cannot coincide. Therefore, there exists and only one unique point on the plane whose distance to the three sides of the equilateral triangle is a fixed value.

What has been discussed above, there is no two different points at the same time make the detection coil voltage vector fully equal, so the reference point of three-dimensional coordinate information and their corresponding detecting coil voltage information into a database, and then at the time of actual location, will be the voltage vector and reference voltage vector in the database, The approximate position of electric vehicle wireless charging receiving coil can be estimated by lookup table method. The specific exact position needs to be realized by a certain algorithm, which will be introduced below.

3. Algorithm for Electric Vehicles Wireless Charging Positioning

The second chapter studies the electric vehicle wireless energy transmission positioning system model based on three coils and the feasibility of three coils positioning method, this chapter will study the offset detection algorithm.

It has been explained in Section 2.2 that when the system is working, the transmitting coil inputs excitation of a certain size to generate mutual inductance between the two coils, and inducted voltage is generated on the receiving coil side. The inducted voltage can be obtained by the following formula:

$$u = \omega M \quad (6)$$

As can be seen from the above equation, when the current at the side of the transmitting coil remains constant, the mutual inductance value and the induced voltage value show a linear relationship. Therefore, the mutual inductance value of the detecting coil can be directly detected to realize the detection of the voltage value. The change of mutual inductance value has a great coupling relationship with the relative position of the transmitting coil and the receiving coil. Therefore, the offset of the coil position can be obtained through a certain coupling relationship only by detecting voltage.

In detecting migration cases before the induction voltage coil need to use wireless charging system between the two coil plane into a grid, and grid intersection set as the reference point, secure a receiving coil in each reference point three detecting coil induced voltage, when stored in three dimensions for $n * n$ numerical matrix, when the actual location, If the receiving coil is offset to a certain position, the proximity of the position to be measured to all the reference points can be described by the membership mathematical model in the fuzzy mathematical theory. So the algorithm of this paper is based on the membership mathematical model of fuzzy mathematics theory. By traversing the residual point voltage matrix obtained before, the four closest points can be found, and then the approximate range of the anchor point coordinates can be obtained. The following is a detailed introduction of the specific calculation method.

When the electric vehicle wireless charging positioning system is working, the three detection coils are kept at the same height and plane with the receiving coil, in an equilateral triangle distribution, and the relative position with the receiving coil is fixed. Firstly, the center point of the receiving coil is placed at each reference point, and the effective values of the three receiving coil voltages are recorded as u_{p1}, u_{p2}, u_{p3} , and these values are stored in the matrix.

In practice, when the position of the receiving coil is offset, the voltage values of the three detection coils can be obtained. However, since the center point of the receiving coil cannot accurately fall on the reference point defined in this paper after every position offset, the accurate offset of the coil cannot be obtained. Instead, the proximity of the center of the receiving coil to several surrounding reference points after each position offset can be described. Set:

$$\begin{cases} u_1 = u_{p1} - u_{r1} \\ u_2 = u_{p2} - u_{r2} \\ u_3 = u_{p3} - u_{r3} \end{cases} \quad (7)$$

Where, u_r is the voltage value of the grid point measured previously, namely the reference point; u_p is the inductive voltage of the detection coil measured after the actual position of the receiving coil is offset; ΔU is used to describe the proximity of the center point of the receiving coil to the reference point of meshing, which is described by membership function in this paper. According to the previous analysis in Section 2.2, the smaller ΔU is, the closer the center point of the receiving coil is to the sub-meshed reference point. In this paper, a membership function model with small distribution is selected. δ can be defined by the membership function of normal distribution:

$$\delta = e^{-a\Delta U^2} \quad (8)$$

After obtaining the voltages of the three detection coils at a certain position, the membership function is used to obtain the δ of each reference point, thus an $n * n$ matrix can be obtained:

$$\Delta = \begin{bmatrix} \delta_{11} & \cdot & \cdot & \cdot & \delta_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \delta_{n1} & & & & \delta_{nn} \end{bmatrix} \quad (9)$$

Due to δ describes specific with all the grid reference point proximity, therefore, its size and specific only related to the distance between the reference point, the reason for out of four is the maximum and the nearest grid reference point, specific ergodic matrix of $n * n$ elements, can get δ values of the four largest elements, write them respectively for the $\delta_A, \delta_B, \delta_C, \delta_D$, These four points, theoretically represent the four meshing reference points closest to the center point of the receiving coil.

Since the meshing reference points were originally designed, the coordinates of points A, B, C and D are known, at this point, the coordinate of the center position of the receiving coil can be obtained from the coordinates of the four meshing reference points and their proximity δ , and the calculation formula is:

$$\begin{cases} x = \frac{x_a\delta_a + x_b\delta_b + x_c\delta_c + x_d\delta_d}{\delta_a + \delta_b + \delta_c + \delta_d} \\ y = \frac{y_a\delta_a + y_b\delta_b + y_c\delta_c + y_d\delta_d}{\delta_a + \delta_b + \delta_c + \delta_d} \end{cases} \quad (10)$$

The obtained x and y are respectively the abscissa and ordinate of the center point of the receiving coil to be solved, so as to realize the positioning of the electric vehicle wireless charging coil. In the following chapters, the algorithm will be used to locate the positioning coil in the wireless charging system of electric vehicles, and the accuracy of the algorithm will be analyzed.

4. Simulation and Experimental Verification of Electric Vehicle Wireless Charging Positioning System

According to the analysis in the previous chapter, for the wireless charging positioning system of electric vehicles, if the current of the transmitting coil and the system frequency are constant, then the position change of the receiving coil can be represented by the voltage change of the three detection coils through a certain algorithm to realize the positioning of the coil of the wireless transmission system of electric vehicles.

In this chapter, relevant simulation, experiment and error analysis will be carried out for the positioning system of electric vehicle wireless energy transmission coil under the condition of asynchronous length, and the positioning accuracy and error sources will be analyzed.

4.1. Simulation Results of Three Detection Coil Positioning System

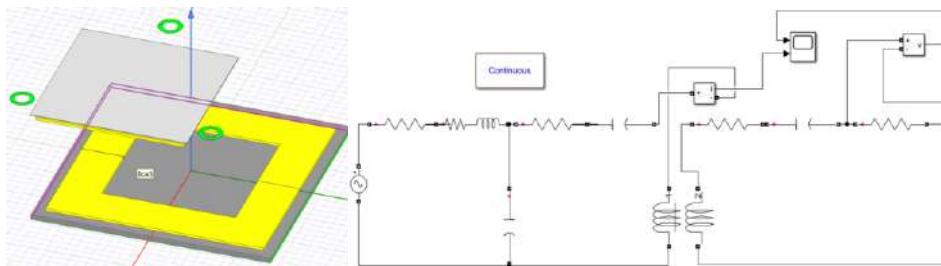


Figure 5. Electromagnetic model of electric vehicle wireless charging. **Figure 6.** LCC-S circuit simulation model.

As shown in Figure 5, the wireless charging positioning system of electric vehicles consists of transmitting coil, receiving coil and three detection coils of exactly the same size and shape. The center point of the transmitting coil in the figure is fixed at the origin of coordinates, and the side length of the square transmitting coil is 50cm. The position of the receiving coil can be moved in the same height plane. The side length of the square receiving coil is 30cm and the height is 17.54cm. The parameters of the detecting coil are mentioned in Section 2.1.2.

In the model established in this paper, the size of the transmitting coil is designed to be 50cm*50cm, the number of turns is 10, and the current is 24.9A. The receiving coil is 30cm*30cm, and the number of turns is 10. The detection coil and the receiving coil are in a height plane, and the three detection coils are arranged around the receiving coil in an equilateral triangle.

Firstly, a region in the receiving coil plane is gridded and $n \times n$ reference points are obtained. In this paper, the size of this region is designed as 35*35cm. Three groups of simulations are carried out, and the step sizes are set as 2cm, 3cm, and 4cm respectively, that is, the distance between two adjacent gridded reference points is 2cm, 3cm, and 4cm. Then the model was simulated, and the step size was set to 1cm during the simulation, that is, a grid of 35*35 could be obtained. Data processing was done on this grid, that is, one of every or several points was taken as the reference point to obtain three groups of data, which were stored in three tables. They respectively represent the mutual inductance between the three detection coils and the transmitting coil when the center point of the receiving coil is located at each grid reference point

The algorithm of this paper is by detecting coil and the transmitter coil between the induced voltage to achieve localization work, therefore need to convert the mutual inductance existing data into data induction voltage, the current value can be built with the help of the LCC-S model simulation in MATLAB, the circuit model is as shown in figure 6.

In the simulation model of LCC-S circuit, the peak value of the AC power supply is 380V, the frequency is 85kHz, the series resistance of the primary side is 0.2Ω, the inductance is 0.02mH, the shunt capacitance of the primary side of the circuit is 0.175μF, and the series capacitance of the secondary side is 0.0377μF. At this point, the current value I of the circuit flowing through the primary coil can be obtained. Using the formula:

$$u = 2\pi \times 85000 \times 32 \times M \div \sqrt{2} \quad (11)$$

The three groups of data obtained previously were calculated according to the above formula, and the induced voltage of the three detection coils when the center point of the receiving coil is located at each meshing reference point could be obtained. They were stored in the three matrices to facilitate positioning calculation by using the algorithm designed. The simulation results are shown in Table 2.

The experimental device is shown in the figure 7, and it can be concluded that the algorithm in this paper can have small errors in both simulation and experiment.

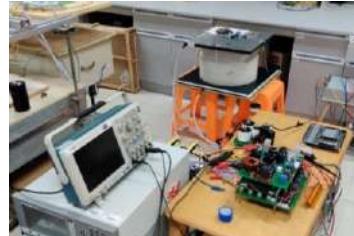


Figure 7. Electric vehicle wireless charging positioning system experimental device diagram.

Table 2. Ten groups of data with step sizes of 2cm, 3cm and 4cm.

The simulation serial number	The theoretical position in 18*18 grid	The actual position in 18*18 grid	Error (cm)
1	(1.5,4)	(1.49,4)	0.1
2	(7.5,7)	(7.55,6.9)	0.13
3	(5.5,10.5)	(5.78,10.41)	0.57
4	(2,16)	(2.02,16.03)	0.05
5	(17.5,8)	(17.10,7.61)	1.12
6	(2.5,10.5)	(2.45,10.66)	0.34
7	(10.5,7.5)	(10.61,7.35)	0.37
8	(8.5,17.5)	(9.10,17.45)	1.2
9	(9,2)	(8.82,2.39)	0.86
10	(7.5,5.5)	(7.58,5.08)	0.86

The simulation serial number	The theoretical position in 12*12 grid	The actual position in 12*12 grid	Error (cm)
1	(1.33,3)	(1.23,2.83)	0.59
2	(5.33,5)	(5.1,4.98)	0.69
3	(4,7.33)	(4.52,7.11)	1.69
4	(1.67,11)	(1.83,11.1)	0.57
5	(12,5.67)	(11.53,5.42)	1.6
6	(2,7.33)	(1.85,7.3)	0.46
7	(7.33,5.33)	(7.05,5)	1.3
8	(6,12)	(6.03,12)	0.09
9	(6.33,1.67)	(6.19,2.02)	1.13
10	(5.33,4)	(5.34,3.59)	1.23

The simulation serial number	The theoretical position in 9*9 grid	The actual position in 9*9 grid	Error (cm)
1	(1.25,2.5)	(1.23,2.68)	0.72
2	(4.25,4)	(4,4)	1
3	(3.25,5.75)	(3,6)	1.41
4	(1.5,8.5)	(2.17,7.83)	3.82
5	(9.25,4.5)	(8.98,4)	2.27
6	(1.75,5.75)	(2.95,5.03)	5.6

7	(5.75,4.25)	(6,4)	1.41
8	(4.75,9.25)	(5.51,9)	3.2
9	(5,1.5)	(4.96,1.57)	0.32
10	(4.25,3.25)	(4,3.23)	1

4.2. Error Analysis of Simulation Results of Detection Coil Positioning System

This paper simulates the situation that the receiving coil is located in ten different positions. That is, 10 serial numbers are input into Visual C++ to simulate the real position of the receiving coil in the electric vehicle wireless energy transmission system, and the theoretical and actual positions of ten groups of data can be obtained in 18*18, 12*12 and 9*9 grids respectively. Their positioning error sizes were analyzed and compared, and four ranges with error sizes of less than or equal to 1cm, 1cm-2cm, 2cm-3cm and greater than 3cm were selected respectively. The results are shown in Table 3.

Table 3. Simulation error statistics under asynchronous length.

error magnitude	less than or equal to 1cm	1cm-2cm	2cm-3cm	greater than
step length 2cm	8	2	0	0
step length 3cm	5	5	0	0
step length 4cm	4	2	1	3

As can be seen from Table 3, when the step size is 2cm, the errors of the ten groups of tests are all within 2cm, and the accuracy of this algorithm reaches the expected standard. The experiment in this paper was only carried out when the step size was 2cm, and the results showed that the ten groups of data were all within the allowable error range. It can be seen that both simulation and experiment can verify the correctness of the algorithm in this paper. Therefore, this algorithm is feasible. When the step size becomes larger, the positioning error will become larger. The following paragraphs will speculate on the error sources of the algorithm.

By analyzing the algorithm, electromagnetic model and data designed in this paper, it can be concluded that the reasons for errors are as follows:

① The algorithm in this paper is established based on the membership mathematical model in fuzzy mathematics theory. δ represents the degree of proximity between the point to be solved and the gridded reference point, which is an inaccurate value. Therefore, the coordinate calculated by multiplying the reference point by their respective degree of proximity is also inaccurate and may cause errors.

② In the electromagnetic model of the wireless charging positioning system of electric vehicles in this paper, the inductive voltage generated by the excitation of the transmitting coil in the three detection coils differs greatly, which may be related to the placement of the three detection coils. In the final calculation, the three groups of data were normalized. That is, the three groups of data are changed into three groups of approximately equal size through scaling and scaling. Although scaling and scaling do not change the magnitude relationship of the original induced voltages at several grid reference points, the accuracy will therefore change. So the error is going to go up.

To sum up, this paper proposes and implements a three-coil based wireless charging positioning algorithm for electric vehicles. The inductive voltage data are processed by the membership degree mathematical model in fuzzy mathematics theory, and the feasibility of the scheme is verified by simulation.

5. Conclusion

Aiming at the requirement of deviation degree of electric vehicle wireless charging coil, this paper designed a positioning algorithm of electric vehicle wireless charging based on three coils. This algorithm was designed based on the membership mathematical model in fuzzy mathematics theory. In this paper, theoretical analysis, simulation analysis and experiment are carried out on the electric vehicle wireless charging system model, and error analysis is carried out on the simulation results. When the step size is 2cm, the errors of ten groups of tests are all within 2cm. The accuracy of this algorithm reaches the expected standard, and the algorithm meets the working requirements of the system, and the scheme is feasible.

References

- [1] Marketsand-Markets. Wireless Charging Market Analysis and Forecast by Technology, Range, Components (Transmitter and Receiver for Various Devices), Uses (Appliances, Defense, Medical, Automotive, Industrial), and Region. [J].2014
- [2] Wu J, Liu X, Brief analysis of electric vehicle charging technology and electric change technology[J]. Automobile practical technology, 2021,46(12):8-10.
- [3] Wireless Power Transfer for Light-Duty Plug-In/Electric Vehicles and Alignment Methodology, Standard SAEJ2954, May 2016. [online]. Available: http://standards.sae.org/j2954_20120122.html.
- [4] Birrell S A, Wilson D, Yang C P, et al. How driver behaviour and parking alignment affects inductive charging systems for electric vehicles[J]. Transportation Research Part C, 2015, 58:721-731.
- [5] Yan M S, Research on Location Technology of Electric Vehicle Wireless Charging System Based on Magnetic Field Induction[D]. Beijing: University of Chinese Academy of Sciences, 2014.
- [6] Zhao J B, Cai T, Duan S X, et al. Relay method suitable for piecewise dynamic wireless charging[J]. Automation of Electric Systems. 2016, 40(16): 64-70.
- [7] Cheng Z Q, Chen H, Wu J D, HE X N. An online wireless charging and positioning system for electric vehicles based on differential inductance[P]. Zhejiang Province: CN109895643A,2019-06-18.
- [8] Fang Q Q. Research on Dynamic Wireless Power Supply System Control of Electric Vehicle Based on Magnetic Positioning Detection Technology[D]. Bejing: Beijing Jiaotong University, 2017, 18-53.
- [9] Allan L, Mohammad N. Wireless Charging System Localization for Electric Vehicles Using RSSI[J]. SAE International Journal of Passenger Cars - Electronic and Electrical Systems, 2015, 8(2):2015-01-0283
- [10] http://www.sae.org/servlets/pressRoom?OBJECT_TYPE=PressReleases&PAGE=showRelease&RELEASEID=109895643A.
- [11] Zhai Z, Sun Y, Dai X, et al. Modeling and analysis of magnetic resonance mode radio energy transmission system[J]. Proceedings of the CSEE, 2013,32(12):155-160.

Study on Type Selection of Relays in Outdoor Terminal Box of Converter Station

Jianying WU¹, Zhen WANG, Xilin YAN

Maintenance & Test Center, CSG EHV Power Transmission Company, Guangzhou, Guangdong Province, 510663, China

Abstract. The outdoor terminal box of converter station's complex working characteristics and working environment easily leads to the failure of relays in it. The paper analyzes the key factors in relay selection from the perspectives of the characteristics of the operating environment. By comparing the characteristics of different component materials and structure designs, the paper proposes suggestions of type selection of relays in outdoor terminal boxes of converter stations.

Keywords. relays type selection, outdoor terminal box, component materials, structure design

1. Introduction

The outdoor terminal box of the converter station is vital in connecting and transferring between the primary equipment and the secondary equipment [1]. Secondary devices such as relays are usually placed in it. As the outdoor terminal box is directly exposed to the external environment without house protection, it is extremely easy to make the relays suffer from damages resulted from dust, moisture, temperature and other factors. When selecting the type of relays, the working environment, component material and structure designs should be considered. Consequently, the relays can be selected reasonably, used properly and related technical performance requirements could be met.

At present, there is scarcely any research on the type selection of relays in outdoor terminal box. The existing research largely analyzed the characteristics of relay failure and studied the measures. It describes the system principle of the relay and the selection of quality grade in the relay application and the principle of the relay derating use in reference [2]. Literature [3] analyzed the influence of ambient temperature, pressure and mechanical stress on the relay and proposed requirements for electrical parameters of relays. Literature [4] analyzed the electrical characteristics of the relays and put forward reliable use measures in coil connection, contact connection and arc suppression. Literature [5] proposed that the parallel contact structure can be used to improve the reliability of relay connection, and the series current-limiting resistance at the output of the contact can inhibit the impulse current.

The researches above make progress on analyzing the mechanism of relay failure, but there are unsolved issues as follows: 1) No comprehensive analysis of the operating

¹ Corresponding author: Jianying Wu, Maintenance & Test Center, CSG EHV Power Transmission Company, No.223, Kexue Avenue, Huangpu District, Guangzhou, Guangdong Province, China; E-mail: wujaneyee92@163.com.

environment stress of relays in outdoor terminal box. 2) No proposed approaches in relay type selection, owing to the complexity of the working environment.

That being said, this paper studies the operating environment characteristics of the outdoor terminal box and its influences on relays. Then the paper analyses the application conditions of different component materials and structure designs, and proposes the approach in type selection of relays to meet outdoor terminal boxes' environmental adaptability.

2. Environmental Analysis of Relays in Outdoor Terminal Box

2.1. Analysis of Operation Environment

The outdoor terminal box of converter station mainly includes CT, PT, circuit breaker, transformer control terminal box, etc. In most cases, the box is made of stainless steel with poor thermal insulation performance, fluctuating temperature and bad sealing [6]. In the DC converter stations, especially in Southern China, the terminal boxes are exposed to high temperature, damp air and dust, which are prone to lead to the potential safety concerns of the relays in the boxes, such as the deteriorated insulation performance, the components corrosion, the aging materials, and even the relay malfunction.

- Temperature and humidity

Converter stations were usually built in mountainous and humid areas, especially in Guangdong with an average air humidity above 77%, due to the abundant rainfall. The temperature and humidity curves of an outdoor terminal box connected with a cable trench in Qiao Xiang Station are shown in Figure 1.

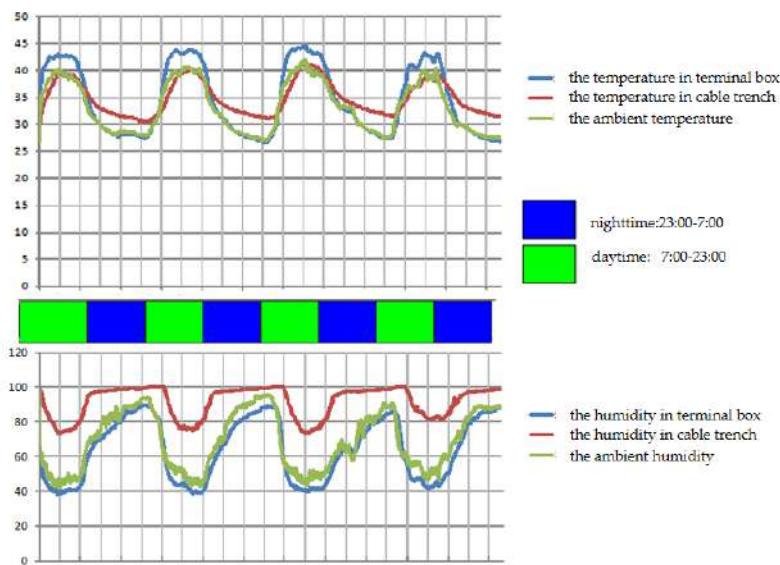


Figure 1. Temperature and Humidity Test Curve of Terminal Box in Qiao Xiang Station.

Test results of the temperature and humidity of the terminal box in the converter station showed that the relays mainly operate in the fluctuating high-temperature and

high-humidity environment. The maximum temperature is over 40°C and the maximum humidity is over 95%.

In rainy seasons, the humid air outside invades through terminal box doors, slots and fans. The secondary live equipment and heaters make the temperature inside the box higher. According to the relationship between saturated humidity of air and the temperature, if the relative humidity is high during the daytime, the moisture in the air saturates rapidly when the temperature drops at night [7]. As the shell of the terminal box dissipates heat faster, the air will condense at top of the terminal box or at the door with lower temperature.

A fan is installed in the terminal box to control internal temperature and draw out the moisture. Then the ventilation may carry external dust into the terminal box. There is also a risk of inlet external dust during open-box inspection and maintenance as Figure 2 shows.

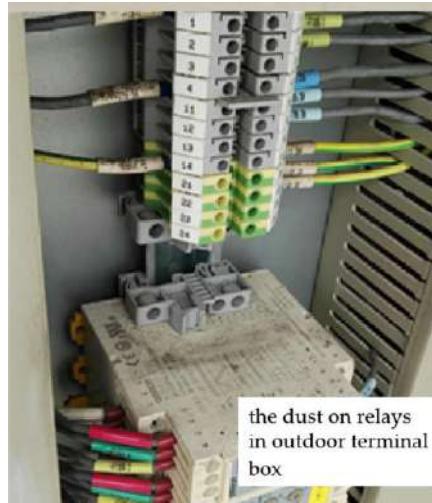


Figure 2. The dust accumulation on the surface of relays.

As it's easy to produce condensed water in the fluctuating high-temperature and high-humidity environment, the dust and the water mixed on the surface of the relays may lead to flashover or short-circuit.

2.2. Influence of Environmental Parameters on Relays

The national standard stipulates the requirements of standard limit values of the relays. The environmental temperature required in the national standard is -25 ~ 40°C, and the average temperature within 24 hours is below 35°C. The monthly average maximum relative humidity of the wettest month should be lower than 90% with the monthly average minimum temperature of 25°C, and there is no condensation on the surface. When the maximum temperature is 40°C, the average maximum relative humidity should be under 50%.

According to the environmental characteristics in outdoor terminal box, the ambient temperature and humidity in the terminal box of some converter stations may fail to meet the standard working conditions of relays. The main environmental parameters and impacts considered in the type selection of relays are shown in Table 1.

Table 1. The main environmental factors and impacts.

Environment	Appearance Failure Features	Electrical Failure Characteristic
high temperature	Softening and melting of insulating materials	The resistance of the coil and the absorption voltage increase. The arc breaking capacity is reduced
fluctuating high and low temperature	Loose structure, change position of moving parts	Poor electrical connection of the contact
humid	Electrochemical corrosion of coils and other metal parts. Plastic shell material aging failure.	The insulation performance decreases and the arcing phenomenon intensifies.
dust deposition	Dust deposits on the contact surface to produce black oxide.	The dust mix with condensation may cause insulation failure.

The main impacts of the environmental parameters on the relays in outdoor terminal boxes are appearance and electrical failure characteristics. The appearance failure characteristics are mainly shown in the metal corrosion of relays components, the aging plastic materials and the malfunction of contact materials. In the long term, appearance failure characteristics will develop into electrical failure such as insulation failures and discounted electrical life. According to the environmental parameters and failure features, the type selection of relays can be standardized from the aspects of relay component materials and structure designs.

3. The Material of Main Components

3.1. The Shell Material of the Relays

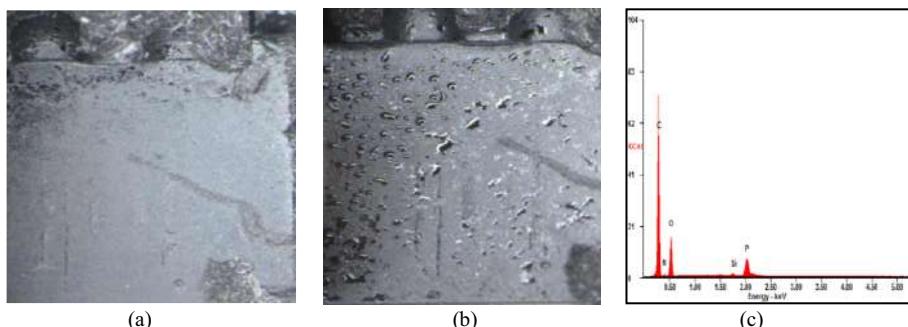
The shell material is vital for the basic protection of relays. The main components consist of plastic substrate and plastic additives. The main components of the rate substrate are made of thermoplastic materials, such as PBT, PET, PC, ABS and PS. The main components of plastic additives include phosphorus flame and bromine flame retardants.

Relay providers will adopt different materials depending on the users' scenario in most cases, and test on physical sustenance, safety features, temperature, and humidity. However, flame retardants with diversified ingredients and plastic substrate may lead to compatibility issues under exposure to high temperature and humidity, then the plastic additives gradually decomposed and precipitated to form oil conductive liquid [8]. The oil leakage and precipitated crystals on the plastic surface of the relays after long periods of operation, which led to risks including erosion of metal terminals, reduction in contact insulation and reliability. Unfortunately, mainstream relay manufacturers did not test on the exposure to the environmental conditions mentioned above.

85°C and 85% RH aging tests on the compatibility of various flame retardants and plastic substrate were conducted on the shell material components of typical relays' plastic shell. The results are shown in Table 2.

Table 2. Compatibility tests (85°C and 85% RH aging tests).

Plastic substrate	Components of flame retardants	Test results	Relay type
PBT+PA plastic alloy	Phosphorus flame retardants	Leakage of conducted oil liquid on relay's plastic surface.	Siemens 7PA 2732
PS	Small amounts of bromine flame retardants stuffed with glass fiber	No compatibility issues between additives and substrate.	Siemens 7PA 2730
PET	Small amounts of bromine flame retardants stuffed with glass fiber	Organic crystals with infinite resistance on relay's plastic surface.	Siemens 7PA2742, Schneider REL91233
PBT	Aluminum hydroxide flame retardants stuffed with glass fiber	No compatibility issues between additives and substrate.	Finder 56.34.9.110.4000

**Figure 3.** Changes of the plastic surface of sample relays before and after tests (Siemens 7PA2732) (a) Plastic surface of sample relays (pre-test) (b) Plastic surface of sample relays (post-test). (c)scanning electron microscope energy spectrum diagram of relay surface liquid.

Phosphorus flame retardants would leak, owing to easily decomposed features and poor compatibility with plastic substrates and other polymer materials [9]. The oil liquid precipitated in the test was detected by electron microscope energy spectrum analysis as Figure 3 shows. The results showed that the main elements of the oil liquid were carbon, oxygen, phosphorus and nitrogen. The oil liquid contained a large number of PO_4^{3-} through ion chromatography analysis. Tests showed that the oil liquid was conductive and was mainly decomposed from the phosphorus flame retardants. The leakage of oil liquid on the surface of relays made of phosphorus flame retardants and PBT+PA plastic substrate led to reduced shell insulation, increasing probabilities of electrochemical erosion on metal terminals, corrosion of terminals and even relay malfunction.

According to the experience and tests of retardants and substrates, relays with phosphorus flame retardants should be avoided considering the high temperature and humidity environments of outdoor terminal boxes. Manufacturers should be required to conduct 85°C and 85% RH aging tests on relays' plastic shell and provide related reports. Tests should proceed for more than 500 hours so that the reliability of relay's shell materials could be tested in exposure of high temperature and humidity.

3.2. Metal Terminal Materials

Terminals are mostly used as reliable physical and electrical connections among wires in relays. It consisted of conductive plates, screws and locator cards as Figure 4 shows. The analysis of the operational conditions of relays in terminal boxes showed that

condensed vapor was likely to drop on the relays, leading to the erosion of terminals. So it's vital to select metal terminals materials to improve the corrosion resistance of terminals.

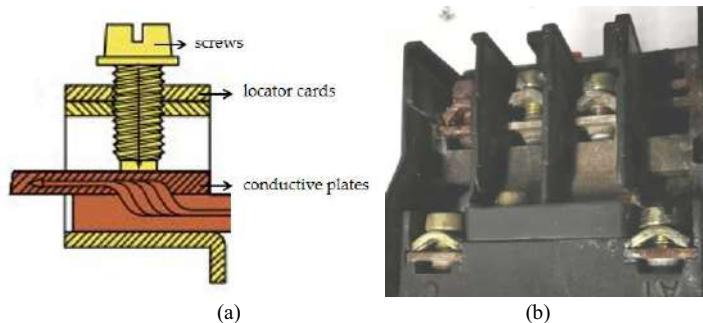


Figure 4. (a) structure of relay terminals (b) corrosion of relay terminals.

- Conductive plates

Conductive plates were supposed to maintain a permanent connection of airtightness and low resistance with wires. They are supposed to have excellent features of conductivity and resistance to erosion.

Copper and copper alloy were preferred connections in relay terminals, as such conductive plates. Copper has excellence in mechanical features, resistance to erosion and conductivity. In addition, same thermal expansion coefficient is shared between copper-made metal components and copper wires. As a result, loose connection caused by the discrepancy of thermal expansion under various temperatures could be avoided.

- Screws and locator cards

Screws and locator cards could firm the wires and were supposed to have excellent features of erosion resistance and adequate physical strength.

They were usually made of copper alloy or steel coated with stain-proof materials. Key material indicators were resistance to erosion and mechanical intensity. In principle, steel without stain-proof coating should not be used so that electrical erosion, unreliable electrical connection and rusted screws, caused by the battery effects between steel components and copper wires under wet conditions, could be avoided.

Relay OEMs usually electroplate the terminals in order to improve the wear resistance, conductivity, reflection and resistance to erosion. Coatings by major relay OEMs were mostly zinc, nickel and tin as Table 3 shows, owing to the trade-off between cost and effectiveness.

Table 3. Terminal materials of relays by mainstream OEMs.

Metal components		Tian Shui	Siemens 7PA	Weidmueller	Finder
Screws	Metal substrates	Steel	Steel	Steel	Stainless steel
	Coating	Zinc	Zinc	Zinc	Zinc
Locator cards	Metal substrates	Steel	Steel	Steel	Stainless steel or copper alloy
	Coating	Zinc	Zinc	Zinc	Zinc
Conductive plates	Metal substrates	Brass(Copper alloy)	Brass (Copper alloy)	Red copper (Copper alloy)	Red copper (Copper alloy)
	Coating	Tian Shui	Tin	Sn-Pb alloy	Nickel

Major features of the coating are as follows.

1)Zinc coating. Zinc coating is the anode for steel, which prevents the erosion of steel. Its effectiveness in protection and ornament is significantly enhanced after passivation and dyeing. Zinc coating is mostly used in screws and guide rails.

2)Nickel coating. Nickel is stable in the air, as the passivation will create a thin layer on the surface which protects the nickel from the erosion of air. Nickel coating is mostly used in screws, cube components and conductors.

3)Tin coating. Tin has good conductivity, welding and extension. It proves to be effective in erosion resistance but it's soft to be easily damaged. Tin coating is mostly used in pins.

According to the analysis above, suggestions on terminals material selection of relays are as follows.

1)Stainless steel and copper alloy were recommended to be used as materials for screws and locate cards, owing to their resistance to erosion.

2)Copper and its alloy were recommended as materials for conductive plates, owing to their conductivity.

3)Nickel coating and zinc coating with thick layers were recommended in relays, owing to their resistance to erosion and cost.

In additions, measures to improve the terminal boxes operating conditions were necessary to mitigate the erosion of relays' metal components.

4. Relay Structure

4.1. Encapsulation Structure

Encapsulation structures is one of the important factors affecting the reliability of relays. The common packaging methods of the relay are open, dustproof, plastic sealing and sealing.

In terms of protection characteristics, sealed relays (including sealed and plastic sealed relays) can isolate contacts and coils from the surrounding medium. This kind of relays can isolate the external environmental climate stress, but the structure and the analysis of the failure are complex. The non-sealed relays (including dust-proof and open relays) is highly sensitive to the environmental changes, and the long-term environmental pollution may lead to the decrease of the reliability of the relays. However, the structure is simple so the installation and maintenance are convenient.

In terms of electrical performance, the electrical durability of non-sealed relays is longer than sealed relays for loads above 2A. In practical applications, non-sealed relays are preferred if the environment permits.

The working current of the relays in the outdoor terminal box is usually between 2A to 10A. So non-sealed relays are preferred considering the service life and the convenience of regular inspection and maintenance as the electrical wiring is very complex in the terminal box. According to the environmental characteristics in the outdoor terminal box, it is recommended to choose the dust proof cover relay or the open relay with the overall shell or environmental protection requirements.

4.2. The Structure of Auxiliary Contacts

Due to the miniaturization of electrical control equipment, the structure of auxiliary contacts should be considered for the relays with multiple sets of auxiliary contacts. Poor structure design may lead to the short circuit in auxiliary contacts. The contact structure of some typical relays is shown in Figure 5.

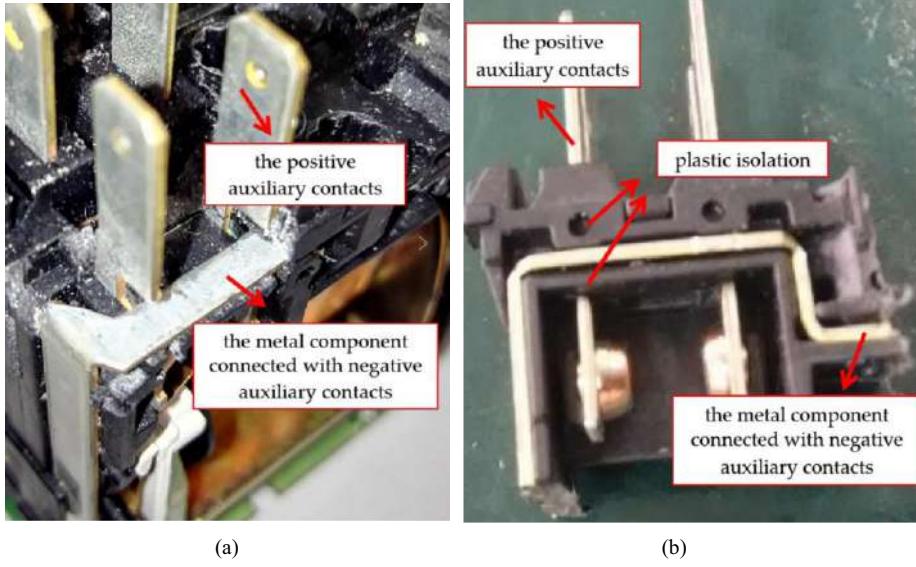


Figure 5. The structure of auxiliary contacts (a) the structure with no-plastic isolation (b) the structure with plastic isolation.

In figure (a), there is no plastic isolation between the positive auxiliary contact and the metal component connected with negative auxiliary contact. The distance between the two is around 1.8 mm with a hollow structure. In figure (b), there is plastic isolation between the positive auxiliary contact and the metal component connected with negative auxiliary contact.

As the conductive pollution may occur in outdoor terminal box, it is recommended to select the relays with plastic isolation between the positive and negative auxiliary contacts which can avoid insulation failure by pollutants.

5. Conclusions

The paper analyzes the working environment of the relays in the outdoor terminal boxes of the converter stations. The influences of environmental conditions such as fluctuating high and low temperature, humidity, dust, should be considered in the type selection of relays.

Through investigations and tests, the paper analyzes the typical materials and structures of relays and proposes several type selection tips, including phosphorus flame retardants should be avoided, stainless steel and copper alloy with coating were recommended to be used in terminals of relays, no-sealed relay with dust proof

cover is preferred and it is recommended to select the relays with plastic isolation between the positive and negative auxiliary contacts.

References

- [1] Zheng, X.X., Li, J.A., Song, S.B. Improvement and application of moisture-proof measures for outdoor terminal box in converter station. *Electric Engineering*, 2018; 15: 125-128.
- [2] Zhou, Z.M., Li, Y., Analyzing selection and application of a relay. *Electric Switchgear*, 2004; (3).
- [3] Li, D.X., Discussion on selection of electromagnetic relay and matters needing attention in use. *Electromechanical Components*, 2015; 02.
- [4] Luo, L., Analysis of selection and reliable operation of electromagnetic relay. *Automation and Instrumentation*, 2014: 02.
- [5] Lv, B., Wei, Y.M., Xiao, W.Z., Preliminary analysis of selection and reliable use of electromagnetic relay, *Electronics World*, 2021; 05:176-177.
- [6] Feng, L., Wang, Z. Study on Outdoor Terminal box moisture-proof measures. *Journal of Physics: Conference Series*, 2019; Volume 1325.
- [7] Xu, H., Yin, H.B., Liu, Q.Y., Solutions to the problem of condensation of substations in the south. *Electrical Engineering*, 2018; 19: 108-116
- [8] Wang, Z., Yang, Y.F., Liu K., Relay insulation failure mechanism of Xing'an HVDC project and regulation measures. *Guangdong Electrical Power*, 2019; 32(08):126-132.
- [9] Yan, H., Yang, J.F., Progress of phosphorus flame retardants' application in plastics. *Plastics Additives* 2015; (6):6-8.

Research on Distribution Network Reliability Investment Estimation Method Based on Sequence Linearization Correlation Analysis

Huazhen CAO^a, Chong GAO^a, Yaxiong WU^a, Hao LI^a, Zijun WANG^b¹, Chengmin WANG^b

^a*Grid Planning & Research Center, Guangdong Power Grid Co. Guangzhou 510030, China*

^b*Electronic Information and Electrical Engineering Department, Shanghai Jiaotong University, Shanghai 200240, China*

Abstract. This paper presents an estimation method of distribution network reliability planning Investment Based on sequence linearization correlation analysis. Firstly, the planning business index closely related to reliability are selected, and the control objectives of reliability index are decomposed into the promotion objectives of each planning business index through sequence linearization correlation analysis. Then, the typical engineering construction scenarios corresponding to each planning business index are constructed, and the investment required to achieve the corresponding promotion objectives of business index is estimated according to the typical scenarios, Finally, the total investment of reliability planning is obtained. The example shows that the method can be applied to the actual distribution network with complex grid conditions and various planning schemes, and can provide powerful guidance for power supply enterprises to improve the efficiency of capital use

Keywords. reliability, distribution network planning, correlation analysis, investment estimation

1. Introduction

With the continuous development of national economy, both industrial activities and people's living conditions require reliable power supply. Power supply reliability refers to the ability of power system to provide power and electricity to power users continuously according to acceptable quality standards and required quantity [1]. With the increasing attention to power supply reliability, researchers begin to study reliability planning. In the planning stage, methods to improve power supply reliability mainly include increasing liaison, increasing segmentation, optimizing grid structure, transformation of distribution automation, etc., which require huge capital investment.

¹ Corresponding author: Zijun Wang, Electronic Information and Electrical Engineering Department, Shanghai Jiaotong University, Shanghai 200240, China; E-mail: 505386064@qq.com.

In order to control the cost, power supply companies must achieve a balance between reliability and economy. Reliability planning investment estimation has become an important part of decision-making in the background, which directly affects the capital plan and profit expectation of power enterprises.

In order to improve the effect of reliability planning, [2] considers the cluster island operation and studies the method of improving the reliability after DG access through two-step planning. Reference [3] applies the fuzzy set pair analysis method to multi-objective power grid planning, and [4] carries out active distribution network expansion planning based on uncertain network theory. Reference [5] presents a method to optimize the loop-based microgrid structure in an active distribution network. Reference [6] focus on contribution of wind farms to distribution network reliability. Reference [7] uses a minimal path based algorithm to compute the reliability of stochastic distribution network to balance reliability and delivery cost.

The existing studies have different explanations for the economy in reliability planning, but they rarely involve how to estimate the reliability investment. The investment data mostly come from post statistics, so it is difficult to estimate the future investment.

In terms of distribution network reliability index and its influencing factors, [8] studies the typical wiring mode of medium voltage distribution network, and [9] focuses on the optimal allocation of energy and analyzes the impact of various factors on reliability and economy. Reference [10] converts the reliability index of power grid into economic index, that is, puts forward the concept of "economic cost". Reference [11] [12] studies the reliability influencing factors in transmission network and hybrid microgrid, which has reference significance. Reference [13] presents the method to assess the reliability and risk implications of post-fault demand response. Reference [14] takes the influence of fault location and delay protection system into consideration. Reference [15] analyses urban distribution network through the application of a time-sequential Monte Carlo simulation.

It can be found that the existing research results on distribution network reliability planning are not precise enough in the economic cost calculation of planning transformation, and can not be targeted to adapt to a variety of planning scenarios. In addition, the existing methods for obtaining the investment data of reliability planning largely rely on manual experience and post estimation, or simply make prediction based on the data of previous years. There lacks a theoretical model that can systematically and scientifically estimate the investment required in the planning scheme to improve the power supply reliability of medium voltage distribution network under specific scenarios.

In order to solve the problems above, this paper is committed to studying and calculating the planning investment estimation method required to achieve the given reliability index. By constructing a typical engineering construction scenario, this paper analyzes the engineering construction cost required for the improvement of each planning business index related to reliability on a micro scale, so as to provide more detailed and Reasonable reliability planning investment estimation method.

2. Reliability Investment of Distribution Network

The estimation problem of reliability planning investment is to estimate the planning investment required to achieve the predetermined reliability target. At present, the common practice is to build a data model for prediction according to the statistical value

of reliability in recent years and the reliability planning investment in the corresponding year. Although this method is simple to implement, it has the following disadvantages: firstly, the lack of historical data samples will affect the estimation accuracy; Secondly, reliability planning investment does not represent the actual investment, and it often takes time to see the effect of reliability improvement after the transformation, which leads to the weak correlation between the statistical value of reliability index and the reliability planning investment in the same year, thus affecting the estimation accuracy; Finally, some changes of management such as the change of reliability index statistical caliber and the adjustment of power supply jurisdiction, will also affect the estimation accuracy.

In order to deal with the above problems and give a more reasonable reliability planning investment estimation method, this paper presents an estimation method of distribution network reliability planning which is based on sequence linearization correlation analysis.

The business index managed by power companies refer to the key index of daily control of each business department. For example, the business index controlled by planning department include transferable power supply rate, cable utilization rate, feeder automation coverage rate, etc. Many business index are closely related to reliability, and business index are usually easier to count and manage. Therefore, power companies prefer to transform reliability index into business index, which is more intuitive and controllable. This paper is also based on business index. Firstly, the planning business index in correlation with reliability are selected, and then the reliability index is decomposed into the promotion targets of relevant planning business index. Then, by constructing a typical scenario for the promotion of each business index, the investment required for the promotion of each business index is estimated, and finally the total reliability planning investment is summarized. The overall idea is shown in Figure 1.

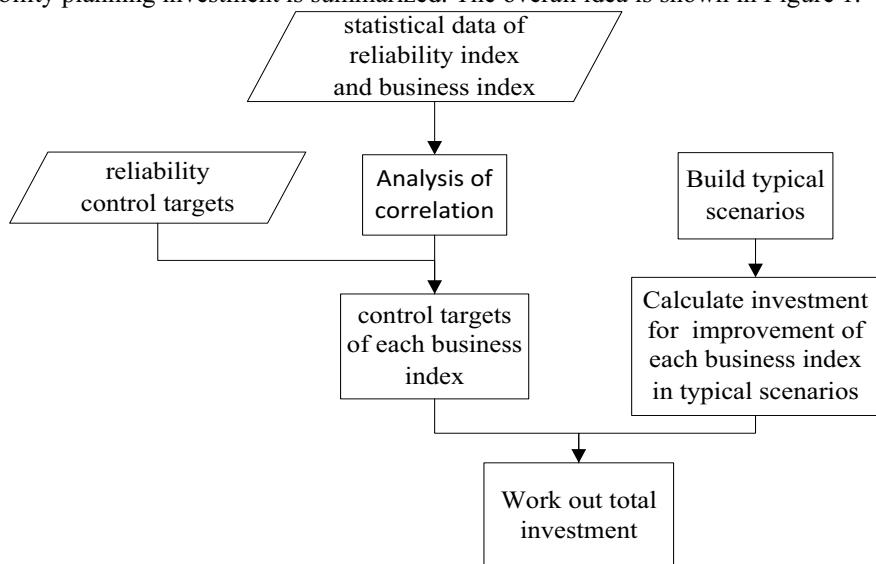


Figure 1. Investment estimation procedure for reliability planning.

The whole process is generally divided into two parts: business index decomposition and typical scenario estimation.

In the business index decomposition stage, the first step is to collect the statistical data of reliability index and business index in historical years, and construct the correlation equation according to these data to obtain the correlation degree of each business index to the reliability index, and then decompose the reliability control targets into the control targets of each business index according to the correlation degree ranking of each business index.

In the typical scenario estimation stage, firstly typical scenarios for the improvement of various business index are built, then the average implementation cost of each typical scenario is estimated according to the collected distribution network project cost, the investment required for the improvement of various business index is worked out, and finally the total investment in reliability planning is summarized.

In the whole process, the overall reliability promotion target is transformed into the promotion target of a single business index, and then the promotion of a single business index is transformed into a typical planning scheme in combination with typical scenarios, so as to realize investment estimation. Compared with the common estimation methods, the calculation of this method is more precise and reasonable, and it is also more suitable for distribution network planning guided by business index.

3. Analysis of Sequence Linearization Correlation

When studying what transformation should be carried out to improve reliability index, it is necessary to determine the correlation between the improvement of various business index and the reliability index to be improved. The improvement of reliability index is affected by various business indexes, which is complex and diverse, and can be realized through the analysis of sequence linearization correlation degree.

Physical systems including power network are nonlinear, and the relationship between the performance of a nonlinear system and a state variable can be described mathematically as follows:

$$y = f(x) \quad (1)$$

where x and y represent the independent variable and dependent variable vector of the physical system respectively. The independent variable is generally the state variable of the physical system, and the dependent variable is generally the performance index of the physical system.

Considering the complexity of nonlinear systems, they are generally solved and analyzed by linearization or continuous linearization. Assuming that y changes only in a small neighborhood that can be linearized, there are:

$$y = f(x_0) + f'(x_0)\Delta x + \varepsilon(\Delta x^2) \quad (2)$$

where $\varepsilon(\Delta x^2)$ represents terms other than primary and secondary after Taylor expansion. When these terms are ignored there are:

$$\Delta y = f'(x_0)\Delta x \quad (3)$$

Written as a matrix:

$$\Delta Y = A \Delta X \quad (4)$$

where A is the incidence matrix, indicating the correlation degree between independent variables and dependent variables.

Generally, when it is difficult to establish a complete mathematical model of the physical system, the analysis method of big data can be used to fit the a-matrix. In this paper, assuming that Y represents the reliability investment cost of distribution network and X represents many factors affecting reliability, a matrix represents the degree of influence on reliability investment cost.

Since the above process is aimed at the linearization of nonlinear functions, it is necessary to determine the reference values X_0 and Y_0 . Then, carry out big data analysis on this reference value. The independent variable and dependent variable are further corrected:

$$x = x_0 + \Delta x, y = y_0 + \Delta y \quad (5)$$

Therefore the modified values of the parameters to be solved of the incidence matrix are obtained. Then the corrected value is analyzed by big data, and the independent variable and dependent variable are corrected again. The X and Y values obtained from the solution are the change sequence from the reference value to the actual value. Repeat the correction process of the deviation until the values stop changing. This process is similar to the Newton method used in solving nonlinear equations.

Since the data of reliability parameters and various business index in the distribution network need to be updated frequently according to the monitoring values to ensure the latest state of safe and reliable operation of the system, it is usually not difficult to obtain a large number of statistical data. Through the correlation analysis of big data, the improvement of reliability index can be effectively linearized and decomposed into the improvement objectives of various business index.

In the stage of distribution network planning, the planning business index with great correlation with reliability are mainly four index: transferable power supply rate, cable utilization rate, average number of sections of overhead lines and feeder automation coverage. For the convenience of discussion, this paper only takes the most commonly used average outage time of users as the reliability index that should be paid attention to in the distribution network planning stage, and carries out the sequence linearization correlation analysis on this basis.

After collecting the statistical values of reliability index and business index in the last 3 ~ 5 years, build a continuous linearized correlation equation as shown in equation (6).

$$R_t = \sum_{j=1}^n K_j X_t^j + K_0 \quad (6)$$

where R_t is the reliability index of year t ; X_t^j is the j -th business index in year t ; N is the number of business index; K_j is the linearization coefficient to be determined for the j -th business index; K_0 is the coefficient to be determined for the constant term.

By plugging the data of X_t^j and R_t of each historical year (M years in total) into equation (6), the incidence matrix shown in equation (7) can be obtained.

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix} = \begin{bmatrix} 1 & X_1^1 & X_1^2 & \dots & X_1^n \\ 1 & X_2^1 & X_2^2 & \dots & X_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_m^1 & X_m^2 & \dots & X_m^n \end{bmatrix} \begin{bmatrix} K_0 \\ K_1 \\ K_2 \\ \vdots \\ K_n \end{bmatrix} \quad (7)$$

The sequence linearization correlation coefficient of each business index to the reliability index can be obtained by using the least square method. The higher the coefficient is, the more significant is the effect of the service index on reliability improvement. In other words, to improve reliability index, priority should be given to improving business index with high relevance.

The process of decomposing the reliability index control value (or planning value) into each business index control value (or planning value) according to the correlation degree of each business index to the reliability index can be expressed by the following equations.

$$\begin{cases} R_{m+1} - R_m = \sum_{j=1}^n K_j (X_{m+1}^j - X_m^j) \\ X_m^j \leq X_{m+1}^j \leq X_{\max}^j \end{cases} \quad (8)$$

where: R_m is the statistical value of reliability index in the m -th year; R_{m+1} is the control value of reliability index in the $m+1$ year; X_m^j is the statistical value of the j -th business index in the m -th year; X_{m+1}^j is the statistical value of the j -th business index in year $m+1$; X_{\max}^j is the upper limit of the j -th business index.

According to the basic principle of distribution network and practical construction experience, the improvement of four index: transferable power supply rate, cable utilization rate, average number of sections of overhead lines and feeder automation coverage will enhance the reliability of power grid, resulting in the change of the control value of reliability index. For example, the value CAIDI (Customer Average Interruption Duration Index) discussed in this paper will decrease with the improvement of business index. CAIDI can be calculated as follows:

$$CAIDI = \frac{\sum_{i=1}^N \lambda_i}{N} \quad (9)$$

where: λ_i is the total interruption time of the i -th user; N is the number of users.

In fact, the specific contribution of each index to reliability improvement is difficult to formulate through theoretical analysis. However, through the sequence linearization correlation analysis, the fine and reasonable linear correlation degree of business index based on big data can be obtained without a complete model of the reliability index affected by various factors. The fine and reasonable linear correlation degree of business index based on big data can be obtained, so as to transform the complex investment estimation problem into a linear optimization problem and guide the reliability planning of distribution network. The correlation coefficient between the reliability index and each business index obtained by solving the incidence matrix by the least square method reflects the impact of the change of each index on the change of reliability index.

4. Investment Estimation Based on Typical Scenarios

After decomposing the reliability index control value into control value of various business index, the cost required for the improvement of business index is estimated by constructing a typical scenario for the improvement of various business index, so as to obtain the planned investment required to achieve the reliability control objectives.

4.1. Definition of Typical Scenarios

While doing distribution network planning, different kinds of transformation are required to improve different business index. In this section, five typical scenarios that often occur in the actual construction and transformation of distribution network are defined. The construction and transformation forms of these typical scenarios are relatively standard and unified. After collecting information about engineering cost, more accurate investment estimates can be obtained. Specific scenarios are defined as:

Scenario 1: Adding a tie point to the existing overhead line. That is to add tie line and tie switch on the existing overhead single radiation line in order to transform it into hand-in-hand wiring mode. In this scenario the switchable power supply rate of distribution network is increased.

Scenario 2: Building a new line and adding a tie point. That is to build a new no-load cable line as a public standby line, which is usually used to transform multiple single radiation lines into "n-1" main and standby wiring.

Scenario 3: Creating a new cable line. That is to build a new cable line in the area with low cable line coverage. In this scenario the cable utilization rate of distribution network is increased. This paper did not take the replacement of existing overhead lines with cable lines into consideration, mainly because except for the needs of municipal engineering, it is generally not allowed to directly change the existing overhead lines into cable lines in present. Optimization can only be performed on the basis of the existing grid.

Scenario 4: Adding a section switch to the existing overhead line. That is to add a section switch to the existing overhead line. In this scenario the average number of sections of the overhead line of the distribution network is increased.

Scenario 5: Distribution automation transformation on existing lines. That is to erect intelligent switchgear and communication network for existing lines, namely distribution automation transformation. In this scenario the feeder automation coverage of distribution network is increased.

In the above five typical scenarios, scenario 1 and scenario 2 correspond to the improvement of switchable power supply rate, scenario 3 corresponds to the improvement of cable utilization rate, scenario 4 corresponds to the improvement of the average number of sections of overhead lines, and scenario 5 corresponds to the improvement of feeder automation coverage. Moreover, scenarios 1 to 4 are separately applied to the overhead network or cable network. Only scenario 5 covers the overhead network and cable network, which need to be discussed separately when estimating the cost.

4.2. Investment Cost Calculation in Typical Scenarios

The cost for the proposed five typical scenarios is calculated as follows.

Scenario 1: Adding a tie point to the existing overhead line

The planning scheme of adding a tie point to the existing overhead line can increase the transferable power supply rate of the distribution network. This scheme is applied to the project of transforming the single radiation network into a "2-1" single ring network.

The formula for investment cost in this scenario is

$$C_1 = L_l (c_o + rc_{or}) + c_{ob} \quad (10)$$

where C_1 is the cost of scenario 1; L_l is the length of the tie line, which usually takes 1/5~1/3 of the power supply radius of the line; c_o is the unit cost of overhead line; R is the construction coefficient, which is between 0 and 1; c_{or} is the unit construction cost of overhead line erection (such as stringing and pole erection); c_{ob} is the unit cost of pole-mounted switch.

Scenario 2: Building a new line and adding a tie point

The planning scheme of building a new line and adding a tie point can also increase the switchable power supply rate of the distribution network, but this scheme is applied to the project of transforming the single shot into the connection form of N supply and one standby. The formula for investment cost is as follows:

$$C_2 = L_l (c_c + rc_{cr}) + c_{cb} \quad (11)$$

where C_2 is the cost of scenario 2; L_l is the length of the tie line; c_c is the unit cost of cable; R is the construction coefficient; c_{cr} is the unit construction cost of cable (such as trenching and landfilling); c_{cb} is the unit cost of switchgear.

Scenario 3: Creating a new cable line

The construction cost of the new cable line includes the cable conductor cost, switchgear cost, box transformer cost and construction cost estimated according to the average line length parameter. The specific calculation formula is

$$C_3 = L_c (c_c + rc_{cr}) + Uc_t \quad (12)$$

where C_3 is the cost of scenario 3; L_c is the length of new cable line; c_c is the unit cost of cable; R is the construction coefficient; c_{cr} is the unit construction cost of cable (such as trenching and landfilling); U is the average number of costumer, usually ranging from 15 to 35; c_t is the unit cost of distribution transformer.

Scenario 4: Adding a section switch to the existing overhead line

Adding a section switch to existing overhead lines can increase the average number of line sections, and the cost is related to the construction of section switches. The calculation formula is

$$C_4 = c_{ob} \quad (13)$$

where C_4 is the cost of scenario 4; c_{ob} is the unit cost of pole-mounted switch.

Scenario 5: Distribution automation transformation on existing lines

Distribution automation transformation for existing lines can improve feeder automation coverage. It can be subdivided into two typical scenarios: overhead line pole-mounted switch transformation and cable line switch station transformation.

The investment cost for feeder automation transformation of an overhead line is

$$C_{5o} = \frac{1}{2} L_o r c_{or} + 2c_{ob} \quad (14)$$

where C_{5o} is the overhead line feeder automation transformation cost in scenario 5; L_o is the total length of overhead line; R is the construction coefficient; c_{or} is the unit construction cost of overhead line feeder automation transformation (such as optical fiber erection); c_{ob} is the unit cost of pole-mounted switch.

The investment cost for feeder automation transformation of a cable line is

$$C_{5c} = \frac{1}{2} L_c r c_{cr} + 2c_{cb} \quad (15)$$

where C_{5c} is the cost of cable line feeder automation transformation in scenario 5; L_c is the total length of the cable line; R is the construction coefficient; c_{cr} is the unit construction cost for automation transformation of cable line feeder (such as optical fiber erection); c_{cb} is the unit cost of switchgear.

4.3. Investment Cost Estimation

After the cost of each typical scenario is calculated, the total investment is estimated according to the current value of each business index and the goal of reliability improvement. The key point of this process is to determine the engineering quantity of each typical scenario of overhead network and cable network.

Combining scenario 1 and scenario 2 we can get the investment estimation of increasing the renewable power supply rate as follows:

$$C_l = C_1 \Delta N_{ol} + C_2 \Delta N_{cl} \quad (16)$$

where C_l is the investment cost for increasing the renewable power supply rate; C_1 and C_2 are the cost of scenario 1 and scenario 2 respectively; ΔN_{ol} and ΔN_{cl} are the number of overhead lines and cable lines that need to be connected.

Based on scenario 3, the investment estimate for increasing the cable utilization rate is

$$C_c = k_c C_3 \Delta N_c \quad (17)$$

where C_c is the investment cost for increasing cable utilization rate; k_c is the conversion factor, since the addition of a cable is not only for improving the reliability of power supply but also for meeting the demand of load growth; C_3 is the cost of scenario 3; ΔN_c is the number of new cable lines.

Based on scenario 4, the investment estimate for increasing the average number of sections of overhead line is

$$C_s = C_4 N_o (X_s^1 - X_s^0) \quad (18)$$

where: C_s is the investment cost of increasing the average number of sections of overhead line; C_4 is the cost of scenario 4; N_o is the number of overhead lines; X_s^1 is the target value of the average number of sections of overhead line; X_s^0 is the current value of the average number of sections of overhead line.

Based on scenario 5, the investment estimate for improving feeder automation coverage is

$$C_a = C_{5o} \Delta N_{oa} + C_{5c} \Delta N_{ca} \quad (19)$$

where: C_a is the investment cost of improving feeder automation coverage; C_{5o} is the cost of implementing feeder automation for overhead lines in scenario 5; ΔN_{oa} is the number of overhead lines requiring feeder automation; C_{5c} is the cost of implementing feeder automation for the cable line in scenario 5; ΔN_{ca} is the number of cable lines that need to implement feeder automation.

To sum up, $C_l + C_c + C_s + C_a$ is the total planned investment required to achieve the reliability control goal.

5. Case Analysis

This section takes the distribution network of a city as an example. The statistical values of the average outage time of customers, various business index and the corresponding reliability planning investment in the last five years are shown in Table 1.

Table 1. Historical data of each index and investment in a city.

Year of data	1	2	3	4	5
transferable power supply rate (%)	68.35	70.83	73.18	76.32	79.22
cable utilization rate (%)	40.23	41.13	41.78	42.23	42.84
average number of overhead line sections	2.65	2.74	2.81	2.88	2.94
feeder automation coverage (%)	8.34	15.23	31.78	50.44	68.23
CAIDI (h/household)	3.13	2.88	2.43	1.91	1.41
Investmen (million yuan)	0.83	0.94	1.15	1.38	1.59

According to the four business index and the average outage time of costumer in each historical year in Table 1, the correlation matrix in the form of equation (7) is listed, and the correlation degree value shown in Table 2 is obtained through the sequence linearization correlation analysis.

Table 2. Correlation of business index and CAIDI in a city.

transferable power supply rate	cable utilization rate	average number of overhead line sections	feeder automation coverage
1.83	0.66	11.09	1.31

It is assumed that the planning goal is to reduce the average outage time of users from 1.41h/household to 1.0h/household. According to the correlation degree between each service index in Table 2 and the user average outage time index, the planning objectives of each service index can be obtained by solving equation group (8), as shown in Table 3. It can be seen from table 3 that the reliability control target can be met only by improving the two planned business index of transferable power supply rate and the average number of sections of overhead lines.

Table 3. Planning objectives of business index in a city.

transferable power supply rate (%)	cable rate (%)	utilization	average number of overhead line sections	feeder automation coverage (%)
Current value	79.22	42.84	2.94	68.23
Target value	86.72	42.84	3.94	68.23

By substituting the above planning improvement objectives into the actual project scenario, the number of overhead lines and cable lines that need to be connected can be obtained. See Table 4 for the number of lines and other calculation parameters required for investment estimation.

Table 4. Setting of other parameters (unit in million yuan per kilometre).

L_l	c_o	c_{or}	c_{ob}
2	56.65	9.86	7.05

R	c_c	c_{cr}	c_{cb}
0.7	65.58	100.45	43.1
L_c	U	c_t	L_o
7.83	38.56	20.06	9.76
N_o	N_c	ΔN_{ol}	ΔN_{cl}
622	467	81	0

According to equations (16) and (18), the investment amount required to improve the transferable power supply rate and the average number of sections of overhead lines in this calculation example can be calculated respectively. The planned investment cost required to achieve the reliability control goal is shown in Table 5.

Table 5. Estimation results of reliability planning investment (unit in 100 million yuan).

C_l	C_c	C_s	C_a	Total
1.09	0	0.44	0	1.53

If the average outage time of costumer is reduced to 1.0h per household according to the traditional regression prediction method, the investment required is about 189 million yuan. It can be seen that there is a certain deviation in the investment obtained by the two estimation methods, which is mainly due to the different principles of the two estimation methods. The estimation results of the method in this paper are greatly affected by the correlation analysis results of the four business index and the unit cost of equipment, while the estimation results of traditional methods are greatly affected by historical data. In this paper, the investment is allocated according to the idea of prioritizing the relevance of planning business index, so the investment amount obtained is more in line with the actual engineering scenario.

6. Conclusion

In this paper, a method for reliability planning investment estimation based on sequence linearization correlation analysis is proposed. According to the analysis results of the sequence linearization correlation between the distribution network planning business index and the reliability index, the planning targets of the reliability index are decomposed into the control targets of each business index, and then the typical scenarios for the improvement of each business index are constructed for investment estimation, so as to obtain a more reasonable reliability planning investment. This estimation method takes reliability and economy as a whole while considering the optimization of reliability improvement. It is a new exploration of distribution network planning and optimization. Moreover, this method can also be used to directly estimate the reliability planning investment according to the planning targets of each business index after the conventional planning of distribution network is completed.

In addition, for different districts and counties or different power supply districts under the same city, the reliability control objectives and promotion strategies will certainly be different. The correlation between each planned business index and reliability under each district and county or power supply district can be analyzed according to the core idea of this paper, and then solved according to the minimum total cost. Thus, a reliability planning model considering differentiated demand and optimal total investment is obtained.

References

- [1] Billinton R, Li W. Reliability assessment of electric power systems using monte Carlo methods [M]. New York: Plenum Press, 1994.
- [2] Hu Di, Ding Ming, Bi Rui, et al. Two-step planning of grid-resource-storage based on cluster partition [J]. Automation of Electric Power System, 2020, 44(2):95-106.
- [3] Jin Huazheng, Cheng Haozhong, Yang Xiaomei, et al. Application of fuzzy set pair analysis in multi-objective transmission network planning considering ATC[J]. Automation of Electric Power System, 2005, 29(21):49-53.
- [4] Li Yan, Hu Zhijian, Zhang Menglin, et al. Active distribution network planning considering economy and reliability based on uncertain network theory[J]. Automation of Electric Power System, 2019, 43(16):68-82.
- [5] Cortes, C., Contreras, S., Shahidehpour, M., et al. Microgrid topology planning for enhancing the reliability of active distribution networks [J]. IEEE Transactions on Smart Grid, 2018, 9(6), 6369-6377.
- [6] Blake, S., Taylor, P., Dent, C., Miller, D., et al. Quantifying the contribution of wind farms to distribution network reliability[J]. Wind Energy, 2014, 18(5), 941-953.
- [7] Xu, X., Niu, Y., Song, Y., et al. Computing the reliability of a stochastic distribution network subject to budget constraint[J]. Reliability Engineering & System Safety, 2021, 216, 107947.
- [8] Liu Xiangjun, MA Shuang, XU Gang. Formation of typical connection mode for distribution network by elementary connection model[J]. Power System Technology, 2012, 36(2):58-63.
- [9] Liu Ziqiu, Hang Minxiang. Optimal energy allocation of microgrid based on reliability and economy by considering characteristics of micro source[J]. Power System Technology, 2014, 38(5):1352-1357.
- [10] Zhu Xukai, Liu Wenyin, Yang Yihan. A new power network planning method by comprehensively considering system reliability factors[J]. Power System Technology, 2004, 28(21):51-54.
- [11] Wang Yizhe, Tang Yong, Dong Chaoyang, et al. A hybrid criterion model for transmission system medium and long terms expansion planning[J]. Power System Technology, 2016, 40(7):2094-2098.
- [12] Wei Yongzhong, Gu Ren, Yang Chen, et al. Optimization model for distribution network reliability investment considering interruptible load[J]. Smart Power, 2019, 47(2):107-112.
- [13] Syrri, A., Mancarella, P. Reliability and risk assessment of post-contingency demand response in smart distribution networks[J]. Sustainable Energy, Grids and Networks, 2016, 7, 1-12.
- [14] Xiao, F., Xia, Y., Zhou, Y., et al. Comprehensive reliability assessment of smart distribution networks considering centralized distribution protection system[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2019, 15(1), 40-50.
- [15] Hernando-Gil, I., Ilie, I., Djokic, S., et al. Reliability planning of active distribution systems incorporating regulator requirements and network-reliability equivalents[J]. IET Generation, Transmission & Distribution, 2016, 10(1), 93-106.

Research on the Construction of Online Culture Trading Platform Based on Netizens' Wishes

Ren SONG, Weiwei WANG¹, Yang LIU

Jilin University of Finance and Economics, Jilin, China

Abstract. According to the statistical data of the 44th China statistical report on Internet development, as of June 2019, the number of Internet users in China has reached 854 million, and the penetration rate has reached 61.2%. The number of mobile Internet users reached 8.47 million, accounting for 99.1% of the total Internet users [1]. The number of Internet users in China continued to grow steadily. With the rapid development of high-tech and Internet, the continuous expansion of Internet users, the continuous optimization of the network environment, the gradual formation of the network society, the network life tends to be a regular state, and the network culture is also formed in this environment, which has become an important part of China's cultural system. This paper takes netizens as the core, studies the construction of network cultural data trading platform based on netizens' behavior. Through the establishment of netizens influence model and online cultural data trading platform, this paper studies the consumption intention of online cultural products and the characteristics of online cultural data trading platform from the perspective of netizens. Based on the data of strong research results, this paper puts forward suggestions and strategies for the construction of network cultural data trading platform and the innovation of network cultural products in China.

Keywords. Network culture, network culture data, trading platform, network communication, network culture development

1. Introduction

Network culture is a kind of cultural form. It takes the network as the medium, the network platform as the carrier, and the culture as the core. Relying on the production and development of high-tech and network technology, it uses certain forms of language, sound and video to spread ideas and culture, express ideas, views, emotions and consciousness. It constructs a new way of Ideological and cultural expression, which influences and changes people's life style, production mode and values [2].

With the rise of network culture, its related production, circulation, service and other links and departments constitute a complex network culture industry [3]. In recent years, its industrial scale has been expanding. In the whole network culture industry chain, the network culture takes the network culture data transaction platform as the carrier, takes the network culture product as the ultimate goal performance. There are a variety of

¹ Corresponding Author, Information Management Center, Jilin University of Finance and Economics, Jilin, China, 23392147@qq.com.

network cultural products, including online literature, online games, online video, online music, etc.

With the increasing abundance of network cultural products, the supply capacity has been significantly enhanced. However, compared with traditional cultural products, consumers' willingness to consume online cultural products is significantly different. Ren Xianliang, deputy director of the office of the central leading group on network security and information and deputy director of the national Internet Information Office, pointed out: to develop network culture, we should strengthen the supply and dissemination of excellent network cultural products, build a platform for network cultural exchange and communication, and share the development of network culture [4], not only from the macro level of politics, system, economy and supervision, but also from the macro level from the perspective of fundamental development. As the carrier of network culture products, network culture data trading platform is the main way of network culture spreading in cyberspace. Its platform features directly affect consumers' online cultural consumption intention products. Therefore, strengthening the construction of network culture data trading platform is of great significance to improve the network culture industry and promote the prosperity and development of network cultural products.

2. Literature review

On the construction of network culture, many scholars put forward different views. Some scholars think that the root cause of the problems in the development of network culture is the lack of government support. The government should take effective measures to promote the development of network culture industry. Some scholars think that the negative effects of technological development will have a certain impact on network culture. Some scholars think that the development and management of network culture industry should pay attention to the strategic level. Some scholars advocated to encourage cultural innovation to realize the benign evolution of the industry. Some scholars think that the role of media in the network culture industry ecosystem is becoming more and more important. In the digital era, it is very important to play the main role of media. Network culture data trading platform is the carrier of network culture display and dissemination, the platform of government support and technical support, and the platform of innovation, strategic development and product operation of the whole network culture industry. Therefore, it is of great practical significance to study the network culture data trading platform and its application. Some scholars think that Internet users, as consumers, have become the core subject of wireless potential release of network culture industry, participate in the innovation process of network cultural products, and improve the innovation ability of network cultural products. Therefore, from the perspective of consumers, the construction of network culture is of great significance.

At present, the research on network culture in China is mostly carried out from three aspects: system supervision, risk system and industrial scale. However, from the perspective of consumer intention, there is little literature on the Internet. This paper focuses on netizens, from the perspective of netizens' wishes, studies the network cultural data trading platform and network cultural products, which is innovative.

3. Research method

3.1. Construction of network culture data trading platform including netizens influence model

The network culture data trading platform is an important part of the network culture industry chain. It is the carrier of network cultural products, and also the platform of spreading and trading network cultural products. If it is an excellent trading platform, it can promote the development and exchange of network culture, and also can promote the survival of the fittest of network cultural products. Like other shopping platforms, online cultural data platform is a trading platform with all the characteristics of the trading platform, which will have a certain impact on the use intention of the Internet. According to the theory of SOR (stimulus organization response), the receiver of external factors will produce conscious or unconscious psychological reactions (including emotional response or cognitive response). After experiencing a series of psychological reactions, the subjects have a certain degree of behavioral response to stimulation [5]?

Through the comparative analysis and in-depth study of many network cultural data trading platforms, this paper divides the characteristics of network data trading platforms into knowledge, economy, interaction, entertainment, vision, personalization and security. Through the stimulation of the characteristics (stimulus factors) of the network data trading platform, it can trigger the internal emotional response of Internet users, including the sense of pleasure and the perception of awakening, and then generate the willingness of Internet users. The characteristic variables of the network culture data trading platform are described in Table 1.

Table 1: main dimensions of characteristic variables of online cultural data trading platform

Variable name	Variable definition
knowledge	The richness and clarity of product related subject information and non-product related peripheral information provided by the platform
Economics	The platform provides consumers with real economic perceived benefits such as product price, price discount and promotion
Interaction	The platform uses the Internet and modern technology to interact with participants
Entertainment	As a kind of media, the platform provides animation, video and other entertainment functions for netizens
Visually	How attractive is the external form of the platform interface
individualization	The platform provides efficient, valuable and targeted services for consumers according to their preferences and behaviors
security	The security degree of the platform for netizens' information, privacy protection and payment process
inspire	The degree to which an individual feels stimulation, excitement, excitement and surprise in a particular situation
Pleasant	The degree to which an individual is satisfied, happy, happy and comfortable in a given situation
Willing to use platform	The subjective will of netizens to use Internet cultural data trading platform

The construction of network data trading platform not only includes the construction of its own platform, but also includes the innovative products of network culture and the consumers' willingness to treat the network cultural products. It is a quantitative index product for us to study the network culture. Online cultural products are intangible products. Consumers have no standard judgment of them. They often rely on their own perception to decide whether to spend or not. Therefore, this study uses perceived value to judge consumers' online cultural consumption intention products. Based on the SOR theoretical model, consumers' perceived response to online

consumption, that is, after generating perceived value, will produce consumption intention. The characteristics of network cultural products are explained in Table 2.

Table 2: the first dimension of the characteristic variables of network cultural products

Variable name	Variable definition
Ease of use	The extent to which consumers think that online cultural products are easy to use
Perceived usefulness	Consumers think that the functional benefits brought by network cultural products
Perceptual Entertainment	The degree of emotional pleasure consumers get from using Internet cultural products
Perceived cost	Consumers' willingness to pay for online cultural products is the perception of price rationality
Perceived risk	Consumers' subjective prediction of various losses and uncertainties caused by the purchase and use of online cultural products
Perceived value	On the basis of perceived profit and loss, consumers' overall evaluation of the benefits brought by network cultural products
Consumption intention	Consumers' subjective willingness to purchase and use online cultural products

Network culture data trading platform is the carrier of network culture, and network culture is finally displayed through network culture products [6]. As the terminal of network culture industry chain, network cultural products directly face Internet users. At the same time, Internet users' willingness to consume online cultural products will also be affected by the use of online cultural data trading platform. Based on the above theory, this paper constructs a relationship model between the willingness and impact of netizens using cultural data trading platform, as shown in Figure 1.

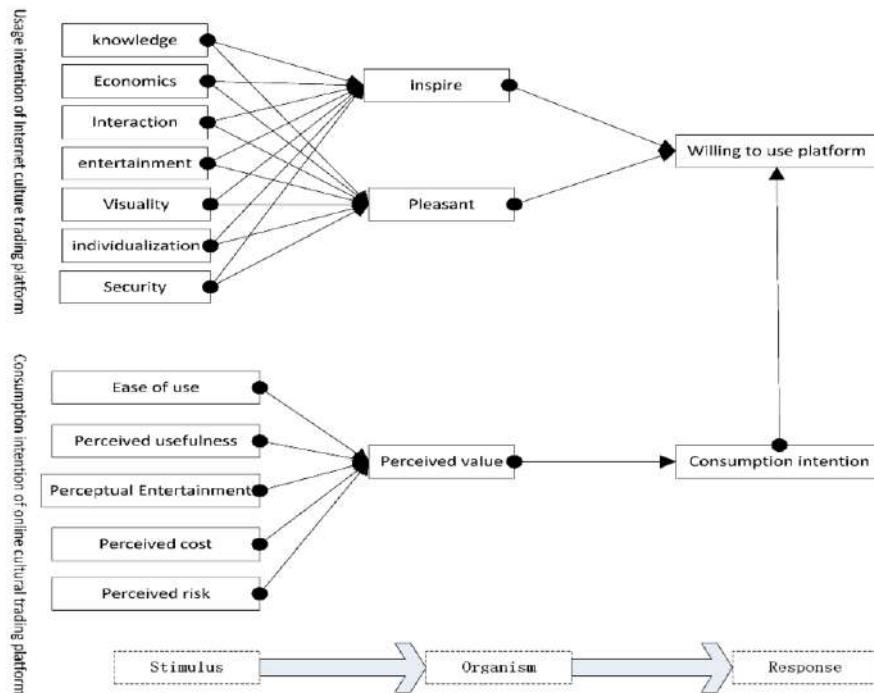


Figure 1 Influence Model of netizens' willingness on online cultural data trading platform

3.2. Construction strategy of Internet cultural data trading platform based on Netizens' wishes

According to the influence model of netizens' willingness of online cultural data trading platform, this paper studies the influence of network cultural data trading platform characteristics on Internet users' consumption intention, the influence of network cultural product characteristics on Internet users' consumption intention, and the influence of Internet users' consumption intention on Internet cultural data trading platform. According to the survey results, based on the wishes of netizens, this paper puts forward the construction and development strategy of online cultural data trading platform.

4. Academic value and application value

Based on the SOR theory, this paper constructs a model of the influence of Internet users' willingness on the Internet cultural data trading platform, which is mainly measured from three aspects, including the influence of the characteristics of the network cultural data trading platform on the Internet users' willingness, the influence of the characteristics of the network cultural products on the Internet users' willingness to consume, the influence of the network users' consumption intention and the network transaction.

According to the influence model of netizens' willingness on online cultural data trading platform, this paper systematically analyzes the definition of main variables. Through the research design and empirical analysis, the research results are obtained, and the research results are discussed and analyzed. Finally, it summarizes the construction strategy of online cultural data trading platform based on Netizens' wishes.

Based on the SOR theory, this paper constructs a model of the influence of netizens on the willingness of online cultural data trading platform, and verifies the model with the method of empirical analysis. On the basis of ensuring the theoretical basis of the model, reliable and effective data are collected through questionnaire survey, as well as the influence of the characteristics of online cultural data trading platform on Netizens' willingness to use, the influence of network cultural product characteristics on Internet users' consumption intention and the influence on the use intention of network cultural data trading platform, which provides a powerful strategy for the development of network cultural data transaction construction and data support. The research method of this paper is more objective than qualitative research.

According to the survey results of netizens' willingness on online cultural data trading platform, this paper puts forward the construction and development strategy of online cultural data trading platform based on Netizens' willingness, which provides a new perspective for the construction of online cultural data trading platform.

5. Conclusion

Based on the systematic analysis and literature review of network culture and its industry development in recent years, this paper studies the construction of network culture data trading platform based on Internet users. Just like the carrier of network cultural products and effective communication channels, network cultural data trading platform plays an important role in the development of the network. In this paper, starting from the Internet

users' willingness to consume network cultural products and the characteristics of network culture platform, starting from the construction of network culture data trading platform, this paper discusses the construction of network culture data trading platform, opens up a new perspective of network culture research, and further enriches the research content in the field of network culture.

From the perspective of netizens, this paper studies the consumption intention of online cultural products and the characteristics of online cultural data trading platform. The research results provide theoretical basis for the construction of network data trading platform and the innovation of network cultural products, and provide data support for the improvement of network culture industry system. It is helpful for network culture enterprises to improve marketing strategy, develop new network culture products, improve data trading platform of network culture, and develop Chinese network culture.

Acknowledgement

This paper is supported by Project of the Department of science and technology of Jilin Province, Research on the Countermeasures of artificial intelligence in the development of innovative education. (Contract number: 20200101084FG). The Education Department of Jilin Province, Research on the influence of intelligent robot in innovation education in Jilin Province. (Contract number: JJKH20210132KJ). The Education Department of Jilin Province, Research on the application of big data technology in the construction of smart campus. (Contract number: JJKH20200142KJ). The Education Department of Jilin Province, Path design and implementation of intelligent financial management in colleges and universities in Jilin Province under the background of artificial intelligence. (Contract number: JJKH20210134KJ). Jilin Province Higher Education Association's 2019 project, Research on the Construction of Entrepreneurship Education Ecosystem and Support System in Colleges and Universities. (No.: JGJX2019D270). Jilin Province Social Science Foundation Project, Research on the Evaluation and Optimization Strategy of the Coordinated Development of Cultural Resources and Tourism Industry in Jilin Province. (No.: 2020C053). The project of teaching reform on higher education in Jilin Province: Practice and research on ideology and politics teaching of network information security course. (No.: JLSI680620190723162538).

References

- [1] China Internet Network Information Center (CNNIC) in 《The 44th China Statistical Report on Internet Development》
- [2] Zhu Ruihai. Research on the dual effects and Countermeasures of network culture [J]. People's forum, 2013(6):174-176.
- [3] Bai Shuying. Outlook on the development of network culture with Chinese characteristics and its theoretical significance [J]. Marxist studies, 2012, 000 (003): 77-82.
- [4] Wang Xingxing, Xie zongxiao. Information Security Management Series II. Information security: the perspective of national strategy [J]. China Quality and standards guide, 2015 (3): 30-32.
- [5] Xu Yue. Research on the influence of information publicity on College Students' habitual energy-saving behavior intention [D].
- [6] Chen Lanlan. Research on the logical construction and survival path of the ecological field of network culture [D]. 2017.

Parametric Optimization of Integrated Circuit Assembly Process: An Evolutionary Computing-Based Approach

Tatjana SIBALIJA¹

Belgrade Metropolitan University, Belgrade, Serbia

Abstract. Strict demands for very tight tolerances and increasing complexity in the semiconductors' assembly impose a need for an accurate parametric design that deals with multiple conflicting requirements. This paper presents application of the advanced optimization methodology, based on evolutionary algorithms (EAs), on two studies addressing parametric optimization of the wire bonding process in the semiconductors' assembly. The methodology involves statistical pre-processing of the experimental data, followed by an accurate process modeling by artificial neural networks (ANNs). Using the neural model, the process parameters are optimized by four metaheuristics: the two most commonly used algorithms - genetic algorithm (GA) and simulated annealing (SA), and the two newly designed algorithms that have been rarely utilized in semiconductor assembly optimizations - teaching-learning based optimization (TLBO) and Jaya algorithm. The four algorithm performances in two wire bonding studies are benchmarked, considering the accuracy of the obtained solutions and the convergence rate. In addition, influence of the algorithm hyper-parameters on the algorithms effectiveness is rigorously discussed, and the directions for the algorithm selection and settings are suggested. The results from two studies clearly indicate superiority of the TLBO and Jaya algorithms over GA and SA, especially in terms of the solution accuracy and the built-in algorithm robustness. Furthermore, the proposed evolutionary computing-based optimization methodology significantly outperforms the four frequently used methods from the literature, explicitly demonstrating effectiveness and accuracy in locating global optimum for delicate optimization problems.

Keywords. Integrated circuit assembly, parametric process optimization, evolutionary algorithms (EAs), teaching-learning based optimization (TLBO) algorithm, Jaya algorithm, artificial neural networks (ANNs)

1. Introduction

An accurate design of process parameters is of utmost importance in semiconductor industry, due to extremely tight tolerances and zero-defect demands for the process outputs (i.e. device characteristics). The objective of parametric optimization is to find an optimal process parameters set that meets requirements for the response mean and reduce its variability. There is a variety of optimization approaches from literature, but not all of them are fully efficient and objective for a multi-response case. The approaches based on statistical techniques are mainly unable to locate a global optimum, since they consider only discrete (local) process parameter values utilized in experimental trials or

¹ Corresponding author; E-mail: tsibalija@gmail.com, tatjana.sibalija@metropolitan.ac.rs

data sets. The exception is response surface methodology (RSM), the most common optimization method that applies a hill climbing or descending tools over a combined response plot [1]. However, it has a tendency to be trapped into local optima for a highly non-linear process with a large number of control factors, and it does not explicitly address correlations among responses and the response variation. The Taguchi's robust parameter design simultaneously assesses both the response mean and variability, taking into account the response specification by using signal-to-noise ratio (SNR) or quality loss (QL) function; the latter shows a loss encountered by the user if the product response deviates from the target [2]. But, the traditional Taguchi method has been designed for a single response. Various approaches have been developed to integrate multiple responses in Taguchi method. Some approaches used principal component analysis (PCA) over SNR or QL data to obtain uncorrelated variables, where only components with eigenvalue higher than 1 are considered [3], thus considering only a part of the original data variability. An improved approach that considered all components was applied to optimize wire bonding process [4]. In [5] PCA was applied directly on the response data, considering all independent components, but giving misleading results since the response specifications are not examined. A number of works applied grey relational analysis (GRA) over SNR data to integrate multiple responses by assuming that all responses are of the same significance [6]; however, the response weights are allocated in a subjective manner. For optimizing wire bonding parameters vs. two responses, conflicting response specifications were addressed by the fuzzy logic [7]. Although it resulted with improved process responses, this approach considered only local solutions.

Approaches based on soft computing, including evolutionary computing techniques, present a viable alternative due to a search across a continuous multi-dimensional space of solutions. There are a few soft computing-based approaches that address parametric optimization problems from the semiconductor industry. For wire bonding process, multiple responses were integrated via GRA, and the process was mapped and optimized using combination of artificial neural networks (ANNs) and genetic algorithm (GA) [8]. The lithography process with two responses was modeled using ANNs and parameters were optimized by particle swarm optimization (PSO) [9]; however, the response specifications and correlations are not explicitly addressed. For the hi-power LED packaging, multiple responses were transformed into QL values and their sum presented the objective function; process was mapped by genetic programming and optimized by PSO [10]. But, the integration of QLs into a single objective was not performed in a totally objective manner. In overall, although effective in addressing various optimization problems, EAs have been frequently criticized due to a significant effect of their hyper-parameters on the accuracy of the obtained solutions.

This work briefly presents an advanced methodology for parametric process optimization, including four metaheuristic algorithms whose performances are benchmarked to obtain the most accurate solution. Its implementation in semiconductor industry, where accuracy is of paramount importance for the device quality and reliability, is illustrated in two use cases showing integrated circuit assembly process (wire bonding).

2. Evolutionary computing-based optimization methodology

The proposed optimization methodology involves three major stages: data pre-processing by statistical methods, process modeling and optimization. In the first stage, experimentally obtained response data are converted into quality loss (QL) values,

according to the type of response. The QL directly shows a response financial significance for the user without imposing any assumptions, as defined by Taguchi [2]:

$$QL = K \cdot \begin{cases} \frac{1}{n} \sum_{i=1}^n y_i^2 & \dots \text{for smaller - the - better type} \\ \frac{1}{n} \sum_{i=1}^n (y_i - t)^2 = \frac{n-1}{n} s^2 + (\bar{y} - t)^2 & \dots \text{for nominal - the - best type} \\ \frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} & \dots \text{for larger - the - better type} \end{cases} \quad (1)$$

Since responses from the same process are correlated, it is necessary to obtain independent variables by applying PCA on the normalized QL data (NQLs). In the proposed approach, the number of independent principal components – PCs ($(j=1, \dots, p)$) corresponds to the number of responses ($i=1, \dots, p$). For each experimental trial ($k=1, \dots, m$), the PC scores ($Y_j(k)$) are obtained based on the elements of the corresponding eigenvector (V_{ij}) and the NQL data, as follows:

$$Y_j(k) = \sum_{i=1}^p NQL_i(k) \cdot V_{ij} \quad (2)$$

GRA is applied over $Y_j(k)$ data to integrate multiple Y_j , i.e. to obtain a grey relational grade ($\gamma_k \in [0, 1]$) using a previously calculated grey relational coefficient ($\varepsilon_j(k)$) and PC weights (ω_j) obtained from PCA:

$$\gamma_k = \sum_{j=1}^p \omega_j \cdot \varepsilon_j(k) \quad (3)$$

The γ value is adopted as a process performance measure: the higher the γ , the better is the process. The details of PCA could be found in [11], and of GRA in [12].

In semiconductor industry the mathematical model of a process is typically unknown. Hence, back propagation (BP) ANNs are engaged to identify the relationship between process parameters (input) and γ (output), using the tangent sigmoid and linear transfer functions. The number of neurons in the hidden layer is varied to obtain the best topology with minimal mean square error (MSE) and maximal coefficient of correlation (R) between original and predicted data. The procedure is explained in detail in [13].

The optimization stage aims to obtain optimal process parameters that maximize process performance (γ). Since metaheuristic effectiveness highly depends on the hyper-parameters, four algorithms are tested: the most frequently used ones (GA and SA) that are highly affected by their settings, and two recent algorithms (TLBO and Jaya) that are free of the algorithm specific hyper-parameters. For the TLBO and Jaya, the only factors to be specified are common for all EAs: population size and total iterations count. In this work, the population size of $5n$ (n is number of design variables, i.e. process parameters) is adopted for all four algorithms, with the total iterations count of 2000.

The GA is the most frequently used metaheuristic in process optimizations. It searches throughout the continual space of solutions using a population, so it belongs to EAs [14]. To improve the current members, it employs several genetic operations: selection, crossover, mutation, migration. Members with the highest objective function are adopted for the next iteration, and iterative procedure is repeated until a total iterations count is met. In this work, three values of the two major operators are tested: (i) selection: *stochastic uniform*, *roulette wheel* and *tournament*, (ii) crossover: *single point*, *two points* and *arithmetic*. Hence, in total 9 GAs are developed for each optimization problem.

The SA algorithm does not belong to EAs, since it is based on a point-to-point search. It mimics the metal annealing process: heating a material to the melting point and slowly decreasing the temperature to maintain a thermal equilibrium. Starting with a sufficiently

high initial temperature, new points are generated using an annealing function. The probability of a new point acceptance is defined by the expression [15]:

$$P(E, T) = \exp\left(-\frac{\Delta E}{K \cdot T}\right) \quad (4)$$

where ΔE is the difference of energy, i.e. difference of the objective function between a new and the old point, T is the current temperature and K is the Boltzmann constant. The temperature decrease is controlled by a temperature function. When a temperature is very low, a reannealing is performed (controlled by reannealing interval) to raise up the temperature. The above procedure continues until a specified number of iterations is reached. Different values of the four major parameters are tested: (i) initial temperature: 10, 100 and 500, (ii) temperature function: *exponential*, *fast* and *Boltzmann* function, (iii) annealing function: *fast* and *Boltzmann* algorithm; (iv) reannealing interval: 10 and 100. Therefore, 36 SA algorithms with heterogeneous settings are run for each problem.

The TLBO algorithm is EA that mimics the teaching-learning behavior in a classroom: the population is composed from students; the design variables are teaching subjects; the student's knowledge (grade) refers to the objective function. First, students are learning from the teacher to improve their knowledge. For the j^{th} subject ($j = 1, \dots, n$) in the k^{th} teaching-learning cycle ($k = 0, 1, \dots, I_{\max}$; I_{\max} is the total iterations count), the updated solution for the i^{th} student is [16]:

$$X_{\text{new},i,j}^k = X_{\text{old},i,j}^k + r(TG_j^k - T_F LG_j^k) \quad (5)$$

where TG_j^k is the teacher grade, LG_j^k is the average students' grade, r is a random number between 0 and 1, and T_F is the teaching factor between 1 and 2. If a new solution is better than the old one, it is adopted for the next stage where student knowledge is further improved by interacting with other fellows. Assuming that u^{th} student learns from the v^{th} student ($u \neq v$), the updated solution in this stage is computed based on the objective functions $f(X_u^k)$ and $f(X_v^k)$ of u^{th} and v^{th} students, respectively [16]:

$$\begin{aligned} \text{if } f(X_u^k) < f(X_v^k) \text{ then } X_{\text{new},L,u,j}^k &= X_{u,j}^k + r(X_{v,j}^k - X_{u,j}^k) \\ \text{else } X_{\text{new},L,u,j}^k &= X_{u,j}^k + r(X_{u,j}^k - X_{v,j}^k) \end{aligned} \quad (6)$$

Based on the objective function assessment for an updated solution, the students with the best knowledge are involved in the next iteration. The process is repeated until a termination condition (i.e. the total number of iterations) is met.

The Jaya algorithm is a very simple EA, based on a straightforward principle: the optimal solution is obtained by approaching the best and moving away from the worst solutions. For the i^{th} design variable ($i = 1, \dots, n$) of the k^{th} population member ($k = 1, \dots, m$) in the l^{th} iteration, the candidate solutions are calculated according to the design variables of the best ($X_{i,\text{best},l}$) and the worst ($X_{i,\text{worst},l}$) members in the population [17]:

$$X'_{i,k,l} = X_{i,k,l} + r_{1,i,l}(X_{i,\text{best},l} - |X_{i,k,l}|) - r_{2,i,l}(X_{i,\text{worst},l} - |X_{i,k,l}|) \quad (7)$$

where $r_{1,k,l}$ and $r_{2,k,l}$ are random numbers between 0 and 1. The objective functions of updated solutions are evaluated; solutions that enhance the objective are adopted for the next iteration. The process is reiterated until the specified iterations count is reached.

Effectiveness of the four algorithms are compared in respect to the accuracy of the obtained solution (the highest objective function) and the convergence rate (minimal number of iterations needed to locate the global optimum).

3. Implementation of optimization methodology in semiconductor industry

3.1. Use case 1: Thermosonic copper wire bonding

Aiming to ensure a reliable performance of a microelectronic device assembled using a copper wire, the objective was to establish a solid bond between copper wire (diameter 50 μm) and aluminum pads at the die that contains integral circuitry, in a thermosonic wire bonding of the power amplifier device. The following control parameters were varied at two levels: contact power (CP_{Cu}), contact force (CF_{Cu}), base force (BF_{Cu}), base power (BP_{Cu}). Two responses were observed at the output: (i) the average ball shear test (BS) showing the average strength of the intermetallic connections between Cu wires and Al die pads (in a device with 41 ball bonds), and (ii) the number of oxide damages (NoOD) found as a failure mode after shear tests in a device. The objective is to find an optimal parameters set to achieve the nominal BS value of $160 \cdot 10^{-2}$ N and to minimize NoOD occurrence that indicates corrosion micro cracks inside the die pad.

The experimental plan was based on the orthogonal array (OA) L_8 including five added replicates [18] (Table 1). The responses were converted into QLs as per formula (1). Application of PCA over NQLs resulted with two independent PCs, integrated in the process performance (γ) using weights obtained from PCA: [0.684; 0.316]. The best BP ANN with topology 4-7-1 established an accurate process model (process parameters vs. γ), with minimal error (MSE=4.4 10^{-4}) and maximal R (99.8).

Four algorithms were employed in the optimization stage. As seen from Table 2, TLBO found a marginally better solution than Jaya; their solutions are significantly better than the solution obtained by SA; the maximal objective found by GA is lower than the best SA algorithm result. Figure 1 shows convergence of the best algorithms (ones that found maximal objectives, as presented in Table 2), i.e. change in the objective function along iterations. It could be seen that GA converged in 260th iteration, SA in 20th iteration, TLBO in 1876th iteration and Jaya in 276th iteration. Therefore, SA was the fastest algorithm followed by GA, but both failed to locate the global solution. TLBO was the slowest; Jaya convergence rate is comparable to GA. Although TLBO was the slowest algorithm, it is important to note that TLBO and Jaya obtained the global or near-to-global solution contrary to GA and SA. The most disperse results are obtained by GA; TLBO and Jaya showed a narrow range of results obtained in multiple runs since they do not have algorithm specific hyper-parameters to be tuned, contrary to GA and SA.

Table 1. Use case 1: experimental results, NQLs and process performance measure

trial no.	Control parameters				Responses		NQL _i (k) i=1,2; k= 1,...,13		γ_k k=1,..,13
	CP _{Cu}	CF _{Cu}	BF _{Cu}	BP _{Cu}	BS	NoOD	NQL _{BS}	NQL _{NoOD}	
1	0	250	200	40	110.15	1	0.846	0.01	0.3830
2	0	250	200	80	175.81	10	0.129	1.00	0.3536
...
13	30	250	350	40	146.57	0	0.112	0.00	0.8235

Table 2. Use case 1: results of four metaheuristic algorithms

Metaheuristic algorithm	GA	SA	TLBO	Jaya
Objective function range	0.9255÷0.9687	0.9702÷0.97054	0.97351÷0.97362	0.97351÷0.97361
Optimal parameters range	[15÷30.1; 400; 299.2÷306.4; 40÷45.5]	[39.3÷40; 395.1÷400; 380.2÷395.6; 62÷69]	[23.3÷23.4; 400; 261.1÷264; 40]	[23.3÷23.8; 400; 261.1÷264.5; 40]
Maximal objective	0.96870	0.97054	0.97362	0.97361
Optimal parameters	[28; 400; 299; 40]	[40; 400; 385; 66]	[23.4; 400; 264; 40]	[23.3; 400; 264.4; 40]
Iterations to max. objective	260	20	1876	276

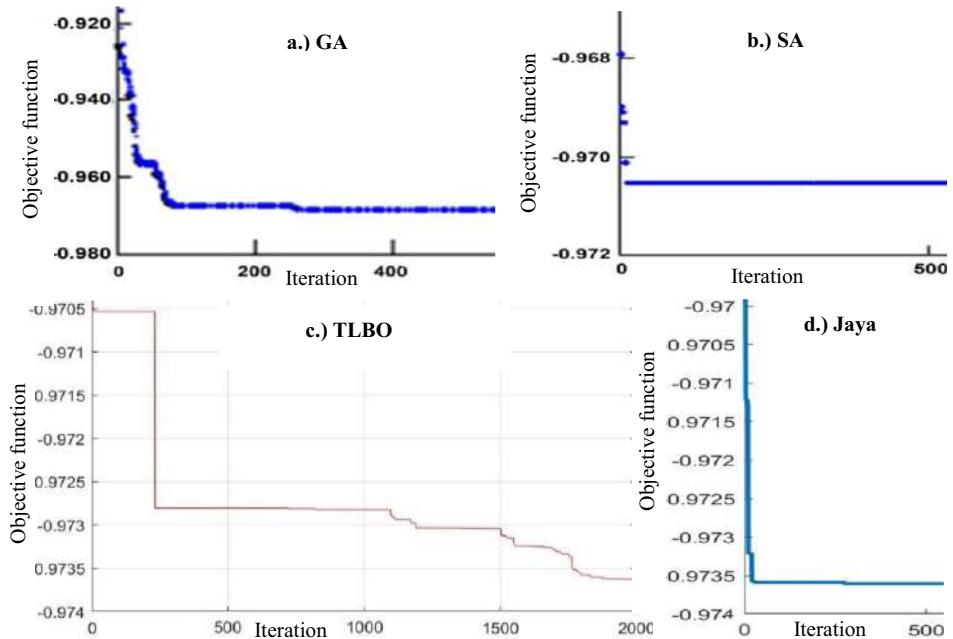


Figure 1. Use case 1 - algorithms convergence: a.) GA, b.) SA, c.) TLBO, d.) Jaya algorithm.

3.2. Use case 2: Thermosonic gold wire bonding

From the control charts it has been detected that behavior of a smaller group of machines in gold wire ($75\mu\text{m}$) bonding significantly deviates from the common machine outputs, i.e. the gold ball bond characteristics. Therefore, an experiment was performed to establish an optimal process parameters for the whole group of machines, considering a gold bond formation at the aluminum die pads. Machine M1 was selected as a representative of a major group with common performance, while machine M2 represented a smaller group of machines with deviant outputs. Two major control parameters were considered at three levels: base power (BP_{Au} : 55, 65 and 75 mW), and base force (BF_{Au} : 85, 100 and $115\ 10^{-2}$ N). Three responses were observed at the output, with the objective to reach a nominal (target) value: the average ball shear test value (BS; target= $270\ 10^{-2}$ N), the average ball diameter (D; target= $193\ \mu\text{m}$), and the average ball height (H; target= $50\ \mu\text{m}$) in a device with 42 bonds. Responses were measured for each individual bond to obtain mean values and assess variability needed to compute QLs.

The experiment was based on OA L₉ with five replicates [19] (Table 3). PCA was applied over NQLs, for M1 and M2 separately. The obtained PCs were synthesized into the process performance (γ) using weights from PCA: [0.667; 0.246; 0.087] for M1, and [0.915; 0.066; 0.019] for M2. In the modeling phase, BP ANN models with topology 2-9-1 showed the best performance for both machines (M1: $\text{MSE}=1.9\ 10^{-4}$, $R=0.97$; M2: $\text{MSE}=1.5\ 10^{-4}$, $R=0.97$). In optimization, all algorithms found the global optimum for M1; GA failed to locate the global solution for M2. GA showed a high dispersion of results obtained by different tunings, while SA results were more homogenous. TLBO and Jaya obtained identical results in multiple runs, showing a high repeatability and robustness. TLBO and Jaya displayed a rapid convergence, slightly faster, in average, than SA; SA converged faster than GA (Table 4). In overall, TLBO and Jaya showed

remarkable performance due to full repeatability of results and fast convergence, while GA and SA results were dispersed (caused by different algorithm specific hyper-parameters), and GA get trapped into a local optimum for M2.

Table 3. Use case 2: experimental results, NQLs and process performance measure

trial no.	Control parameters		Responses			NQL _i (k) i=1,2,3; k= 1,...,14			γ_k k=1,...,14	
	BF _{Au}	BP _{Au}	BS	D	H	NQL _{BS}	NQL _D	NQL _H		
M1	1	55	85	229.6	182	58.9	0.96	0.89	0.95	0.5263
	2	65	85	260	187	54.8	0.17	0.51	0.38	0.5871

	14	65	115	277.4	201	43.3	0.14	0.52	0.56	0.6362
M2	1	55	85	232.2	180	58.7	0.99	0.90	0.99	1.0000
	2	0	85	262.4	188	55.8	0.17	0.27	0.56	0.3733

	14	65	115	281.7	199	43.85	0.18	0.21	0.49	0.5769

Table 4. Use case 2: results of four metaheuristic algorithms

	Metaheuristic algorithm	GA	SA	TLBO	Jaya
M1	Objective function range	0.8807±0.88120	0.8810±0.88120	0.88120	0.88120
	Optimal parameters range	[85; 99±100]	[85; 99±99.5]	[85; 99]	[85; 99]
	Maximal objective	0.88120	0.88120	0.88120	0.88120
	Optimal parameters	[85; 99]	[85; 99]	[85; 99]	[85; 99]
	Iterations to max. objective	8	21	8	15
M2	Objective function range	0.7081±0.71280	0.7578±0.75801	0.75801	0.75801
	Optimal parameters range	[85; 95±100]	[85; 98.5±99]	[85; 99]	[85; 99]
	Maximal objective	0.71280	0.75801	0.75801	0.75801
	Optimal parameters	[85; 95]	[85, 85]	[85, 85]	[85, 85]
	Iterations to max. objective	710	3	3	3

3.3. Comparison with optimization methods from the literature

Benefits of the suggested evolutionary-based optimization methodology are demonstrated in comparison with the four frequently used methods from the literature: RSM [1], methods proposed by Fung&Kang [3], Liao [5] and Yang *et al.* [6]. The proposed methodology (with any of the four metaheuristics) surpasses the benchmarked methods in both studeis (Table 5), due to: (i) search over a continous space to detect a global solution; (ii) weaknesses of the methods from literature in terms of adressing correlations among responses, developing a process performance in a completely objective manner or a tendency towards local solutions, as discussed in the introduction.

Table 5. Use cases 1 and 2: comparison of results obtained by different optimization methods

Use case	Optimization methodology	RSM [1]	Fung & Kang (2005) [3]	Liao (2006) [5]	Yang et al. (2014) [6]	Proposed method with GA	Proposed method with SA	Proposed method with TLBO	Proposed method with Jaya
1	Maximal objective (process performance)	0.8514	0.9613	0.9286	0.9613	0.96870	0.97054	0.97362	0.97361
	Optimal parameters	[20; 350; 300; 45]	[30; 400; 350; 40]	[30; 400; 350; 80]	[30; 400; 350; 40]	[28; 400; 299; 40]	[40; 400; 385; 66]	[23.4; 400; 264; 40]	[23.3; 400; 264.4; 40]
M1	Maximal objective (process performance)	0.6303	0.5637	0.5768	0.5637	0.88120	0.88120	0.88120	0.88120
	Optimal parameters	[65;100]	[65; 115]	[75;115]	[65; 115]	[75; 100]	[85; 99]	[85; 99]	[85; 99]
M2	Maximal objective (process performance)	0.6379	0.6379	0.5876	0.5876	0.71280	0.75801	0.75801	0.75801
	Optimal parameters	[65;100]	[65;100]	[75;115]	[75;115]	[85; 95]	[85, 85]	[85, 85]	[85, 85]

4. Conclusion

Application of the four algorithms, within the proposed methodology, in two wire bonding studies demonstrated superior results of TLBO and Jaya over GA and SA, especially for the solution accuracy. The GA showed the worst performance, due to inability to find a global optimum, high dispersion of the results obtained with different hyper-parameters and slow convergence. The *stochastic uniform* selection with *single* or *two point* crossover gave better results than the other combinations. The SA performed better than GA; the most beneficial SA settings are: initial temperature 100 or 500 with reannealing interval 10, with *Boltzmann* annealing and *Boltzmann* or *fast* temperature function. The TLBO and Jaya algorithms showed remarkable results, demonstrated also in a recent optimization study [20]. Therefore, they could be recommended for future applications. In this work they were applied with $5n$ population size (n is the number of design variables) and 2000 iterations, so these setting could be recommended for future applications for medium-size problems. Since TLBO needed a large number of iterations to reach the global solution in the first study, a higher iterations count might be demanded for more complex problems with a bigger number of design variables.

References

- [1] Myers RH, Montgomery DC. *Response surface methodology: process and product optimization using designed experiments*. Wiley, New York, 2002.
- [2] Taguchi G. *Introduction to quality engineering*. Asian Productivity Organization, New York, 1986.
- [3] Fung CP, Kang PC. Multi-response optimization in friction properties of PBT composites using Taguchi method and principle component analysis. *J Mater Process Technol* 2005;17:602–610.
- [4] Lin CT, Lin Cj. Multi-response optimization of wire bonding process for evaluating alternative wire material. *Microelectronics J*. 2020;106:104925.
- [5] Liao HC. Multi-response optimization using weighted principal component. *Int J Adv Manuf Technol* 2006;27:720–725.
- [6] Yang YS, Shih CY, Fung RF. Multi-objective optimization of the light guide rod by using the combined Taguchi method and Grey relational approach. *J. Intell. Manuf.* 2014;25:99–107.
- [7] Tsai JT, Chang CC, Chen WP *et al*. Optimal Parameter Design for IC Wire Bonding Process by Using Fuzzy Logic and Taguchi Method. *IEEE Access* 2016;4:3034–3045.
- [8] Tsai TN. A hybrid intelligent approach for optimizing the fine-pitch copper wire bonding process with multiple quality characteristics in IC assembly. *J Intell Manuf* 2014;25:177–192.
- [9] Li TS, Hsu CM. Parameter optimization of sub-35 nm contact-hole fabrication using particle swarm optimization approach. *Expert Syst Appl* 2010;37:878–885.
- [10] Hsu CM. Improving the lighting performance of a 3535 packaged hi-power LED using genetic programming, quality loss functions and particle swarm optimization. *Appl Soft Comput* 2012;12:2933.
- [11] Ramsay JO, Silverman BW. *Functional Data Analysis*. Springer, New York, 2005.
- [12] Lu M, Wevers K. Grey System Theory and Applications: A Way Forward. *J. Grey Syst.* 2007;10:47–54.
- [13] Sibalija T, Majstorovic V. *The advanced multiresponse process optimization – an intelligent and integrated approach*. Springer, Switzerland, 2016.
- [14] Sivanandam SN, Deepa SN. *Introduction to Genetic Algorithms*. Springer, Heidelberg, 2008.
- [15] Spall J. *Introduction to stochastic search and optimisation*. John Wiley & Sons, Inc., New Jersey, 2003.
- [16] Rao RV, Savsani VJ, Vakharia DP. Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design* 2011;43(3):303–315.
- [17] Rao RV. Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *Int. J. Ind. Eng. Comput.* 2016; 7(1):19–34.
- [18] Sibalija T, Majstorovic V, Miljkovic Z. An intelligent approach to robust multi-response process design. *Int J Prod Res* 2011;49:5079–5097.
- [19] Sibalija T, Majstorovic V. Novel Approach to Multi-Response Optimisation for Correlated Responses. *FME Transactions* 2010;38: 39–48.
- [20] Sibalija T, Kumar S, Patel GCM *et al*. A soft computing-based study on WEDM optimization in processing Inconel 625. *Neural Comput & Applic* 2021; DOI:10.1007/s00521-021-05844-8.

Smart Manufacturing and Jobs

Juvenal Mendoza-Valencia¹

Instituto Politécnico Nacional, México

Abstract. The proper use of new technologies, together with a good management of these by the workers, makes the workplace become a melting pot of new ideas, where professional development and motivation [1] make a difference with past and present times where monotonous and repetitive work was the constant. With the use of artificial intelligence, the cloud and big data, a set of intelligent sensors is necessary, which can monitor a critical variable within the production line, which requires an encrypted communication infrastructure and with information that can be monitored in real time, which allows a better performance of the company, making it a more pleasant place for staff and with working conditions that help to overcome them when performing complex tasks. Due to the educational level that many of the workers currently have and the fact of being born in a technified society, it causes them to be more interested in carrying out activities where they develop their creativity and skills, resulting in economic and psychological well-being.

Keywords. Manufacturing, worker and artificial intelligence

1. Introduction

The companies that take advantage of new technological developments in a creative way are those that remain in the market; for this they must promote among their staff the quality of having critical thinking, problem solving, analytical reasoning and collaborative work [2], for this it is necessary that the organization has a horizontal structure to be democratic, in this way they can have the conditions to adapt to change.

If we change our business approach from short term to long term, we can be in line with technological development, it follows a series of steps which range from: development, growth, maturity and obsolescence, in this way we can seek the staff of the different positions depending on the activity carried out table 1.

Table 1.- Job relationship vs technology development curve

<u>Market Stall</u>	<u>Technological level</u>
Direction	Onward development
Management	Growth forward
Technical	Maturity
Operative	Maturity

Source: self made.

If we consider the importance of table 1, we can look for personnel in the different areas where the greatest emphasis of hiring is the interest of future

¹Corresponding author; E-mail: juvenalmv69@gmail.com

knowledge of new technologies, so that you can relate what is current with future trends, which it will allow you to anticipate changes.

2. Method

For the following research, a field study was carried out in SMEs located in the metropolitan area of Mexico City, the purpose was to analyze their organizational culture, the degree of motivation of their workers and commitment of management and direction in this regard; For this purpose, interviews were conducted with the personnel of the three hierarchical levels. Likewise, a bibliographic search was carried out that covered topics of motivation in employment and technologies of industry 4.0, with the aim of analyzing their possible impacts on the factories of the future.

3. Workers past, present and future

When the industrial revolution began at the end of the 18th century, the vast majority of the population belonged to the agricultural sector, so knowledge revolved around it, the concern centered on famines, diseases and wars [3], At the beginning of the 20th century, industrial processes are consolidated and mass production is born, where workers are specialized in an activity within it, which is accompanied by an increase in goods, the uncertainty about what will happen in the future decreases, for what is more confident, being conducive to the functioning of the economy in general.

In the twentieth century, the population became more urban and only 2% were in agricultural activities [3], those born in it seek to develop a technical skill, to be able to insert themselves in industry, those who achieve it are guaranteed well-being, this is known as a welfare state policy [4].

The rhythm of work is determined by the market, so there is a duality between the possibilities of the factory and its demands. For this purpose, we have to look for indicators that allow us to know how its behavior is; This is reflected in an increase or decrease in the speed of the production chain; so the worker must be willing to bear the stress that it causes.

It is said that in the industrial world nothing is more constant than change, which means that there are times where there is a lot of work and others where it is scarce, which causes a cycle of hiring and firing, the worker must strive to be of the essential part of the production line, so that it has a permanence in the time and in this way reduce its economic uncertainty.

Over time, new technological developments appear, the requirements that they demand, make the skills of the workers have to change, from there the speed of change will be determined, which allows establishing training cycles, start-up and execution of them, This is where the difference between avant-garde companies and followers is achieved, the former develop revolutionary ideas, socialize them and set the standard for price, quality and variety; causing the latter to be forced to adopt said changes.

The process of socialization of the implementation of new technologies is a slow process, which implies the search for the worker's profile, many hours of training and

understanding of the potential of technology in the field of production; Many times workers, technicians and engineers do not understand their purpose, so at some time they work with very little vision of what really happens.

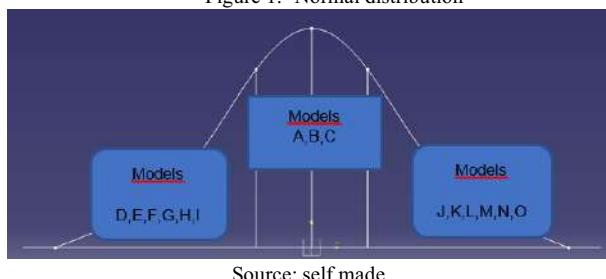
Leading companies will have enough time to put their ideas into practice on how to use new technologies and train staff to do so; on the other hand, those that do not do so will work with a lot of stress when they implement these changes, due to the fact that they have new parameters that mark the advances in the world of technology.

Many psychologists and sociologists have generated the idea that individuals fear change, being the reason why technology takes time to be exploited in factories; more recent studies indicate that this is true with certain generations like the baby boomer and X; but the millennials and the Z look for new experiences that break the monotony of daily work [4].

In recent studies, it is said that only 15% of workers are committed to their tasks [4], we have a big problem to motivate the workforce, so it is natural that companies prefer less developed countries to install their factories; But when their population becomes urban and therefore industrial, they will have productivity problems again, because future workers are no longer interested in the positions that are offered.

We must bear in mind that a production line has a normal statistical distribution in the demand for its products, the reason is due to the different seasons of the year, which means that 80% of production is concentrated in some models and 20% corresponds to the rest, this means that 80% of the time will be practically doing the same work and 20% will be doing different things (fig. 1), so the recruitment of new talented workers will be more and more problematic [5].

Figure 1.- Normal distribution



Source: self made.

If we take into account that every day, as a result of the processes of globalization, production volumes tend to increase by millions of units, this means that routine and repetitive operations will be practically all the time, so that workers, technicians and engineers will be subjected to great stress, to be able to stay in the workplace and if not, the only solution is to leave it, which explains why staff turnover is high 12% annually.

Hence the importance of making a good analysis of the production line, to determine with certainty which areas are dangerous, with a lot of physical effort, monotonous and that demand very little creativity; having this information, it is possible to plan the increase in capacity of the line demanded by the market, the intelligent technology that supports it and the profile of the worker that will keep it running.

Once it is decided to implement intelligent manufacturing in the lines, the human resources department can begin to study the jobs, looking in historical data for the

following points: worker interest in the activity, education demanded by the position, stress, creativity required and innovation in problem solving; when the results are obtained, a comparison is made with the percentage of rotation that was had before using the smart equipment: if it is greater, equal or less; consequently determining if their expectations are met.

In this way, we can influence the study programs in high school and universities, emphasizing the fields of interest shown by the personnel in the industrial area, with which the future profile of the students can be adjusted to the trends of the intelligent manufacturing, the purpose is to ensure that the resources invested in education are increased to have better workers, engineers and entrepreneurs, who can be leaders in the industrial branch.

4. Artificial intelligence and employment

If we start from the fact that all technological progress has as its primary objective the enhancement of human capabilities, which will allow the benefits of Science and Technology to spread in society and that it will determine the ways in which it will use it in its daily life, thus generating general well-being.

In order for artificial intelligence to take place, it is necessary that there be a communications infrastructure inside the factory, which allows wired and wireless communication, it is also very important to establish the technological structure by levels: operational, field, control, managerial and business [6], which will generate a set of data, for this purpose we will require a genetic control algorithm to transform them into different phases: 1st. an analysis, 2nd. in information, 3rd. knowledge, 4th. control and 5th improvement [7], each time an iteration occurs, the system will make small modifications to process them in the genetic algorithm, producing an evolution in decision-making over time [8].

If the man-made objects can be related to the environment, transfer the data to the cloud, analyze it, determine a better way to execute the work with the use of genetic algorithms and make a transfer to the emitting object; will result in better performance, over time all our devices will be operating optimally [9].

As there are millions of objects operating in the greatest diversity of places, environments, conditions and people's needs, in the cloud together with their analytics, they will be able to make probabilistic models such as: control charts, Kpi, Cpk, mean time between failures and maintenance 4.0; In order to take advantage of the wealth of massive information, Big data will be used [10], the results will be adjusted to reality, give recommendations for use, maintenance and replacement, with which the confidence of the production line will tend to 100%.

If we have already defined what we are going to do with the data generated by intelligent machines and the evolution in time that it entails, it cannot be said that up to this point the task is already completed, the human being performs: a geographical mobility, he has social relations, has his own point of view, understands public and business policies, market needs and understands his social environment; which are changing dynamically, that is, it is a unique being in the world and therefore has an

infinite wealth of information, which it would be necessary to consider to reduce the duality between what the factory offers and what the market demands.

For this purpose it is necessary for the company to develop a network intelligence to have the information required by each of the members, with the aim of communicating in real time, to know what are: the activities to be developed, the incidents, the models to be produce, volume, cancellations, etc.; These computer programs must be executed on 6G smartphones in order to meet the communication demands, which must run at speeds of 1×10^{-9} s [11], which allows supporting internal and external requirements of the IIoT.

As the work of data collection, analysis and execution will be in charge of intelligent machines, the human environment is focused on the development of innovation and collaborative work; For this purpose, the members of the areas of the company must be mature enough to leave behind the problems that arise in current companies such as: envy, hatred, vanity and pride; which do not allow the group to work together; those in charge of production have to be solving these rather than looking to the future.

Smart work centers will be more concerned with analyzing the psychological conditions of their members in order for their cooperative spirit to emerge and in this way generate innovative developments whose purpose is an original point of view, a product of the culture of their members. participants.

If we rely on artificial intelligence, as a way to put these ideas into practice on the production line, we will therefore have original products that attract the attention of the consumer, due to the practicality of their ideas, ease of maintenance and solutions to problems. specific that are presented at a given time, so that the buyer can be satisfied with the property purchased.

The new technologies of the 4th industrial revolution such as: artificial intelligence, big data and the cloud; They must allow us the emergence of new relations of production, where the human spirit will set the standard [12] if the purpose is to ensure that all internal and external participants in the production chain have the common benefit as their objective, it will be possible to leave behind practices such as the search for places with low wages; By changing our purpose, the markets where the plants settle will develop, creating new consumers.

With an increase in demand, the work of producing the goods is left to the intelligent system; Therefore, the workers will focus on solving fundamental problems such as: level of flexibility of the plants, ecological impact, recycling, caring for the environment, assigning the good to the customer at the time he makes the purchase, working with the social part , benefits and damages that our product may have, governmental aspects, future trends [13].

When we have a work group that relies on artificial intelligence, boring, repetitive, monotonous and dangerous jobs are left to machines [14]; Therefore, the culture that each of its members has becomes very valuable, due to the fact that it defines the ways of being of the people where they are: their dreams, joys, their fears, anguish, hopes, longings, etc. where the pursuit of happiness is the reason for their lives.

The word work will change its meaning to self-realization, own identity, cultural expression of the group and melting pot of new ideas, so that workers will be more identified with the company and their commitment in a percentage of 80 to 90% and the

staff turnover of 2 to 3%, this should be reflected in an increase in productivity, quality and variety.

Currently it is said that 84% of the value of a company is represented by the intangible part and 16% by the tangible; Until now, the creation of innovative products has been carried out outside the factories, so the people involved in their development lack knowledge of how this is produced, so the result is products that do not respect the life of the worker they are polluting and when their life cycle ends, it becomes an industrial waste problem, so it is common to see them anywhere.

We must seek that in future smart factories there is no such dualism and that creators and innovators have direct contact with smart manufacturing and in this way seek to solve the problems that arise at the time of good manufacturing, transportation, marketing, consumption and recycling; with which a new sustainable economy can be generated, reflected in cleaner environments for the future [15].

If the innovators work hand in hand with the workers, they will provide mutual feedback and 100% ecological products will come out throughout their distribution chain, artificial intelligence will be able to support activities with a lot of stress due to a punctual increase in demand and will find solutions that help to complex decision making, with multiple variables and short response time.

We need quality jobs, with complex activities, the creation of new consumers, the distribution of wealth to guarantee social stability and unrestricted respect for the dignity of the person, all of which will translate into a new social approach, which makes people breathe in to more challenges big.

5. Results

Companies have a short-term focus, which limits them to prepare for the future, for this purpose they seek to have a competitive advantage in the area in order to carry out the work they carry out on a daily basis and in this way they hope to stay in the market; which sometimes depends on their geographical location, the service that the product demands, the opportunity time, cost of the good and variety, so if they take advantage of them, their permanence can result.

With regard to the worker, the managerial part is the one that is most aware of the importance of permanence, promotion, development and organizational culture; Therefore, they establish strategies to achieve said purposes; but if for the achievement of these it is necessary to disburse additional amounts of money to which it is normally destined, the management is strongly opposed, they consider it an expense, due to the fact that at any time they can go and lose the investment.

6. Discussion

Workers are being influenced every day by the appearance and use of new intelligent technologies, they allow them to relate horizontally and vertically in the company, having by the mere fact of doing so, a combinatorial power for solving problems to

almost unlimited real time; Not wanting to recognize this new reality leaves us at a disadvantage to those companies that do.

For the use of new technologies, what has to be considered is the degree of maturity of these, in this way the real potential is known, with which the possible application can be defined with certainty; Likewise, it is more attractive for the worker to develop an activity within the factory that gives him the opportunity to grow as a person, interact and develop his full potential.

7. Conclusion

We are in the 4th industrial revolution where new technologies are appearing, the most important for the manufacturing area is Artificial Intelligence, for this purpose we must define what we want from this set of new technological advances, in order not to repeat the error that it was committed in the 3rd industrial revolution where we only focused on putting it into practice, forgetting that it should serve the human being; so after 40 years we have problems of: distribution of wealth, ecological, health, social, etc.; at this moment we have in our hands the power to make a change where all participants benefit.

We also have to consider that once the guidelines that will mark said transformation are established, a dynamic that supports them is generated and over time it only grows, so it is impossible to make any change once it is underway.

It is worth noting the fact that, in the history of industrial revolutions, when the human being is not at the center (1st and 3rd), technological advance only benefits some people who have a close relationship with the centers of power and When the human being is the main objective, there is a benefit for society, as is the case of the 2nd industrial revolution and that could be the case of the 4th.

Acknowledgements

To the National Polytechnic Institute and the UPIICSA Postgraduate Studies Section and especially to the Master of Industrial Engineering, for all the support received for the preparation of this article.

References

- [1] Agrawal A., Gans J. Goldfarb A Prediction machines, 1^a ed. USA: Harvard Business Review; 2018.
- [2] Brynjolfsson E., McAfee A. The second machine age, 1^a ed. USA: W. W. Norton and company, Ing; 2014.
- [3] Noah Y. Homo Deus, 1^a ed. México: Debate; 2019.
- [4] Clifton J. Harter J. It's the manager moving from boss to coach, 1^a ed. USA: Gallup press.
- [5] Chopra S., Meindl P. Supply Chain Management, 5^a ed. USA: Pearson; 2013.
- [6] Veneri G., Capasso A. Hands-on Industrial Internet of things, 1^a ed. UK: ed. Packt; 2018.
- [7] Pérez M. Big Data, 1^a ed. México: Alfaomega; 2016.
- [8] García A. Inteligencia Artificial, 1^a ed. México: ed. Alfaomega; 2017.

- [9] Caiza G., Saeteros M., Oñate W., Garcia M. Fog computing at industrial level, architecture, latency, energy, and security, ELSEVIER 2020: Heliyon 6 (2020) e03706.
- [10] Ponce P. Inteligencia Artificial, 25^a ed México: Alfaomega; 2019.
- [11] Zhou Y., Liu L., Wang L., Hui N. Service aware 6G: an intelligence and open network base don convergence of communication, computing and caching. ELSEVIER 2020 May; B.V. on behalf of KeAi Communications Co. Ltd..
- [12] Phuyal S., Bista D., Bista R. Challenges, opportunities and future directions of Smart manufacturing. ELSEVIER 2020 sustainable Futures 2 (2020) 100023.
- [13] Beier G., Ullrich A., Niehoff S., Reibig M., Habich M., Industry 4.0: How it is defined from a sociotechnical perspective and how much sustainability it includes. ELSEVIER 2020 Jornal of cleaner Production 259 (2020) 120856.
- [14] Sgorbossa F., Grosse E., Neumann W. Human factors in produccion and logistics systems of the future. ELSEVIER 2020 Anual Reviews in control 49 (2020) 295-305.
- [15] Knapcikova L., Balog M. Industry 4.0: Trends in management of intelligent manufacturing System, 1^a ed. Switzerland: Springer;2019.

Uncertain Machine Load Forecasting Based on Least Squares Support Vector Machine

Qiaofeng MENG¹

CPC Wuhan Municipal Party School & Wuhan Administration College, Wuhan, China

Abstract. Machine state is a very important constraint for job shop scheduling. For the uncertainty machine state, the paper proposes a machine load forecasting method based on support vector machine. The method reduces complexity and improves efficiency by eliminating a large number of unrelated input factors and selecting a small number of input parameters with strong correlation. The efficiency of the algorithm is verified by the production workshop instance.

Keywords. Machine load forecasting, support vector machine, particle swarm optimization

1. Introduction

Based on past machine load data, predicting machine load data for a certain day in the future is a short-term forecast problem. Traditional load forecasting methods such as trend extrapolation [1], time series method [2], exponential smoothing method [3], regression analysis [4] And the gray model method [5], etc., usually need to establish a mathematical model, which is more difficult to solve complex problems [6]. The load forecasting method based on time-frequency analysis classifies the load by frequency variation, including Fourier analysis method [7] and wavelet analysis method [8]. The load prediction methods described by the dynamic process include chaos theory [9] and Kalman filter algorithm [10]. The knowledge-based expert system [11] generalizes the factors affecting the load into knowledge and predicts future loads through some rules. The knowledge and rules of this method are more difficult to determine, the model is more relevant to specific problems, and the model is not easy to popularize and apply. The artificial neural network [12] is a nonlinear system composed of a large number of simple neurons. But the forecasting process of it is easy to fall into local minimum values. And the learning process is time consuming and easy to cause over-fitting[13]. The support vector machine can predict small sample data, the method is simple and has high efficiency, and the structured learning unit is adopted, thus avoiding the over-fitting problem and receiving extensive attention and research.

In this paper, the least squares support vector [14] machine nonlinear regression is used to predict the machine load. Compared with the standard support vector machine [15], the least squares support vector machine has the following characteristics[16]: the

¹ Corresponding author; E-mail: 1343854830@qq.com.

equality constraint replaces the inequality constraint; There are fewer model parameters required and the model is easier to determine[17].

The next section of the paper proposes an uncertain machine load forecasting scheme including forecasting process. The third section is the simulation experiment with actual production workshop data. In the fourth section the full paper is summarized.

2. Uncertain Machine load forecast

The machine load in this paper is defined as the sum of the processing times of all the processes assigned to a machine. A very critical step in the flexible job shop scheduling problem is to arrange the process onto the right machine. In the actual production workshop, some machines have more work-pieces to be processed, while some machines have fewer tasks or even idle for a long time. On the one hand, the completion time of the operation is prolonged, and on the other hand, other problems are caused. More power consumption and wear and tear of some machines are serious. To solve these problems, it is necessary to consider the load of the machine when arranging the process to the machine, and to balance the load of each machine as much as possible to achieve maximum production efficiency. Actual production is a real-time dynamic process, and there are many uncertainties. For example, orders can come at any time in the customized production mode. This requires the scheduling system to consider the instantaneous static machine load and possible future orders when arranging the work-piece processing machine. The dynamic load enhances the robustness of the scheduling system and reduces the number of rescheduling.

The prediction model of the least squares support vector machine is shown by the formula 1, where α_i is Lagrange multiplier, $K(x, x_i)$ is a function of the mapping from low-dimensional space to high-dimensional feature space, The calculation method of α_i , b is as described in the paper[18].

$$y = \sum_{i=1}^l \alpha_i K(x, x_i) + b. \quad (1)$$

The machine load consists of two parts: one is the existing work, and the other is the job that arrives randomly in the future. Since the arrival of the job is random, the real-time machine load is also uncertain. This paper predicts the future machine load based on known machine load data. When the machine is assigned, the current determined machine load and the predicted machine load are considered. The rule-based method, such as the machine with less total machining time, is more likely to be selected. To generate a robust scheduling scheme. For uncertain machine load factors, such as the arrival of random orders, etc. have a greater impact on scheduling, this paper predicts the uncertainty of machine load.

The key steps of the NPSO-SVM algorithm presented in this paper are described below.

Step 1. Preprocess historical data to generate training data according to the principle of similarity;

Step 2. Initialize. Select the radial basis kernel function, initialize the regularization parameter γ , the kernel function width σ , local Search threshold L_1 ;

Step 3. Using the improved PSO algorithm for parameter optimization;

1) initialization;

- 2) Calculate the individual fitness values in the population, and update the global optimal value and the individual optimal value;
- 3) Calculating the variance of the fitness value of the population Var ;
- 4) If $Var < L_1$, perform a local search;
- 5) The population evolved to the next generation;
- 6) Repeat (2)-(5) until the end condition is met.

Step 4. Calculate α_i , b , construct the prediction function $f(x)$, and predict the machine load with the prediction function.

3. Simulation

The improved particle swarm optimization algorithm was used to perform optimization experiments on the regularization parameter and the kernel width coefficient.

The optimal regularization parameter γ and the kernel width coefficient σ is set as follows.

Single shift: $\gamma = 100, \sigma = 3.3$. Double shift 1: $\gamma = 120, \sigma = 1.8$. Double shift 2: $\gamma = 118, \sigma = 7.9$. Overtime shift 1: $\gamma = 100, \sigma = 4.5$. Overtime shift 2: $\gamma = 100, \sigma = 4.2$. Three shifts: $\gamma = 100, \sigma = 3.8$.

The machines in the actual production workshop usually work according to the shifts of the operators. The shifts for each machine in this paper:

- 1) Single shift (08:30-12:00, 13:00-17:30)
- 2) Double shift 1 (07:00-15:00, 15:00-22:00)
- 3) Double shift 2 (07:00-19:00, 19:00-07:00)
- 4) Overtime 1 (08:30-12:00, 13:00-19:30)
- 5) Overtime 2 (08:30-12:00, 13:00-21:30)
- 6) Three shifts (07:00-15:00, 15:00-23:00, 23:00-07:00)

The daily shift of the equipment is set to one of six shifts, and the processing of the work-pieces is arranged according to the shift of the equipment. Shifts have a significant impact on machine load, so this article is tested separately for each shift.

In order to evaluate the quality of the solution, the average relative prediction error at t moment and the average relative prediction error shown in the formula 2 and formula 3 are used as Evaluation criteria, where $MAE(t)$ represents the average relative prediction error at t moment, MAE is the average relative prediction error, and $L_k(t)$ is the load of k for the machine predicted at t moment, $\hat{L}_k(t)$ is the actual load of the machine k , l is the number of machines, and N is the predicted number of moments.

$$MAE(t) = \frac{1}{n} \sum_{k=1}^l \frac{|L_k(t) - \hat{L}_k(t)|}{\hat{L}_k(t)} \times 100\%. \quad (2)$$

$$MAE = \sum_{t=1}^N MAE(t). \quad (3)$$

In order to verify the effectiveness of the NPSO-SVM algorithm proposed in this paper, the algorithm is proposed by the hybrid wavelet transform, artificial neural network algorithm (WT-ANN)[19] and the hybrid wavelet and support vector machine (C-WSVM) algorithm were compared[20]. The three algorithms were optimized for 30

experiments. According to the data of a factory from May 1, 2017 to December 1, 2017, the machine load at each hour of the day on December 2 is predicted, and the forecast is performed according to the shift. The single shift prediction result using the WT-ANN algorithm is shown in the figure 1, and the average relative error is 1.83%. The single-shift prediction result using the C-WSVM algorithm is shown in the figure 2, and the average relative error is 1.44%. The single-shift prediction result using the NPSO-SVM algorithm proposed in this paper is shown in the figure 3, and the average relative error is 1.09%. The average relative error of the double shift 1 is 3.31% using the WT-ANN algorithm. The average relative error of the double shift 1 is 2.04% using the C-WSVM algorithm. The average relative error of the double shift 1 is 0.76% using the NPSO-SVM algorithm. The average relative error of the double shift 2 is 3.52% using the WT-ANN algorithm. The average relative error of the double shift 2 is 2.57% using the C-WSVM algorithm. The average relative error of the double shift 2 is 1.25% using the NPSO-SVM algorithm. The average relative error of the overtime 1 is 3.32% using the WT-ANN algorithm. The average relative error of the overtime 1 is 2.17% using the C-WSVM algorithm. The average relative error of the overtime 1 is 1.41% using the NPSO-SVM algorithm. The average relative error of the overtime 2 is 3.37% using the WT-ANN algorithm. The average relative error of the overtime 2 is 2.35% using the C-WSVM algorithm. The average relative error of the overtime 2 is 1.42% using the NPSO-SVM algorithm. The average relative error of the three shifts is 5.05% using the WT-ANN algorithm. The average relative error of the three shifts is 3.3% using the C-WSVM algorithm. The average relative error of the three shifts is 1.9% using the NPSO-SVM algorithm. It can be seen from the figure and the experimental data that the prediction error of the NPSO-SVM algorithm proposed in this paper is better than the prediction error of WT-ANN algorithm and C-WSVM for the six shifts.

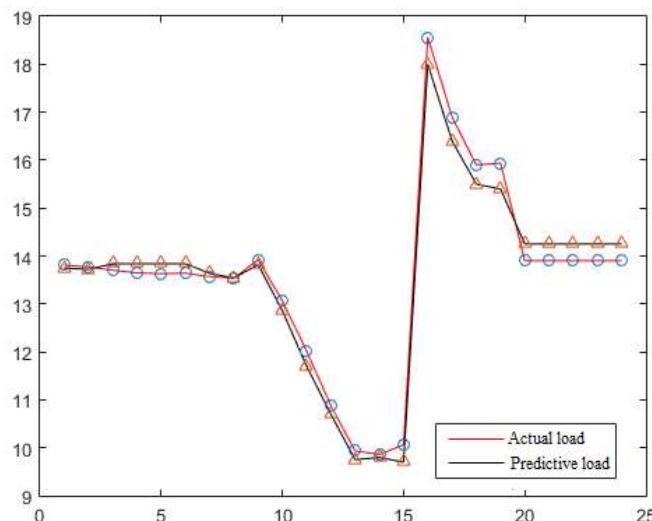


Figure 1. Prediction results using WT-ANN of Single shift.

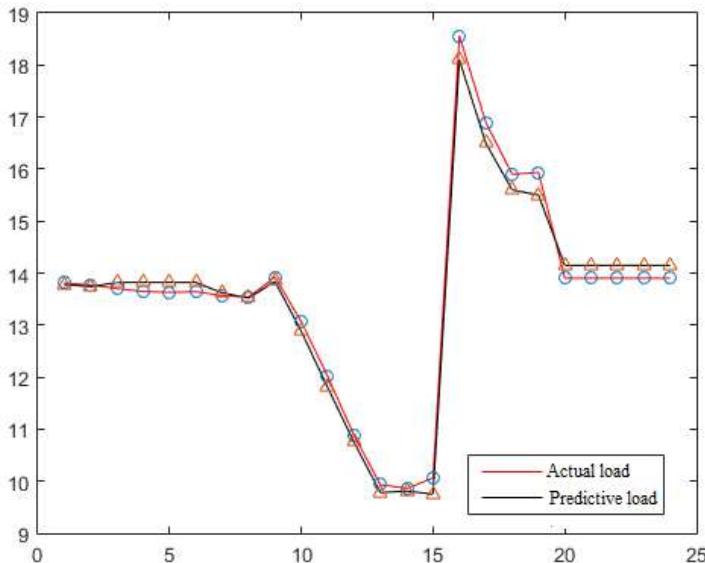


Figure 2. Prediction results using C-WSVM of Single shift.

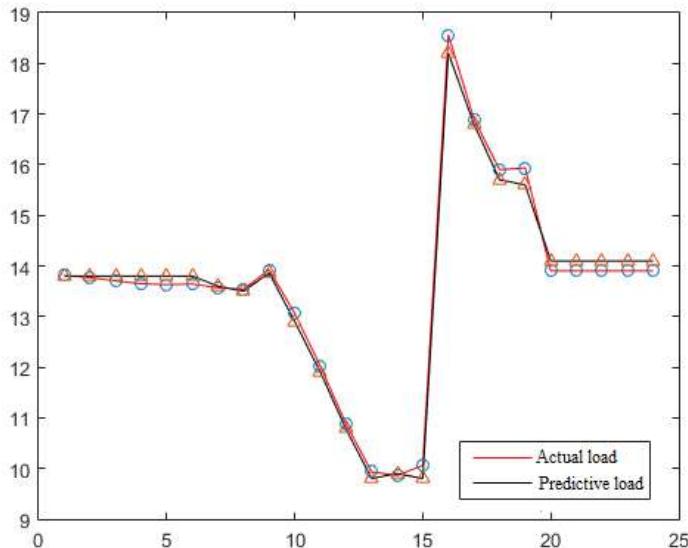


Figure 3. Predictin results using NPSO-SVM of Single shift .

The algorithm performance of the NPSO-SVM algorithm proposed in this paper, the WT-ANN algorithm and the C-WSVM algorithm are shown in the Table 1. As can be seen from the table, the NPSO-SVM algorithm proposed in this paper in the predicted average relative error MAE and the time performance is superior to the WT-ANN algorithm and the C-WSVM algorithm.

Table 1. Algorithm performance comparison

Shift type	WT-ANN		C-WSVM		NPSO-SVM	
	MAE(%)	Time(s)	MAE(%)	Time(s)	MAE(%)	Time(s)
Single	1.83	1468	1.44	1255	1.09	1163
Double1	3.31	1389	2.04	1219	0.76	1109
Double2	3.52	1393	2.57	1276	1.25	1187
Overtime1	3.32	1428	2.17	1302	1.41	1210
Overtime2	3.37	1430	2.35	1295	1.42	1196
Three shifts	5.05	1485	3.3	1356	1.9	1276

4. Summary

Aiming at the uncertain machine load, a machine uncertainty load forecasting method based on least squares support vector machine is proposed. The improved particle swarm optimization algorithm is used to predict the machine load. The results of simulation shows that this method is more effective than the other two popular methods.

References

- [1] Xia Jingjing. Research on load combination forecasting and third-party logistics of coal storage [PhD thesis]. Wuhan: Huazhong University of Science and Technology, 2013.
- [2] Hirata Y, Aihara K. Improving time series prediction of solar irradiance after sunrise: Comparison among three methods for time series prediction. *Solar Energy*, 2017, 149:294-301.
- [3] Shen H D, Wan Z K. A Dynamic Prediction Mechanism Based on Exponential Smoothing Method. *Computer Technology Development*, 2017, 27(7): 6-9.
- [4] Braun M R, Altan H, Beck S B M. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy*, 2014, 130(5):305-313.
- [5] Ming J, Fan Z, Xie Z, et al. A modified grey verhulst model method to predict ultraviolet protection performance of aging B.mori, silk fabric. *Fibers Polymers*, 2013, 14(7):1179-1183.
- [6] Xie M, Hong G Y, Wohlin C. A study of the exponential smoothing technique in software reliability growth prediction. *Quality Reliability Engineering International*, 2015, 13(6):347-353.
- [7] Anistratov D Y. Fourier Analysis of the Multilevel QD Method with α -Approximation For Time-Dependent Radiative Transfer Problems. *Transactions of the American Nuclear Society*, 2016, 115:491..
- [8] Yang X, Sun M. Power Consumption Prediction Based on Multi-Scale Continuous Wavelet Analysis. *Power System Clean Energy*, 2017, 33(6):67-70.
- [9] Li Dongdong, Zhai Zishan, Lin Shunfu, et al. Short-term load forecasting of microgrids based on chaotic time series method. *Journal of Electric Power Systems and Automation*, 2015, 27(5): 14-18.
- [10] Wang L, Khoshnevisan S, Luo Z, et al. Extended Kalman Filter for the Inverse Analysis of a Supported Excavation Based on Field Monitoring Data for Improving Predictions of Ground Responses. *Geotechnical and Structural Engineering Congress*, 2016, 1380-1391.
- [11] Hossain I, Hossain A, Choudhury I A, et al. Fuzzy Knowledge Based Expert System for Prediction of Color Strength of Cotton Knitted Fabrics. *Journal of Engineered Fibers Fabrics*, 2016, 11(3):2016.
- [12] Li Kangji, Su Hongye, Chu J. Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: a comparative study. *Energy Build* 2011;43(10):2893-9.
- [13] Doucoure B, Agbossou K, Cardenas A. Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. *Renewable Energy*, 2016, 92:202-211.
- [14] Sudheer C, Maheswaran R, Panigrahi B K, et al. A hybrid SVM-PSO model for forecasting monthly streamflow. *Neural Computing & Applications*, 2014, 24(6):1381-1389.
- [15] Wang G J, School B. Time Series Forecast of Stock Price Based on the PSO-LSSVM Predict Model. *Science Technology & Industry*, 2017.

- [16] Wang Cheng, Yang Sen, Meng Chen, et al. Research on a power supply combination fault prediction based on improved QPSO optimized SVR. *Computer Measurement and Control*, 2014, 22(5): 1342-1344.
- [17] Kang Cao. Research on short-term electric load forecasting model based on least squares support vector machine [PhD thesis]. Chengdu: Southwest Jiaotong University, 2012.
- [18] Wang G J, School B. Time Series Forecast of Stock Price Based on the PSO-LSSVM Predict Model. *Science Technology & Industry*, 2017.
- [19] Hooshmand R A, Amooshahi H, Parastegari M. A hybrid intelligent algorithm based short-term load forecasting approach. *International Journal of Electrical Power Energy Systems*, 2013, 45(1):313-324.
- [20] Wang J, Shi Q. Short-term traffic speed forecasting hybrid model based on Chaos-Wavelet Analysis-Support Vector Machine theory. *Transportation Research Part C Emerging Technologies*, 2013, 27(2):219-232.

Edge Loss for Remote Sensing Image Super-Resolution

Jiaoyue LI^a, Weifeng LIU^b, Kai ZHANG^c and Baodi LIU^{b,1}

^a*College of Oceanography and Space Informatics, China University of Petroleum, China*

^b*College of Control Science and Engineering, China University of Petroleum, China*

^c*School of Petroleum Engineering, China University of Petroleum, China*

Abstract. Remote sensing image super-resolution (SR) plays an essential role in many remote sensing applications. Recently, remote sensing image super-resolution methods based on deep learning have shown remarkable performance. However, directly utilizing the deep learning methods becomes helpless to recover the remote sensing images with a large number of complex objectives or scene. So we propose an edge-based dense connection generative adversarial network (SREDGAN), which minimizes the edge differences between the generated image and its corresponding ground truth. Experimental results on NWPU-VHR-10 and UCAS-AOD datasets demonstrate that our method improves 1.92 and 0.045 in PSNR and SSIM compared with SRGAN, respectively.

Keywords. Image super-resolution, remote sensing image, edge loss, generative adversarial network

1. Introduction

Image SR is to repair a high resolution (HR) image from its low resolution (LR) counterpart. Remote sensing images often have a low spatial resolution but have high requirements for the clarity of small objects in the vision, so using image super-resolution to process remote sensing images has immense application value.

In the past, quantities dictionary-based algorithms had been applied to super-resolution of remote sensing images, which was usually troublesome to capture sufficient deep features to restore satisfactory images.

The high-frequency information is very beneficial for restoring high-quality images. Fortunately, the appearance of residual structure makes it possible to obtain deep features using deep learning-based approaches. Various deep learning-based methods have emerged at the historic moment, and they can successfully seize much high-frequency information to repair the image. In particular, the Generative Adversarial Network [1] is a very successful deep network model and it is trained through adversarial methods to learn the distribution of real data, and the training purpose of generator and discriminator is exactly the opposite. The generated image confirms that a good perception effect can

¹ Corresponding Author: E-mail: thu.liubaodi@gmail.com.

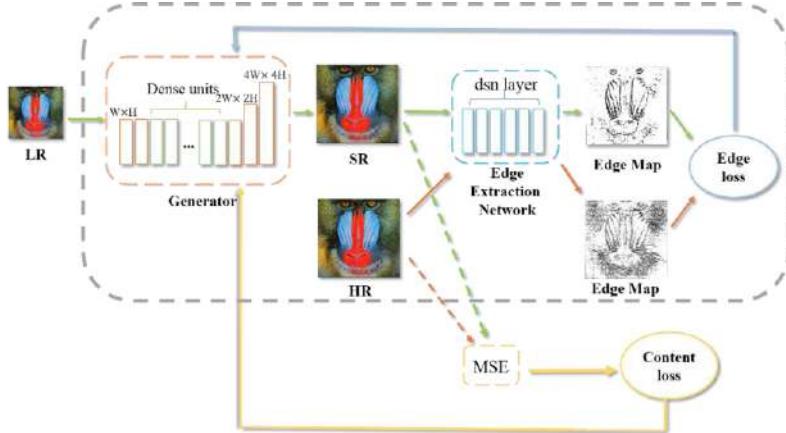


Figure 1. Overview of SREDGAN. First, the images generated by the generator and the ground truth images are input into an edge extraction network. Second, the edge maps created by the two kinds of images are differentiated to obtain edge loss. Finally, we propagate the edge loss to the generator and minimize the loss function to make it show solicitude for edges.

be obtained by using a generative confrontation network, but there is often a common phenomenon that the edges are blurred and the artifacts are in evidence.

Superb high-resolution images with sharp edges and realistic shapes, moreover, the edge information contained in remote sensing images is more abundant, and it becomes more important to increase the attention to the edge information acquisition during the training process. Therefore, we pose a densely connected edge-based generative adversarial network for image super-resolution (SREDGAN). The generator of SREDGAN draws support from a densely connected network to extract multiple features, and the network's training is adjusted continuously by a hybrid loss.

In general, the main contributions of our work are summarized in the following three points:

- Take edge maps into account, which is of great help to the sharpness of images, and replacing traditional edge extraction with pre-trained edge extraction networks.
- Designed an end-to-end network SREDGAN that is easy to train.
- Prove that our approach has achieved excellent results by comparing it with other advanced SR methods.

2. Approaches

We advance the SREDGAN model to generate a high-quality image with distinct edges and neural contents. First of all, input the LR image I^{LR} into generator G and output the SR image I^{SR} . Then, the I^{SR} and high-resolution images I^{HR} are sent to the discriminator D to distinguish the similarity between I^{SR} and I^{HR} . Our purpose is to find the appropriate loss function to make the generated images real. An overview of the generator of the proposed model is showcased in Figure 1.

For the most SR task, minimize the pixel-level and perceptual-level differences between I^{SR} and I^{HR} is widely adopted. However, it leads to uneven edges inevitably. In

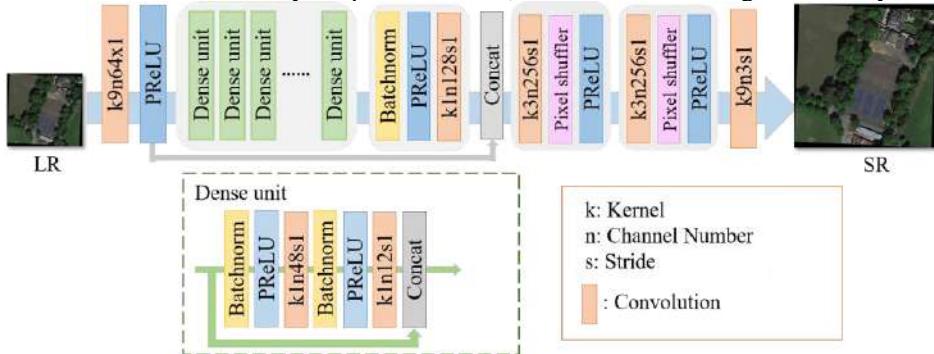


Figure 2. Structure of the generator which has five parts: shallow feature extraction, deep feature extraction, parameter transition, upsampling, image reconstruction.

this paper, we intend to add an edge loss to relieve the problem. The whole model is explained in Section 2.1, Section 2.2 plains describes our proposed edge loss and other functions of SREDGAN.

2.1. Structure of SREDGAN

Our generator is divided into five parts, the input image goes through a convolutional layer to get shallow features, and then the deep feature is extracted by sixteen dense units with a growth rate at 12. After several dense units, the parameters increases rapidly, which would bring substantial computational costs to the subsequent processing. Therefore, a 1×1 convolution is added to decrease parameters. To help gradient propagate, a long skip connection joins shallow features to deep features. Sub-pixel convolution [2] is a method to enlarge image size. It combines individual pixels on a multi-channel feature into a unit on a feature; the pixels on each feature are equivalent to sub-pixels on the reconstructed feature. We use two sub-pixel convolution layers to upscale the length and width of the input images by four times in our network. At last, a convolution implements the reconstruction of HR images. Our generator network is pictured in Figure 2.

Discriminator aims to identify which image is true and which one is false; it guides the generator to show solicitude for one direction for training, so a suitable discriminator is also very essential. SR intends to restore realist and high-resolution images. Therefore, it requires more high-frequency features. VGG [3] can extract deep features, and its structure is simple, so it is a pretty choice to use the VGG network as the discriminator framework, which is the similar to the discriminator of SRGAN [4].

2.2. Design of Loss Functions

To solve the situation that the image edges are indistinct, and the contents do not conform to the real, we construct an edge loss. The purpose of edge loss is to decrease the edge discrepancies between SR images and HR images. Although training with L1 or L2 loss will get a pleasant effect under objective criteria, visual results obtained are not ideal. The network does not put more emphasis on learning edge distribution, so the unreasonable contents and artificial scene frequently arise, and the edges are irrational

and unequivocal. Edge loss harmer at making object edges authentic and distinguishable. It makes the edges been paid close attention to network training.

Table 1. Results for test datasets, the data from label 1 to label 5 is the result of UCAS-AOD, the data from label 6 to label 10 is the result of NWPU-VHR-10, and the last column of the table represents the average.

No.	Bicubic	SRResNet	SRGAN	SRGAN2	SREDGAN
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
(1)	25.45/0.779	25.81/0.817	23.64/0.726	25.22/0.779	26.01/0.821
(2)	25.46/0.905	27.79/0.906	25.42/0.821	27.12/0.863	27.98/0.909
(3)	26.86/0.825	29.08/0.851	27.49/0.816	28.22/0.830	29.27/0.853
(4)	28.74/0.850	31.93/0.912	29.55/0.875	30.30/0.879	32.16/0.913
(5)	27.67/0.837	30.98/0.839	29.16/0.811	29.80/0.809	30.68/0.838
(6)	25.79/0.677	26.96/0.783	25.87/0.740	25.66/0.746	27.23/0.781
(7)	23.70/0.637	26.26/0.761	24.47/0.698	24.07/0.686	26.27/0.705
(8)	24.55/0.585	31.44/0.939	30.33/0.897	30.68/0.896	31.08/0.919
(9)	25.90/0.575	26.91/0.738	25.41/0.693	25.78/0.701	26.63/0.726
(10)	27.90/0.573	28.97/0.768	27.83/0.756	28.53/0.766	30.97/0.819
AVE	26.20/0.724	28.61/0.831	26.91/0.783	27.54/0.796	28.83/0.828

Nowadays, there are amounts of methods about edge extraction, although non-deep learning ways are easy, the kind of edge it extracts is too single. The salient disadvantage of the traditional methods is that their convolution kernels are unique and usually sensitive to a specific type of edge. HED [5] is a product of deep learning, and it has six outputs dsn1, dsn2, dsn3, dsn4, dsn5, and dsn-fuse, the first five represent edge detection maps of different feature layers respectively, and the last one indicates the fusion of the first five outputs. In our works, we have modified some convolution settings. We observe that the edges from the dsn1 layer is the most comprehensive. With the increase of network depth, the rougher the edge is, because HED [5] aims to detect the outline of an object, not the tiny edges. However, subtle edges are what we demand in our task, so we choose the dsn1 layer as a feature extraction layer. We input the generated images and HR counterparts into the HED network and minimize their feature discrepancies in the dsn1 layer. We named the error edge loss, and Eq. (1) shows the details.

$$L_{edge} = \frac{1}{WH} \sum_{i=1}^C \|h_i(I^{HR}) - h_i(I^{SR})\|_2 \quad (1)$$

Where $h_i(I^{HR})$ represents the dsn1 layer feature with the input of HR image, and $h_i(I^{SR})$ expresses the feature from SR image. Similar to the method of perceptual loss [6], the parameters of the HED network are trained in advance.

To ensure that the color of the generated image is as realistic as possible, we also take advantage of the common L2 loss and adversarial loss.

3. Experiments

3.1. Experimental Settings

We train our models on the RAISE dataset [7], which contains 8,156 common high-resolution images, and we get their low-resolution version pairs by Bicubic interpolation. Since large amount of training data is powerful to SR tasks, we flip images randomly and crop the size of HR images to 96×96 as well as low-resolution images to 24×24. We adopt remote sensing image datasets, NWPU-VHR-10 [8] and UCAS-AOD [9] as test datasets.



Figure 3. Result of the different methods on NWPU VHR-10.

Adam is chosen for optimizing training, and we set β_1 to 0.9 and set β_2 to 0.99. For optimum result, we assign the weight of content loss (L2 loss), adversarial loss and edge loss to 1, 0.0001 and 0.001, the weighted sum of these three losses constructs the total generator loss. The learning rate for the first 100000 iterations is set to 10^{-4} , and its remaining 100000 iterations are tenfold. We employ Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) as the criterion for testing.

3.2. Experimental Settings

We compare our method with Bicubic, SRResnet [4], SRGAN [4], SRGAN2 and our SREDGAN, and SRGAN2 is SRGAN with our edge loss. Table 1 and Figure 3 show the evaluation results and generated images of the two test datasets.

We can find the images generated by Bicubic are greatly blurry. Although the images generated by SRResnet have achieved excellent results in Table 1, the image is still not clear enough. Under the evaluation criteria of PSNR and SSIM, the results of SRGAN and SRGAN2 are not very satisfactory, but they have a good observation effect. Synthetically, the images generated by the generative adversarial network have higher resolution than the non-GAN network, but their results on PSNR and SSIM are slightly lower than those of non-GAN network.

In order to verify that increasing the edge loss helps the network to restore the image, we compared SRGAN and SRGAN2. The experimental results show that the results of SRGAN2 are better than those of SRGAN, whether in objective indicators or direct observation. The experimental comparison results of SRGAN and SREDGAN show that our generator model is more suitable for image feature extraction.

Comparing SRResnet and SREDGAN, we found that under the action of network structure and loss function, the PSNR and SSIM values of SREDGAN are close to the results of SRResnet, which shows that extracting rich features and paying more attention to the edges of the image will double the clarity and accuracy of the image.

4. Conclusion

In this paper, we propose a new model SREDGAN that pays more attention to edge information, and we adopt advanced edge extraction method to embed into an end-to-end network. With edge loss, models can generate an image whose edge is clear and close to the ground truth. Extensive experimental results demonstrate the effectiveness of our proposed methods for remote sensing super-resolution.

Acknowledgements

The Natural Science Foundation of Shandong Province, China(Grant No.ZR2019MF073, ZR2018MF017, No.ZR2017MF069), the Open Research Fund from Shandong Provincial Key Laboratory of Computer Network (No.SDKLCN-2018-01), Qingdao Science and Technology Project (No.17-1-1-8-jch), the Fundamental Research Funds for the Central Universities, the Major Scientific and Technological Projects of CNPC (No.ZD2019-183-008), and the Creative Research Team of Young Scholars at Universities in Shandong Province (No.2019KJN019).

References

- [1] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. Proceedings of the Advances in Neural Information Processing System;
- [2] Shi W, Caballero J, Huszár F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016 1874-1883.
- [3] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
- [4] Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z. Photo-realistic single image super-resolution using a generative adversarial network. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017 4681-4690.
- [5] Xie S, Tu Z. Holistically-nested edge detection. Proceedings of the IEEE international conference on computer vision; 2015 1395-1403.
- [6] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. European conference on computer vision. 2016; 694-711.
- [7] Dang-Nguyen DT, Pasquini C, Conotter V, Boato G. Raise: A raw images dataset for digital image forensics. Proceedings of the 6th ACM multimedia systems conference; 2015: 219-224.
- [8] Cheng G, Zhou P, Han J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Transactions on Geoscience and Remote Sensing. 2016; 54(12): 7405-7415.
- [9] Xia GS, Bai X, Ding J, Zhu Z, Belongie S, Luo J, Datcu M, Pelillo M, Zhang L. DOTA: A large-scale dataset for object detection in aerial images. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2018 3974-3983.

Phase Tracking Sequences for 5G NR in 52.6-71 GHz Band: Design and Analysis

Alexander MALTSEV^{a,1}, Andrey PUDEEV^b, Seonwook KIM^c, Suckchel YANG^c,
Seunghwan CHOI^c, Sechang MYUNG^c

^a*Nizhny Novgorod State University, Nizhny Novgorod, Russia*

^b*LG Electronics Russia R&D Lab, Moscow, Russia*

^c*LG Electronics, Seoul, Korea*

Abstract. This paper presents a novel approach to the phase tracking reference signal (PTRS) design for phase noise impact compensation in the 5G NR communication systems intended to work in a new 52.6 GHz to 71 GHz frequency band. For detailed problem illustration, the phase noise compensation algorithms are discussed and explained, from the basic common phase error (CPE) compensation to the MMSE-base inter-carrier interference (ICI) filtering. Performance of the different phase noise compensation algorithms is investigated for the baseline PTRS accepted in the current 5G NR specification and compared with the newly proposed approach to the PTRS design. This approach is based on nulling the subcarriers adjacent to the reference signals to minimize influence of the ICI on the estimation process. It was shown that new nulling PTRS design outperforms currently used distributed PTRS structure. In addition, numerical results represent a trade-off between the filter size and the amount of the allocated training resources to achieve better performance. It was shown that proposed PTRS structures and processing algorithms give ICI compensation level very close to optimal scheme and thus, different approaches (such as time domain compensation) may be required for further progress.

Keywords. 5G NR, beyond 52.6 GHz, CPE, de-ICI filtering, OFDM, PTRS design, phase tracking, phase noise

1. Introduction

Recent advancements in the communication technologies and an ever-growing data traffic capacity demands drive the 5G NR towards higher frequency bands and subjected a new work item on the support of the frequencies from 52.6 GHz to 71 GHz [1]. It is noted [2] that one of the key challenges for the OFDM systems in the millimeter-wave bands is a phase noise that causes severe inter-carrier interference (ICI) and prohibits the operation of the spectrally efficient higher order modulations without applying a specific phase tracking and ICI compensation algorithms. The performance of these algorithms strongly depends on training sequence design and functions. Although such phase tracking sequences are already specified in the 5G NR standards, the expansion to the 52.6-71 GHz band may require reconsideration of the currently implemented sequence structures.

¹ Corresponding author: Alexander Maltsev, e-mail: maltsev@rf.unn.ru.

In 5G NR specifications [3], phase tracking reference signal (PTRS) is introduced for mitigating phase noise impact and is defined as the known pseudo-random QPSK-modulated training signal. PTRS can be allocated within a time-frequency resource grid with configurable parameters for periodicity (or density) in time and frequency. In particular, the frequency domain parameter K and the time domain parameter L correspond to PTRS mapping for which PTRS resource element is allocated for every K -th resource block in frequency domain and for every L -th OFDM symbol in time domain. Variation of these parameters may affect the accuracy of the phase estimation and phase noise compensation.

The impact of phase noise on the OFDM system performance is well studied in the literature using theoretical and simulation analysis based on a different noise models [4]. For OFDM systems, phase noise causes so-called Common Phase Error (CPE) shift for all subcarriers in the signal band and, secondly, to the loss of subcarrier orthogonality due to the ICI effect. For reliable system operation, especially using the higher order modulations, these effects should be compensated. It should be noted that for lower order modulation, typical values of additive (thermal) noise are significantly larger than phase noise influence, and the latter can be neglected.

The CPE term may be efficiently estimated with the uniform PTRS grid specified in the Rel-15 [3] using relatively simple processing. This may be especially efficient in the multipath environment with the frequency-selective channel.

On the contrary, the ICI term is more difficult to estimate and compensate than the CPE term. Such processing typically requires per-symbol matrix inversion operation to obtain LS or MMSE estimation of the phase noise realization and further (de)convolution of the received signal to obtain noise-free data in the frequency domain – for example as in [5][6].

A perspective direction may be a design of the specialized PTRS structures, which is more suitable for the ICI compensation in the sense of performance and implementation complexity. In the recent works [7][8] a block PTRS structure is proposed and analyzed. With such approach, the PTRS resource elements are grouped in chunks, so the neighboring REs interference can be known. A new approach, based on similar principle but different implementation is proposed in present paper. Considered PTRS structure design is based on the adjacent subcarrier nulling and allows not only reducing the number of operations but also isolating PTRS from the random data influence.

An important role in the phase noise impact on the OFDM signals plays the ratio between the phase noise bandwidth and the OFDM subcarrier spacing (SCS) which in 5G NR can be largely varied [9]. The part of the phase noise inside the SCS affects only the given subcarrier and causes the phase shift error (or CPE). The part of the phase noise outside the SCS affects neighboring subcarriers and causes ICI.

Figure 1 shows the power spectral density (PSD) of the phase noise, according to the model currently accepted as a baseline for simulations and analysis [10]. For convenience, 5G NR SCS values (including those that are not supported in current NR specification but discussed to be introduced for frequency range from 52.6 GHz to 71 GHz) are shown along the frequency axis.

It can be seen that the amount of out-of-SCS noise corresponding to the ICI term is different for different subcarrier spacing, so the higher SCSs are much less susceptible

to the ICI effect of the phase noise. These qualitative observations further will be verified by qualitative analysis using the link layer simulations of the 5G NR system.

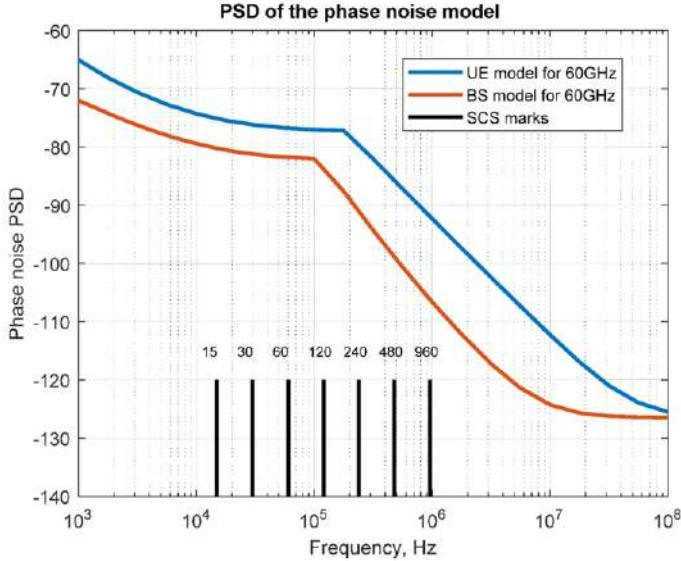


Figure 1 PSD of the phase noise based on model adopted in 3GPP document [10]

2. Phase noise impact on the OFDM signals

Analysis of the phase noise influence on an OFDM system performed in a number of works [4][5][6] show that the overall phase noise impact can be divided in the two components – the CPE, and the ICI – see Eq. (1), in details derived in [11]. The first term \bar{J}_0 is common for all subcarriers on the given OFDM symbol, but largely fluctuating for different symbols. The second term describes influence of the adjacent OFDM subcarriers on the given one due to signal spreading caused by multiplicative phase noise.

$$Y_k = H_k S_k \bar{J}_0 + \sum_{i=0, i \neq k}^{N-1} H_i S_i \bar{J}_{k-i} + \eta_k^{AWGN}, \quad (1)$$

where S_k is the symbol transmitted at the subcarrier with index k , through channel with H_k transfer function value at the same subcarrier and affected by additive noise η_k^{AWGN} . Convolution is performed over all N subcarriers, and term \bar{J}_k is the phase noise realization in frequency domain at the k -th subcarrier defined as:

$$\bar{J}_k = \frac{1}{N} \sum_{n=0}^{N-1} e^{j\phi_n} e^{-j2\pi nk/N} \quad (2)$$

In the general communication system phase noise occurs at the transmit side for the carrier frequency upconversion, and at the receive side for the downconversion, so strictly, there are two different phase noises – transmit and receive. For the flat AWGN channel, obviously, they are combined in one, with the same distribution and increased

power. For the frequency selective channel, the transmit noise is affected by the channel transfer function and thus, may not be described by simple convolution. Fortunately, as shown in [11], the representation of the phase noise impact as convolution in the frequency domain for the case of frequency-selective channels is applicable inside the channel coherence bandwidth, which is typically much larger than phase noise bandwidth, and thus representation (1) can be freely applied for the description of the phase noise impact on the 5G NR communication system in the considered channels with the delay spread up to tens of nanoseconds. Moreover, the PN spectrum realization has specific symmetry, where the J_i and J_{-i} tones taken at the symmetric positions around the DC component J_0 have equal imaginary parts and real parts with equal magnitudes, but inverted signs [11]. This important property will be used for improvement of the phase noise realization estimate.

Comparing the phase noise bandwidth with the OFDM signal subcarriers spacing (see Figure 1), or, in time domain, comparing the channel coherence time and OFDM symbols duration, we can observe that phase noise is not correlated on the adjacent symbols and thus, phase noise should be estimated and compensated on every OFDM symbol. The 5G NR specification allows changing also time density of the PTRS elements, but for considered SCSs/OFDM symbols durations, PTRSs should be allocated on every OFDM symbol for proper phase noise effects compensation. However, for the purposes of the carrier phase drifts tracking in the cases when phase noise is negligible, the PTRSs allocation can be much sparser.

While the first term in (1) may be rather easily mitigated by finding a common phase shift for the OFDM symbol, the compensation of the ICI may require more advanced algorithms and pilot structures. Current 3GPP 5G NR specifications define the PTRSs as a grid of selected resource elements (REs), sparsely distributed in the frequency domain in every 2nd or 4th resource block of 12 subcarriers (i.e., one PTRS element every 24-th or 48-th subcarrier). Whether it is enough or not, or should we consider another PTRS structures besides distributed grid – this is the subject of present paper studies.

3. Per-symbol Phase noise compensation approaches

3.1. Common phase error (CPE) approximation

The common phase error term \bar{J}_0 in (1) is equivalent to the same phase shift of the time-domain representation of the signal, constant during the OFDM symbol. So, the approximation of the phase noise with CPE in frequency domain is equivalent to the piece-wise constant per symbol phase noise approximation in the time domain [12]. Obviously, this approximation becomes more accurate for smaller symbol durations (i.e., for larger SCSs).

3.2. Linear trend approximation

The next logical step in the accuracy improvement of the phase noise evaluation and compensation is using a linear time domain approximation of the phase noise. In practice, such schemes can be implemented by exploitation on the cyclic prefix (CP)

properties on the OFDM signal, but in this preliminary analysis we assume ideal cases both for piecewise-constant and piecewise-linear approximations.

Figure 2 illustrates the principle and relative accuracy of the constant and linear approximations for different SCSs / OFDM symbol duration values. It can be seen that for higher subcarrier spacing value / shorter OFDM symbols, the constant approximation becomes more and more accurate.

For better understanding of the effects, we will plot the original phase noise PSD estimation versus the PSDs for the noise with compensated constant and linear trends for different SCS sets. The PSD is estimated over large number of realizations using Welch periodogram method [13]. The estimated PSDs in Figure 3 illustrate that estimation and compensation of the CPE works like filtering low frequency part of the phase noise, from the zero frequency to the SCS value. It can be seen that using the 960 kHz SCS allows more efficient ‘filtering’ with the larger bandwidth. The linear trend compensation allows even better phase noise compensation, but all these methods do not compensate ‘out-of-band’ noise which implies the noise in the frequencies larger than the SCS and causes ICI, as it can be seen in Figure 3.

So, phase noise compensation algorithms that operate on a per-symbol basis have limited bandwidth and thus, cannot fully compensate the noise impact. However, with the increasing of the SCS value, simple CPE compensation becomes more and more efficient since phase noise quickly degrades with frequency (approximately 30 dB per decade).

For the baseline subcarrier spacing values, as well as for newly proposed, additional intra-symbol processing is required for the compensation of the higher frequencies phase noise components that will cause ICI.

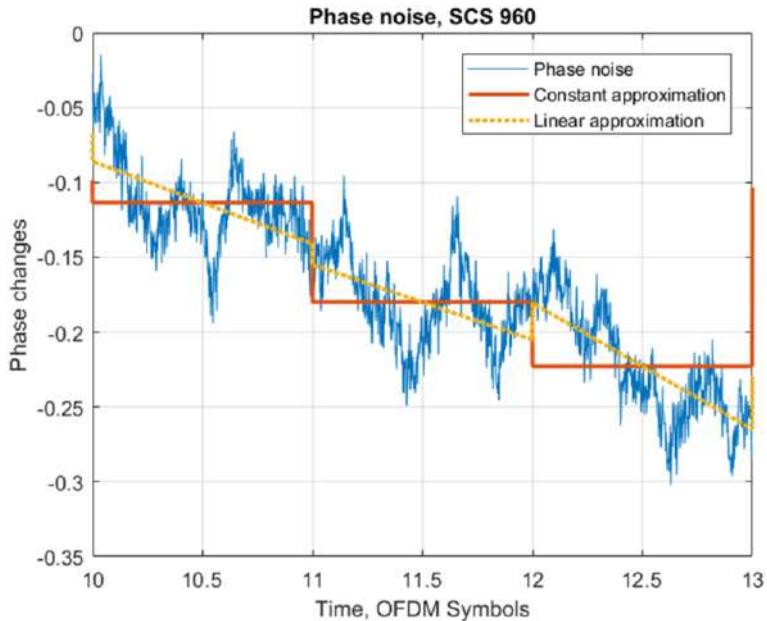


Figure 2 Piecewise constant linear approximations of the phase noise for SCS 960 kHz

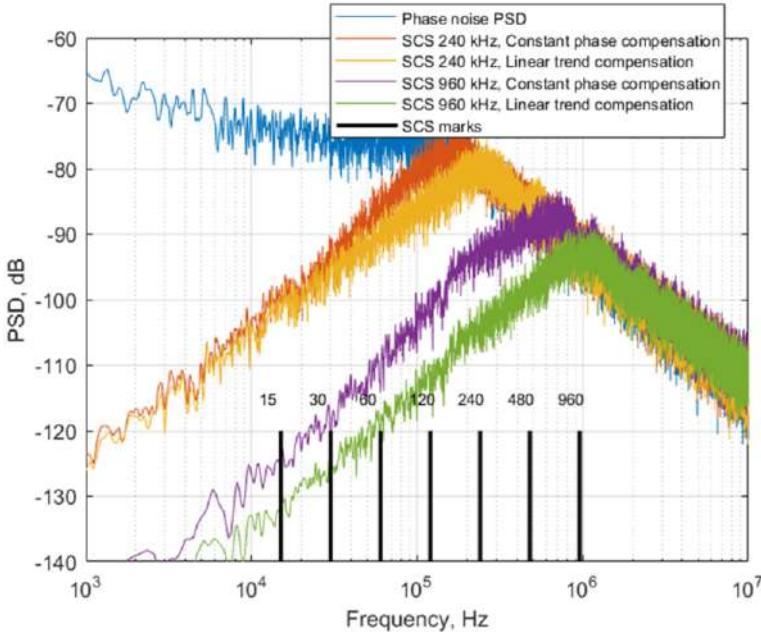


Figure 3 Power spectral density comparison of different trend compensation techniques

4. Inter-carrier interference compensation schemes

For the full compensation of the phase noise impact, the exact knowledge of the frequency domain phase noise realization at the OFDM symbol, given by Eq.(2) is required. However, as it can be seen from PSD, frequency domain representation \bar{J}_i has the most power in the center and several neighboring subcarriers, whilst the noise influence outside those several subcarriers are much smaller and well behind the additive phase noise. Thus, it is convenient to represent the Eq. (1) as a convolution of the received signal with the ‘significant’ part of the phase noise, occupying u positions on the both sides of the considered k -th subcarrier:

$$Y_k = \sum_{i=-u}^u \bar{J}_i H_{k-i} S_{k-i} + \eta_k^{AWGN} + \eta_k^{ICI} \quad (3)$$

or, substituting the product of transmitted signal and channel at k -th subcarriers as X_k and equivalent AWGN and residual ICI noise as W_k :

$$Y_k = \sum_{i=-u}^u \bar{J}_i X_{k-i} + W_k \quad (4)$$

In this representation, the significant part of the phase noise is represented as $2u+1$ tap filter. Since phase noise realization in frequency domain has unity amplitude, the convolution in Eq. (4) is reversible and can be done by filtering with the complex conjugate vector \bar{J}_i .

However, the direct solution of the Eq. (4) regarding \bar{J}_i requires the contiguous pilot structures [12], whilst in the 5G NR specifications the uniformly sparse grid is considered. To avoid the limitations of the signal structure, instead of estimation of the phase noise realization \bar{J}_i using Eq.(4), with the necessity to know adjacent values of X_k we can look for the filter that directly transforms the received signal Y_k , corrupted by the phase noise into the known ideal reference signal values. In this case problem can be formulated as a convolution of the received signal, which is known both at the pilot and data locations with the estimated filter:

$$X_k = \sum_{i=-u}^u J_i Y_{k-i} + Q_k \quad (5)$$

Solving the Eq. (5) should provide filter with $2u+1$ taps that can restore the original signal X_k from the received signal Y_k corrupted by the phase noise. Obviously, increasing the number of filter taps may increase compensation efficiency, since less and less noise appears in the uncompensated part Q_k . However, with limited observation points, increasing the number of estimated parameters may quickly decrease estimation accuracy, while gains from adding more taps are very small.

Convolution (5) can be rewritten in the matrix form using special matrix structure for the set of L observed subcarriers:

$$\begin{bmatrix} X_{k_0} \\ \vdots \\ X_{k_{L-1}} \end{bmatrix} = \begin{bmatrix} Y_{k_0+u} & \dots & Y_{k_0-u} \\ \dots & \dots & \dots \\ Y_{k_{L-1}+u} & \dots & Y_{k_{L-1}-u} \end{bmatrix} \begin{bmatrix} J_{-u} \\ \vdots \\ J_u \end{bmatrix} + \begin{bmatrix} Q_{k_0} \\ \vdots \\ Q_{k_{L-1}} \end{bmatrix} \quad (6)$$

or:

$$\mathbf{X} = \mathbf{YJ}_u + \mathbf{W} \quad (7)$$

Problem (7) has well-known in the literature (for example, [14]) solutions that can be selected on the base of prior knowledge.

The simplest solution, least squares (LS) minimized the squared error of filter tap estimate without any knowledge of the phase noise and AWGN parameters:

$$\hat{\mathbf{J}}_u^{LS} = (\mathbf{Y}^H \mathbf{Y})^{-1} \mathbf{Y}^H \mathbf{X} \quad (8)$$

Minimum mean square error (MMSE) solution:

$$\hat{\mathbf{J}}_u^{MMSE} = \mathbf{R}_{JJ} \mathbf{Y}^H (\mathbf{Y} \mathbf{R}_{JJ} \mathbf{Y}^H + \mathbf{R}_{WW})^{-1} \mathbf{X}, \quad (9)$$

where $(\cdot)^H$ denote Hermitian transpose, \mathbf{R}_{JJ} is the frequency domain phase noise correlation matrix, and \mathbf{R}_{WW} is equivalent noise (residual ICI and AWGN) correlation matrix. Calculation of these values is described in the Appendix A.

MMSE requires prior knowledge of the correlations and noise statistics, has larger complexity in terms of matrix inversions operations, but provides more reliable solution, equally suitable for low and high SNR cases.

Both the LS and MMSE estimates can be further improved by exploitation of the symmetry property of the phase noise frequency domain realization and also de-ICI filter. For this, taps of the estimated filter should be averaged taking into account their symmetry around the center tap.

5. PTRS with subcarrier nulling allocation

It should be noted that both aforementioned algorithms of the phase noise compensation using PTRS are agnostic to the resource allocations – i.e. pilots can be allocated with any pattern along the signal bandwidth. The most efficient way is allocating pilot resources uniformly, to cope with channel frequency selectivity and ensure reliable phase tracking.

However, in this case PTRS values are affected by the ICI from adjacent data subcarriers, induced by the phase noise and other sources. Due to randomness of the data, the full cancellation of the ICI is possible only for decision-directed schemes, with successive decoding and feedback. The computational complexity of such schemes is prohibitive, especially for the millimeter-wave transmissions which typically have very large bandwidth.

Instead, it is possible to allocate PTRSs in a small groups of 3, 5 or generally $(2n+1)$ subcarriers, with only the center subcarrier active and modulated, and the edge subcarriers nulled. The power from the nulled subcarriers can be effectively reallocated to the center one, increasing the SNR and avoiding the ICI at the same time.

Figure 4 shows the basic principle of distributed PTRS allocation, currently used in 5G NR specification (in Figure 4(a)) and the newly proposed PTRS with nulling structure (in Figure 4(b)). It should be noted that typically the distance between PTRSs is no less than 24 subcarriers, but for illustrative purposes, the distance is decreased.

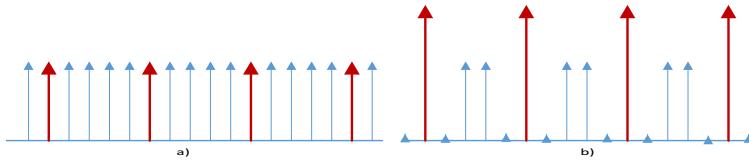


Figure 4 Conventional distributed PTRS allocation (a) vs. PTRS allocation with nulling (b)

It should be noted, that despite nulling of some PTRS subcarriers, they are still counted as reference signals and should be taken into account in the calculations. On the one hand, those subcarriers are included in the calculation of the $(\mathbf{Y}^H \mathbf{Y})$ matrix that is subject to the inversion. Nulled elements on off-diagonal positions make this matrix better conditioned, which is very important in the case of small data allocations, when number of available PTRSs is small. On the other hand, nulled subcarriers significantly decrease the number of multiplications on the final step of Eq. (8) – multiplication by the reference signal \mathbf{X}_k .

6. Computation complexity analysis

Computation complexity comparison of the considered algorithms may be done by calculating the number of complex multiplications (cmult) required for each algorithm, assuming that $M \times N$ by $N \times M$ matrix multiplication requires $M^2 N$ scalar operations. It is also assumed that matrix inversion requires the number of multiplications roughly equal to N^3 . Computational complexity for each approach is calculated assuming that

matrix operations are performed in optimal order, in accordance with corresponding Eq. (8) and (9). Additionally, the number of operations required for application of the filter to the received data is calculated as product of number of data subcarriers by filter size and the shown in a separate column, but also included in the total.

Table 1 shows the number of the complex multiplication required to apply corresponding algorithm, last two columns contain percentage comparison with respect to the base LS scheme. It can be seen that for the large allocations sizes (larger RB number) MMSE becomes very computationally burdened, while for the smaller allocations and larger filter sizes it can be comparable to the LS. The nulling scheme demonstrates a small improvement, since the main part of the complex operations for the LS case involves the application of the filter to the data sequence, and estimation of the filter taps requires much less computations.

Table 1 Computational complexity summary

# RB	PTRS K	# Taps	# RE	De-ICI filter, CMULT	LS, total, CMULT	LS+Null, total, CMULT	MMSE, total, CMULT	LS+Null, %	MMSE %
8	2	3	96	288	376	365	504	-3%	34%
8	4	3	96	288	350	345	374	-1%	7%
8	2	7	96	672	1292	1271	1484	-2%	15%
8	4	7	96	672	1178	1167	1218	-1%	3%
64	2	3	768	2304	2756	2671	39620	-3%	1338%
64	4	3	768	2304	2548	2505	7668	-2%	201%
64	2	7	768	5376	7592	7421	48552	-2%	540%
64	4	7	768	5376	6680	6595	12824	-1%	92%

7. Simulation assumptions and results

The performances of the considered PTRS allocations were investigated for the SCS of 120, 480 and 960 kHz in the 400 MHz bandwidth. Two PTRS configurations were considered: the ‘dense’ case with PTRS allocated every 2nd resource block (K=2) and ‘sparse’ case with PTRS allocated every 4th resource block (K=4). For example, for SCS 480 kHz, for the dense case 32 PTRS resource elements were allocated out of 384 total REs for every OFDM symbol, both for baseline and nulling allocations.

To keep the comparison between two different PTRS allocations fair, the amount of resource elements allocated for PTRS (including nulled REs) was the same for every compared scheme, so the amount of active and nulled subcarriers for the case of nulling was equal to the number of PTRS REs in the baseline 5G NR case. Thus, the number of active PTRS REs is decreased for the case of nulling scheme. The power is also equalized, by boosting the center active pilots.

An OFDM signal with a TX-side phase noise was propagated through a frequency selective channel described by TDL-A model [15], with normalized delay spread of 10ns and Jakes Doppler spectrum corresponding to 3 km/h. RX-side phase noise is applied after the channel and the thermal AWGN noise is added further. With two receive antennas, 2x2 MIMO system but with a rank-1 transmission was modeled, with the optimal transmit beamforming and receiver MRC processing. Transmit beamforming was performed per precoding block of 24 contiguous subcarriers, assuming SVD decomposition of known channel. At the receiver, per precoding block channel estimation is performed, with the LS fitting to the adaptively selected number of channel impulse response taps.

To access block error rate (BLER) metric, the transmitted signal is encoded / decoded with the LDPC algorithm in accordance with the 5G NR specification [16]. Modulation-coding scheme 28, corresponding to the 64-QAM modulation with coding rate 0.9258 (see [3]) is used. For convenience, simulation parameters and assumptions were summarized in Table 2.

Table 2 Simulation parameters and assumptions

Parameters		Assumption
Carrier frequency		60GHz
Bandwidth		400 MHz
SCS		480 kHz / 960 kHz
Data allocation (PDSCH)		64 / 32 RBs per slot
PTRS configuration:	K=2, L=1	32 / 16 REs per symbol
Number of REs	K=4, L=1	16 / 8 REs per symbol
Channel model/Pathloss		TDL-A, 10 ns DS, 3 km/h
Transmission scheme		2x2 MIMO
Channel estimation		LS DFT fitting
Modulation and coding		64-QAM, LDPC R=948/1024

Besides the considered practical CPE and LS / MMSE ICI compensation algorithms, two additional reference cases are also evaluated. The first reference case is the straight forward ideal situation without phase noise – the maximum achievable performance in the case of full phase noise cancellation. The second reference case shows the maximum achievable noise cancellation for a given ICI cancellation filter size. It can be called Clairvoyant algorithm for n-taps filter size. Degradation of this reference curve in comparison with ideal no noise case shows the amount of phase noise not covered by de-ICI filter. This gap can be improved only by increasing the filter size.

On the contrary, difference between Clairvoyant and practical LS / MMSE algorithms for a given filter size is caused only by the taps estimation errors and can be reduced by increasing the estimation accuracy – with application of advanced algorithms and increasing the PTRS density.

MCS 28 is selected as an edge performance mode with high spectral efficiency that requires relatively high SNR values to function. With this MCS, ICI compensation in addition to simple CPE processing is required to achieve target 1% BLER. With lower MCS values, CPE and phase noise compensation schemes may have only a quantitative

difference with a couple of dB gap, while for selected MCS the difference may be qualitative.

Figure 5 and Figure 6 shows the BLER performance under described assumptions for the different SCS values.

The left graph shows BLER curves for the case of 480 kHz SCS, while the right corresponds to 960 kHz SCS, both for baseline value of PTRS density parameter equal to 4, unless curve legend states another. There are five different algorithms/cases depicted: the black line stands for the ideal case of the phase-noise free system, while the next cyan set of curves correspond to the next degree of idealization: the Clairvoyant algorithm with ideal estimation of the fixed size ICI compensation filter.

The blue and magenta curves show the ICI compensation with filters estimated with the LS and MMSE algorithms, respectively, while red lines with the worst performance illustrate the simplest CPE compensation approach.

Inside this categorization, the curves with different additional parameters like number of filter taps or density parameter values can be distinguished by markers and dashed/dotted line style.

It can be seen that with nulled PTRS allocation, simple LS algorithm reaches the same performance as more complex MMSE algorithm.

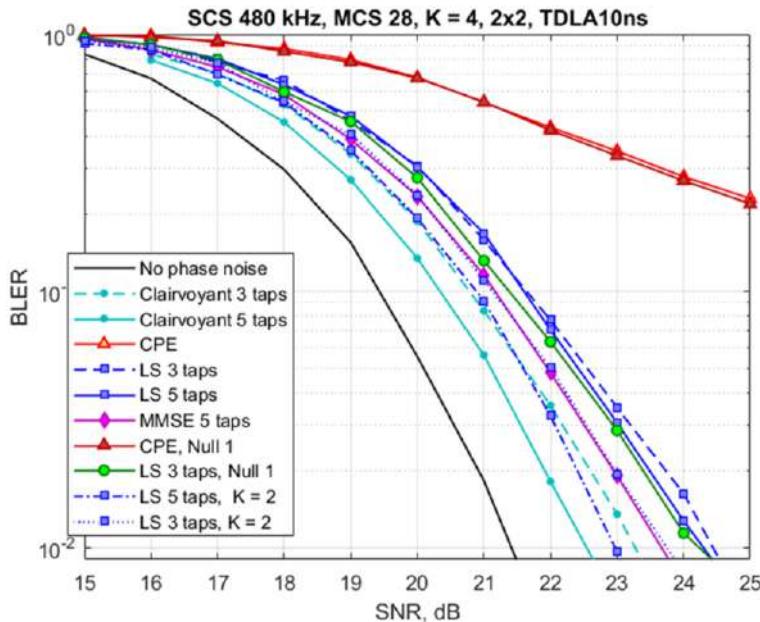


Figure 5 BLER vs. SNR performance comparison for SCS 480 kHz

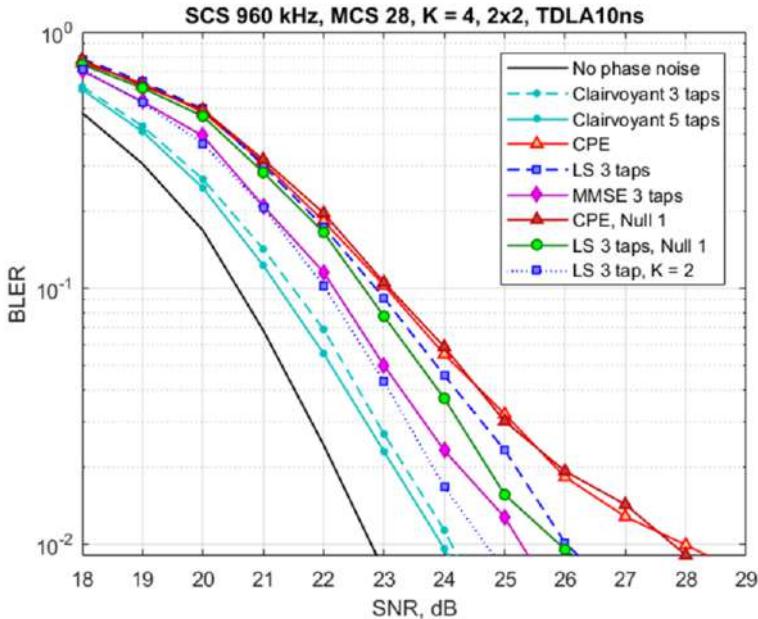


Figure 6 BLER vs. SNR performance comparison for SCS 960 kHz

Comparing the Clairvoyant mode with the LS algorithm for different SCS and K-parameters, it can be observed that there are two edge cases possible.

With relatively large number of PTRSs available (smaller SCS, smaller PTRS spacing K = 2), the filter coefficient estimation accuracy is high, and the practical curves is very close to Clairvoyant ones for given number of taps, and the performance is limited with filter size. Compensation can be improved by increasing filter size; however, this will lead to a decrease in the coefficient estimation accuracy. Therefore, some balance is required between these parameters for achieving an optimal technical solution.

This leads to the opposite case when number of PTRSs is not enough for reliable estimation of the filter coefficients. This case is realized with larger SCS with fixed BW, larger PTRS spacing (K=4) and for the case of large number of filter taps to be estimated. In this case, we can even have the performance degradation in comparison with smaller filter size cases. So, for given number of available PTRSs there is some optimal filter size.

Near this optimal point, where filter estimation accuracy matters, we can improve the BLER performance by exploiting some advanced PTRS allocation schemes, such as nulling or estimation approaches (e.g., MMSE). It can be seen that MMSE can improve the performance in case of small number of PTRSs, but inefficient in case of higher density. At the same time, amount of large size matrices inversion which is required for MMSE calculation, makes it prohibitively complex for practical implementation.

We can see that the difference between the best practical schemes and ideal Clairvoyant mode for given filter size may be as small as 0.5 dB. At the same time, increasing the filter size provides diminishing gains with each increase, with simultaneous coefficients estimation accuracy dropped.

Proposed Nulling PTRS allocation scheme improves the filter taps estimation accuracy by partially avoiding the noise-induced ICI at pilots allocations and provides BLER improvement, about 0.2 dB. It should be noted that BLER curves of the practical algorithms are pretty close to ones of the ideal Clairvoyant algorithm, so further improvement is not expected.

8. Conclusion

In this paper, we have investigated performance of various practical phase noise compensation algorithms operating on the baseline 5G NR PTRS structure and newly proposed PTRS with nulling allocation. It was shown that simple CPE compensation may not be enough for higher order MCS support in the presence of the phase noise, and ICI compensation algorithms are needed to make these modes feasible.

The performance of ICI compensation via Wiener filtering strongly depends on the filter size. It was shown that with dense PTRS placement, the LS and more computationally complex MMSE algorithm have similar performance, very close to the theoretical limit for given filter size.

In the practical case of limited number of available PTRSs, the key role in the compensation performance plays the filter coefficients estimation accuracy that can be increased either by using MMSE, or by exploiting proposed Nulling scheme that allows partial ICI avoidance. Although Nulling scheme provides minor gain about 0.2 dB, it does not require any changes in the LS algorithm and provides a comparably small improvement in computational complexity, due to nulling part of the PTRSs used in calculations.

Generally, the phase noise compensation efficiency depends on the compensation filter size, and with the appropriate number of PTRSs, the increased filter size can be reliably evaluated. However, increasing filter size provides diminishing gains, due to quickly decaying of the phase noise PSD, whilst increasing the PTRS overhead gives stable spectral efficiency drop. Finding the combinations of the filter size and PTRS density, that maximizes system spectral efficiency may be the direction for further studies. Besides, as we demonstrated that the practical filtering ICI compensation techniques are very close in performance to the ideal Clairvoyant algorithm; other approaches like time-domain phase compensation should also be considered.

Appendix A

For the implementation of the MMSE algorithm, knowledge of the phase noise correlation matrix in frequency domain is required. By definition:

$$\mathbf{R}_{JJ}(p, q) = E\{J_p J_q^*\} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} E\{e^{j(\phi_n - \phi_m)}\} e^{-\frac{j2\pi}{N}(pn - qm)}$$

The expression under the sum, the mathematical expectation of the random value exponent $E\{e^{j(\phi_n - \phi_m)}\}$ has the meaning of the characteristic function [13] at parameter $t=1$ of the Gaussian process. Whole expression, in short is a 2D FFT from this characteristic function. So, we can find that:

$$E\{e^{j(\phi_n - \phi_m)}\}|_{t=1} = \psi(1) = e^{-\sigma_{mn}^2/2},$$

where $\psi(t) = \exp\left\{\mu it - \frac{\sigma^2}{2}t^2\right\}$ is Gaussian characteristic function [14].

The expression under the exponent σ_{mn}^2 is a dispersion of the difference between delayed samples of phase noise process that depends on autocorrelation function.

$$\sigma_{mn}^2 = \sigma_m^2 + \sigma_n^2 - 2\rho_{mn}\sigma_m\sigma_n = 2\sigma^2(1 - \rho_{mn}) = 2A(0) - 2A(|m - n|)$$

An autocorrelation function of the phase noise process is a Fourier transform from the PSD, which is a base of the phase noise model and generally should be known for given transmitter and receiver.

Figure A1 shows the structure of the investigated correlation matrix. It can be seen that only the center values can be taken into account while MMSE algorithm is used.

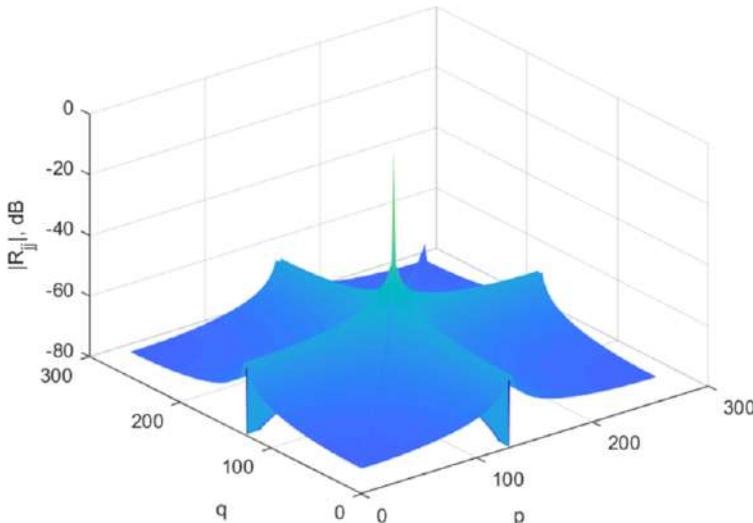


Figure A1 Phase noise frequency domain correlation matrix

References

- [1] RP-202925, “Revised WID: on Extending current NR operation to 71 GHz”
- [2] Maltsev, A., Maslennikov, R., Khoryaev, A. Influence of phase noise on OFDM data transmission systems. Radiophysics and Quantum Electronics, 2011, vol. 53, no. 8, p. 475–487.
- [3] 3GPP TR 38.214 V15.10.0 (2020-06), “Physical layer procedures for data (Release 15)”
- [4] Petrovic, D., Rave, W., and Fettweis, G., “Intercarrier interference due to phase noise in OFDM—Estimation and suppression,” in Proc. IEEE VTC Fall, Sep. 2004, pp. 2191–2195.
- [5] Syrjälä, V., Levanen, T., Ihlainen, T., and Valkama, M., “Pilot Allocation and Computationally Efficient Non-Iterative Estimation of Phase Noise in OFDM,” IEEE Wireless Comm Letters, Apr. 2019.

- [6] Petrovic, D., Rave, W. and Fettweis, G. , "Effects of Phase Noise on OFDM Systems With and Without PLL: Characterization and Compensation," IEEE Trans. Comm., Aug. 2007.
- [7] T. Levanen, O. Tervo, K. Pajukoski, M. Renfors and M. Valkama, "Mobile Communications Beyond 52.6 GHz: Waveforms, Numerology, and Phase Noise Challenge," in IEEE Wireless Communications, vol. 28, no. 1, pp. 128-135, February 2021, doi: 10.1109/MWC.001.2000185.
- [8] Y. Qi, M. Hunukumbure, H. Nam, H. Yoo and S. Amuru, "On the Phase Tracking Reference Signal (PT-RS) Design for 5G New Radio (NR)," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), 2018, pp. 1-5, doi: 10.1109/VTCFall.2018.8690852.
- [9] 3GPP TS38.211 V15.8.0(2019-12) "Physical channels and modulation (Release 15)"
- [10] 3GPP TR 38.803 V14.2.0 (2017-09) "Study on new radio access technology: Radio Frequency (RF) and co-existence aspects"
- [11] Lomayev, A., Kravtsov, V., Genossar, M., Maltsev, A., Khoryaev, A., "Method for Phase Noise Impact Compensation in 60 GHz OFDM Receivers", Radioengineering, vol. 29, no. 1, April 2020
- [12] R1-2005922 "On Phase Noise Compensation for OFDM"
- [13] Proakis, John G., and Dimitris G. Manolakis. Digital Signal Processing. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1996.
- [14] Kay, S. M. *Fundamentals of Statistical Signal Processing*. Vol. 1. Englewood Cliffs, NJ: Prentice-Hall, 1998. ISBN: 0-13-345711-7
- [15] 3GPP TR 38.901 V16.1.0 (2019-12), "Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)"
- [16] 3GPP TR 38.212 V15.9.0 (2020-06), "Multiplexing and channel coding (Release 15)"

Investigating Automated Hyper-Parameter Optimization for a Generalized Path Loss Model

Usman Sammani SANI^{a,c}, Daphne Teck Ching LAI^{a,b,1} and Owais Ahmed MALIK^{a,b}

^a*Digital Science, Faculty of Science, Universiti Brunei Darussalam*

^b*Institute of Applied Data Analytics, Universiti Brunei Darussalam*

^c*Department of Electrical Engineering, Bayero University, Kano, Nigeria*

Abstract. This work aims at developing a generalized and optimized path loss model that considers rural, suburban, urban, and urban high rise environments over different frequencies, for use in the Heterogenous Ultra Dense Networks in 5G. Five different machine learning algorithms were tested on four combined datasets, with a sum of 12369 samples in which their hyper-parameters were automatically optimized using Bayesian optimization, HyperBand and Asynchronous Successive Halving (ASHA). For the Bayesian optimization, three surrogate models (the Gaussian Process, Tree Structured Parzen Estimator and Random Forest) were considered. To the best of our knowledge, few works have been found on automatic hyper-parameter optimization for path loss prediction and none of the works used the aforementioned optimization techniques. Differentiation among the various environments was achieved by the assignment of the clutter height values based on International Telecommunication Union Recommendation (ITU-R) P.452-16. We also included the elevation of the transmitting antenna position as a feature so as to capture its effect on path loss. The best machine learning model observed is K Nearest Neighbor (KNN), achieving mean Coefficient of Determination (R^2), average Mean Absolute Error (MAE) and mean Root Mean Squared Error (RMSE) values of 0.7713, 4.8860dB, and 6.8944dB, respectively, obtained from 100 different samplings of train set and test set. Results show that machine learning can also be used to develop path loss models that are valid for a certain range of distances, frequencies, antenna heights, and environment types. HyperBand produced hyper-parameter configurations with the highest accuracy in most of the algorithms.

Keywords. Path loss, environment, hyper-parameter optimization, feature importance

1. Introduction

Path loss is the decrease in the strength of radio frequency signal strength as it travels from the transmitter to the receiver. For effective design, expansion and monitoring in mobile networks, knowledge of the propagation characteristics of radio signals is required in order to determine the base station transmitting power and antenna height for a given cell radius [1],[2]. Propagation models are used in predicting path loss, which are mathematical expressions used in determining path loss based on frequency, antenna heights, distance etc. Path loss models are classified as either Deterministic models,

¹ Corresponding Author, School of Digital Science, Universiti Brunei Darussalam, Gadong, Brunei Darussalam; E-mail: daphne.lai@ubd.edu.bn

Empirical models, or Stochastic models. Empirical models are the most widely used but their accuracy may diminish if used in an environment different from the one they were developed [3]. Machine learning models are also used to predict path loss and the accuracy of such models outperforms that of empirical models. Such models have parameters that are set before training and their performances depend on them. These parameters are called hyper-parameters and a user might try to use the default setting of a software package or try to optimize their values. There are manual [4] and automated [5] ways to optimize hyper-parameters. The present study focuses on automated optimization of the hyper parameters of a generalized path loss model for multiple frequencies and environments.

Based on analysis of existing machine learning models, it was observed that the following optimization approaches were applied: no hyper-parameter optimization [6], manual tuning [7] and automated hyper-parameter tuning [8]. In terms of the environment type, existing models were developed for urban [9], suburban [10], or rural [11] environments individually, while others in combination [12]. The problem is that if we choose to work at specific frequencies and environment, then for a wide change in frequency another model is required for predicting path loss. As such several models must be developed for the unique environments and frequencies. This can limit the incorporation of machine models in Radio Frequency Planning tools.

Automated hyper-parameter optimization is carried out to increase accuracy, reduce time to tune hyper-parameters manually and to have consistent repeatability in observed results. The traditional way of hyper-parameter tuning is either through Grid search or Random search. Grid search has dimensionality curse for large search space and Random search has limited accuracy [13]. This work is aimed at developing an automatically optimized generalized model that can predict path loss for different environments and frequencies. The contributions of this work are:

- a. Development of generalized path loss models that consider various frequencies and environments, with a clutter height feature used in differentiating the environments.
- b. Use of state-of-the-art hyper-parameter optimization algorithms to automatically optimize the hyper-parameters of machine learning models used in the prediction of pathloss.

2. Methodology

This section describes the dataset used and its preparation, optimization algorithms used and the implementation process.

2.1. Data preparation

Dataset used in this work is a combination of four datasets from different environments, comprising of rural, suburban, urban, and urban high rise terrains, making a total of 18,720 data points. The datasets are made up of path loss values and eleven features namely, latitude, longitude, elevation, distance, frequency, transmitting antenna height (ht), receiving antenna height (hr), difference in latitude between transmitting and receiving antenna (distance_x), difference in longitude between transmitting and receiving antenna (distance_y), elevation at transmitting antenna height position (tAntennaElev) and clutter height. Three of the datasets used are public datasets available

at [14], [15], and [16], and were originally used in [17], [18] and [19], respectively. The fourth dataset is a set of measurements carried out by the authors. The properties of the datasets are given in Table 1. Out of the 9480 data points for the urban area, 3,129 samples were extracted representing a stratified sample of 33% of the total so as to balance the samples per environment. A total 12,369 samples were used. Figure 1 shows the distribution of each feature by environment.

Table 1. Dataset properties

Dataset	Frequency (MHz)	Data points	Description
1 [17]	1835.2, 1840.8, 1864 and 1836	9308	Urban area
2 [18]	868	5624	3349 samples for urban high rise and 2275 for rural area
3 [19]	1800	3616	Sub urban area
4	2140	172	Urban area

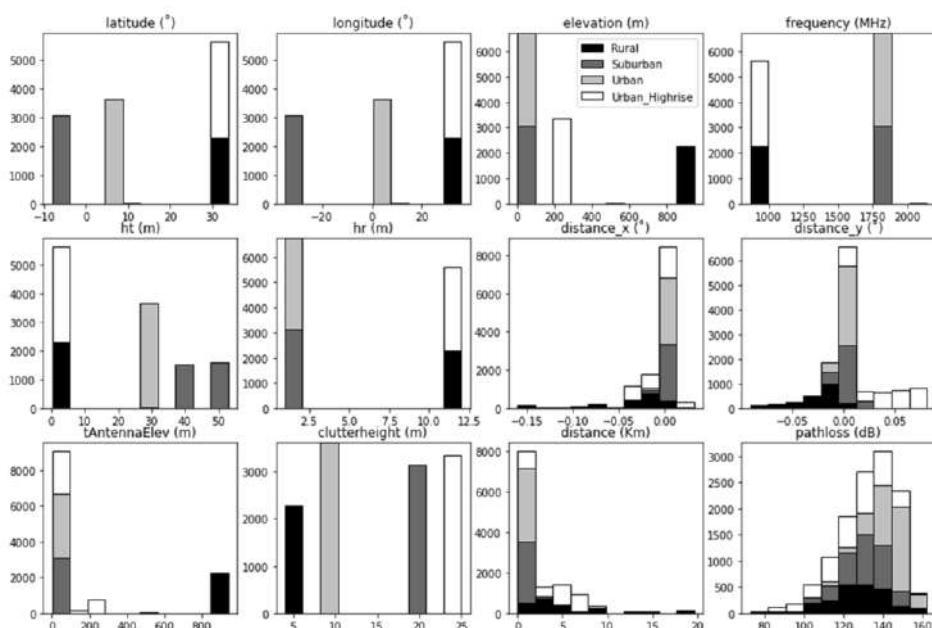


Figure 1. Distribution plots of features with respect to terrains

2.2. Optimization Algorithms

Optimization algorithms used are Bayesian optimization, Hyperband and Asynchronous Successive Halving (ASHA). Bayesian optimization is a sequential model-based optimization technique that takes prior information, use the present sample and then produce posterior information based on a criterion set by a utility function, U known as the acquisition function. It is used in solving functions in which their computation of extrema is expensive, evaluation of their derivative is hard, or they are non-convex. Hyperband and ASHA are early stopping approaches that allocate more resources to promising hyper-parameter configurations than unpromising ones in order to avoid waste. Resources refer to time or number of iterations [13].

2.3. Implementation

The performances of five machine learning algorithms are investigated, namely, KNN, RF, a single hidden layer MLP with Adam optimizer as weights updater, Gradient Boosting (GB) and Extreme Gradient Boosting (XGB) under their best hyper-parameter settings. Six different methods were used for the determination of the hyper-parameters which included Random search, Bayesian optimization using three different types of surrogates (GP, RF and TPE), HyperBand and ASHA. Eighty percent of the data points were used as train set in optimizing the hyper-parameter and twenty percent were used to test the performance. The features in both the train and test sets were scaled to have a mean of zero and standard deviation of one. The hyper-parameter values were evaluated using a 10-fold cross validation with the best score and time taken by each of the algorithms recorded. Each of the optimization algorithms was set for 50 trials/evaluations from which the best trial was selected. Five python packages were used to implement the automated hyper-parameter optimized machine learning process: XGBoost, Scikit-learn for the remaining machine learning models and random search, Scikit-Optimize for Bayesian search using GP and RF surrogate models, Hyperopt for Bayesian optimization using TPE surrogate and Optuna for HyperBand and ASHA.

The best hyper-parameter configuration from each search method was later used in the respective algorithm for performance evaluation. The performances of 100 different samplings of the train set and test set were averaged in each case and the mean MAE, mean RMSE, and mean R^2 were recorded. To examine the improvement in performance provided by the optimization processes, each of the algorithms was also run using the default hyper-parameter settings in the Scikit-learn package or XGBoost.

3. Results and Discussion

The following subsections present the results from various aspects of the study and the discussions about the results obtained. For result analysis, Bayesian optimization with Gaussian process is abbreviated BGP, BRF as an abbreviation for Bayesian optimization with RF surrogate model and BTPE for Bayesian optimization with TPE surrogate model.

3.1. Hyper-parameter Optimization

The performance metrics from the evaluation of the hyper-parameters resulting from the search algorithms, and the time taken by each of the search algorithms to complete 50 evaluations are presented in Table 2. Experiments were conducted with an Intel® Core™ i7-8700 CPU @ 3.20GHz × 12, with 15.6 GB of Random-Access Memory (RAM) and Linux Ubuntu release 20.04 as the operating system. Variations were observed across the five machine learning algorithms based on the hyper-parameter optimization algorithm that spent the highest and least amount of time to complete 50 iterations of the search as well as the algorithm that produced the highest score/least error as presented in Table 3. From Table 2, reduction in RMSE of 0.3693dB, 1.6318dB, 0.5681dB, and 0.922dB in KNN, MLP, RF, and XGB, respectively was achieved by the best optimization algorithm with respect to the default setting.

Table 2. Results of Hyper-parameter optimization

Algorithm		Time (mins)	Mean MAE (dB)	Mean RMSE (dB)	Mean R ²
KNN	Default	N/A	5.0357	7.2637	0.7460
	Random	4.1552	5.2967	7.5349	0.7270
	BGP	5.7399	5.0954	7.3435	0.7406
	BRF	4.7528	5.1499	7.3850	0.7374
	BTPE	2.8291	5.1739	7.3995	0.7364
	Hyperband	6.7423	4.8860	6.8944	0.7713
	ASHA	4.0985	4.8995	6.9203	0.7694
MLP	Default	N/A	7.0158	9.0367	0.6150
	Random	4.1380	6.9376	8.9774	0.6206
	BGP	7.1437	6.7411	8.7105	0.6435
	BRF	31.9863	5.5831	7.4049	0.7415
	BTPE	21.9405	5.9917	7.8796	0.7080
	Hyperband	12.3198	6.4114	8.3240	0.6740
	ASHA	9.8036	6.1387	8.0014	0.6987
RF	Default	N/A	6.8148	9.0916	0.6049
	Random	6.6318	7.0879	9.2457	0.5913
	BGP	9.1431	6.7877	8.9476	0.6205
	BRF	14.8185	7.0726	9.2129	0.5938
	BTPE	7.0694	6.9779	9.0628	0.6074
	Hyperband	5.4783	6.4035	8.5235	0.6573
	ASHA	3.4771	6.6433	8.7038	0.6416
GB	Default	N/A	7.0895	9.2150	0.5966
	Random	37.2973	7.2513	9.5577	0.5633
	BGP	20.2965	7.2310	9.5234	0.5691
	BRF	32.9644	7.2808	9.6443	0.5582
	BTPE	43.2859	7.1539	9.4587	0.5727
	Hyperband	26.7169	7.2611	9.4831	0.5626
	ASHA	2.6171	8.4016	11.0389	0.3834
XGB	Default	N/A	7.8449	10.4265	0.4842
	Random	3.3760	7.9297	10.5898	0.4688
	BGP	7.3779	7.9464	10.6292	0.4654
	BRF	7.7370	7.7580	10.3710	0.4911
	BTPE	4.3214	7.8279	10.4070	0.4862
	Hyperband	2.5661	7.2991	9.5045	0.5725
	ASHA	3.0020	9.5639	13.1130	0.1695

Table 3. Summary of performance

Algorithm	Highest Accuracy	Highest Speed	Lowest Speed
KNN	Hyperband	BTPE	Hyperband
MLP	BRF	Random	BRF
RF	Hyperband	ASHA	BRF
GB	Default	ASHA	BTPE
XGB	Hyperband	Hyperband	BRF

KNN is the best algorithm for this problem as it yielded the least values of the error metrics and a higher R² value. This is due to its stability, especially for large number of 14 neighbors as presented in Table 4. Other algorithms consisting of MLP and Tree based algorithms are less stable, and therefore overfit. HyperBand returned hyper-parameter values with the highest accuracy in most of the machine learning algorithms. Tables 4 to 8 present the default hyper-parameter values in the packages and values obtained from the various optimization methods.

Table 4. Hyper-parameter values of KNN

Parameter	Default	Random	BGP	BRF	BTPE	Hyperband	ASHA
Number of neighbors	5	4	4	4	4	14	13
Algorithm	auto	Ball tree	Ball tree	Kd tree	Kd tree	Ball tree	Ball tree
Leaf size	30	21	24	34	28	19	40
Minkowski Distance metric power	2	8	2	4	5	6	6

Table 5. Hyper-parameter values of RF

Parameter	Default	Random	BGP	BRF	BTPE	Hyperband	ASHA
Number of trees	100	214	162	292	192	260	173
Maximum depth	None	39	74	90	49	81	65
Maximum features per split	Auto	6	9	6	4	10	9
Minimum samples split	2	6	2	18	18	58	73
Minimum samples leaf	1	6	7	12	2	18	20

Table 6. Hyper-parameter values of MLP

Parameter	Default	Random	BGP	BRF	BTPE	Hyperband	ASHA
Number of neurons	100	108	21	115	96	32	102
Activation	relu	relu	tanh	logistic	tanh	logistic	logistic
Alpha	0.0001	0.0010	0.0009	0.0031	0.3391	0.1000	0.0080
Epsilon	1e-8	0.0010	0.2744	0.0059	0.8714	0.8518	0.5340
Beta_1	0.9	0.1000	0.4420	0.5526	0.8559	0.6329	0.7819
Beta_2	0.99	0.9900	0.0225	0.7798	0.8223	0.4865	0.5990

Table 7. Hyper-parameter values of GB

Parameter	Default	Random	BGP	BRF	BTPE	Hyperband	ASHA
Number of trees	100	176	63	188	153	40	215
Maximum depth	3	55	87	75	69	45	24
Maximum features per split	None	5	6	8	7	4	8
Minimum samples split	2	47	45	64	60	40	56
Minimum samples leaf	1	9	9	4	20	26	95
Loss	LS	LS	LS	LS	LS	Huber	LAD

Table 8. Hyper-parameter values of XGB

Parameter	Default	Random	BGP	BRF	BTPE	Hyperband	ASHA
Number of trees	100	68	253	246	281	184	119
Learning rate	0.3	1	0.2742	0.2755	0.1893	0.2449	0.0476
Booster	gbtree	gbtree	gbtree	dart	gbtree	gbtree	gbtree

3.2. Feature Importance

Feature importance from the best KNN model was also considered using the permutation importance method as depicted in Figure 2. It was observed that all features are relevant in the path loss prediction as none was scored zero in feature importance. A novel feature in this model is the “tAntennaElev” that stands for elevation at the location of transmitting antenna. The reason for using it is that since a multiple transmitter model is considered, the altitude of the transmitting antenna relative to the receiving antenna becomes variable also because different transmitting antenna positions have different elevation heights above sea level. The model was observed to give the feature a moderate priority greater than that of some of the features used in earlier empirical models such as frequency and the heights of the antennas. The height of the receiving antenna had the least priority.

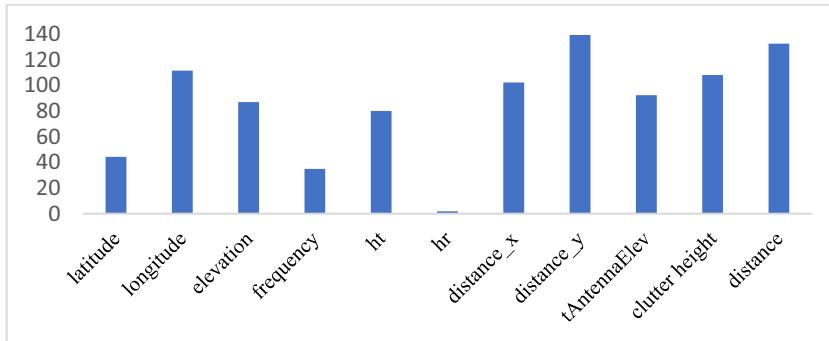


Figure 2. Feature importance of the best KNN model

3.3. Performance by Environment

The best optimized machine learning algorithm obtained is KNN. The R^2 value, MAE and RMSE from each of the environments were computed from the model's predictions on test data as presented in Table 9, with three scenarios. In the first scenario, 100 different samplings of Train-Test split were used in testing the model with unoptimized/default hyper-parameter configuration and results were averaged. In the second scenario, a single Train-test split was used while 100 different samplings were used in the third scenario as in the first scenario. It was observed that results in the first scenario had the least accuracy while the second scenario had the highest accuracy. Meanwhile, the third scenario had an accuracy in between the two other scenarios. The second scenario had the best accuracy because a single Train-Test split was used. This Train-Test split was also used during the optimization of hyper-parameters, while in the third scenario, the hyper-parameters obtained using the Train-Test split in the second scenario were used in checking performance using 100 different Train-Test splits. Thus, results in the second scenario are due to overfitting to the single Training data. Therefore, result from the third scenario should be used in estimating the performance of the obtained hyper-parameters as it reveals how the hyper-parameter setting responds to variations in training and testing data. The prediction plots per environment are shown in Figure 3. Figure 3(a) shows the measured path loss and that predicted by the model in the rural environment. It will be observed that large distances were covered, especially in the rural environment. This is because the dataset used contains measurements carried out by [18] from a Long Range Wide Area Network (LoRaWAN) and the distance covered in rural environment was larger due to little obstacles or structures. For the urban highrise in Figure 3(d), even though the measurements are from a LoRaWAN, the distance covered is smaller than in the rural environment due to the presence of high density of buildings. Figure 3(b) and Figure 3(c) represents the suburban and urban environments, respectively. Low distances were covered in these environments. It was observed that the values predicted by the model in suburban environment have the least fitting to the actual measured data as reflected in its performance metrics in Table 2. In addition to the frequencies in Table 1 in which the model is valid, the range of transmitting antenna height for which it is valid is from 0.2m to 53m and that of receiving antenna is from 1m to 12m as shown in Figure 1.

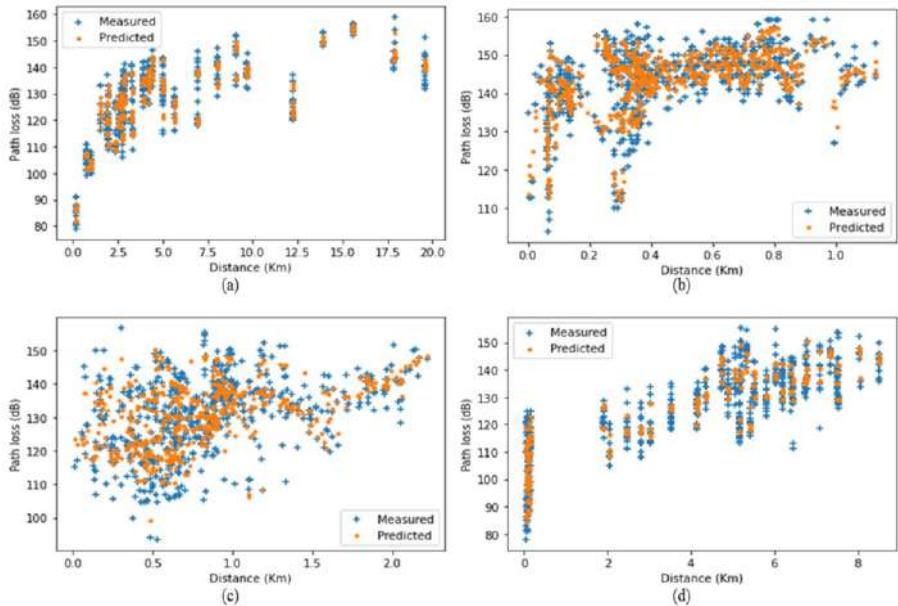


Figure 3. Predicted and measured path loss for (a) rural (b) suburban (c) urban (d) urban highrise environments

Table 9. Optimized and un-optimized performance metrics for the best KNN model

Environment	Un-optimized (100 runs)			Optimized (1 run)			Optimized (100 runs)		
	Mean R ²	Mean MAE	Mean RMSE	R ²	MAE	RMSE	Mean R ²	Mean MAE	Mean RMSE
Rural	0.8952	3.3787	4.9395	0.9491	2.5720	3.4109	0.8875	3.4985	5.1140
Suburban	0.1918	5.8191	8.3359	0.4704	4.5385	6.4173	0.2704	5.5744	7.9179
Urban	0.5518	5.2174	7.3314	0.7492	4.3772	5.8047	0.5846	5.1744	7.0588
Urban High rise	0.7419	5.2329	7.6560	0.8529	4.1689	5.9110	0.7818	5.0031	7.1369

4. Conclusion

We developed a novel path loss model valid for various environments, antenna heights, and frequencies, using the clutter height recommended in ITU-R P.452-16 as feature that differentiates among the environments. The model was obtained by comparing the performances of five machine learning algorithms in which state of the art optimization techniques were used in optimizing their hyper-parameters. To the best of our knowledge, no existing work has applied these techniques in the hyper-parameter optimization of machine learning algorithm for path loss models. We demonstrated that HyperBand optimization produced much improved results. The performance of each of the hyper-parameter optimization algorithms was observed to be dependent on the machine learning algorithm whose hyper-parameters are being optimized, as expected. Hyper-parameters obtained using HyperBand gave the best results in most of the algorithms. The best machine learning algorithm observed is K Nearest Neighbor resulting in an R² value of 0.7713. We demonstrated how machine learning models that consider similar properties adopted by the empirical approach, such as range of distances, antenna heights and frequencies can be developed, but with much improved results. We also

demonstrated that evaluation of hyper-parameters repeatedly with various versions of Train-Test splits from the same dataset reveals the adaptive response of the hyper-parameters to variations in training data. Otherwise, training using only one trainset could result in an overfit. Our method improved accuracy, reduced the time for the optimization of hyper-parameters and the chance of having an overfitted model.

References

- [1] Mahasukhon P, Sharif H, Hempel M, Zhou T, Wang W, Ma T. Propagation path loss estimation using nonlinear multi-regression approach. In: Proceedings IEEE International Conference on Communications. Cape Town, South Africa; 2010.
- [2] Popoola SI, Oseni OF. Empirical Path Loss Models for GSM Network Deployment in. Int Ref J Eng Sci. 2014;3(6):85–94.
- [3] Ogbulezie J, Onuu M. Site specific measurements and propagation models for GSM in three cities in Northern Nigeria. Am J Sci Ind Res. 2013;4(2):238–45.
- [4] Ebhota VC, Isabona J, Srivastava VM. Signal power loss prediction based on artificial neural networks in microcell environment. In: Proceedings 2017 IEEE 3rd International Conference on Electro-Technology for National Development, NIGERCON 2017. Owerri, Nigeria; 2018. p. 250–7.
- [5] Probst P, Boulesteix AL, Bischl B. Tunability: Importance of hyperparameters of machine learning algorithms. J Mach Learn Res. 2019;20(March).
- [6] Benmus TA, Abboud R, Shatter MK. Neural network approach to model the propagation path loss for great Tripoli area at 900, 1800, and 2100 MHz bands. In: Proceedings 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering, STA 2015. Monastir, Tunisia; 2015. p. 793–8.
- [7] Popoola SI, Faruk N, Oloyede AA, Atayero AA, Surajudeen-Bakinde NT, Olawoyin LA. Characterization of Path Loss in the VHF Band using Neural Network Modeling Technique. In: Proceedings 2019 19th International Conference on Computational Science and Its Applications, ICCSA 2019. Saint Petersburg, Russia: IEEE; 2019. p. 166–71.
- [8] Lin KP, Hung KC, Lin JC, Wang CK, Pai PF. Applying least squares support vector regression with genetic algorithms for radio-wave path-loss prediction in Suburban environment. Lect Notes Electr Eng. 2010;67 LNEE:861–8.
- [9] Mom JM, Mgbe CO, Igwue GA. Application of Artificial Neural Network For Path Loss Prediction In Urban Macrocellular Environment. Am J Eng Res. 2014;03(02):270–5.
- [10] Popescu I, Naftornita I, Kanatas A, Moraitsis N. Neural Model for Path Loss Prediction in Suburban Environment. 2001;(July).
- [11] Echie JO, Oyedum OD, Ajewole MO, Aibinu AM. Comparative analysis of basic models and artificial neural network based model for path loss prediction. Prog Electromagn Res M. 2017;61:133–46.
- [12] Ayadi M, Ben Zineb A, Tabbane S. A UHF Path Loss Model Using Learning Machine for Heterogeneous Networks. IEEE Trans Antennas Propag. 2017;65(7):3675–83.
- [13] Yu T, Zhu H. Hyper-Parameter Optimization: A Review of Algorithms and Applications. 2020;1–56. Available from: <http://arxiv.org/abs/2003.05689>
- [14] GitHub - timotrob/SVR_PATHLOSS: The Database relative to paper A Proposal for Path Loss Prediction in Urban Environments using Support Vector Regression [Internet]. [cited 2021 Jan 2]. Available from: https://github.com/timotrob/SVR_PATHLOSS
- [15] LoRaWAN Measurement Campaigns in Lebanon | Zenodo [Internet]. [cited 2021 Jan 2]. Available from: <https://zenodo.org/record/1560654#.X-9B-VUzbIV>
- [16] GitHub - lamvng/Path-loss-prediction: Applying Multilayer Perceptron to predict Path Loss value. [Internet]. [cited 2021 Jan 2]. Available from: <https://github.com/lamvng/Path-loss-prediction>
- [17] Timoteo RDA, Cunha DC, Cavalcanti GDC. A proposal for path loss prediction in urban environments using support vector regression. In: Proceedings of The Tenth Advanced International Conference on Telecommunications. 2014. p. 119–24.
- [18] El Chall R, Lahoud S, El Helou M. LoRaWAN network: Radio propagation models and performance evaluation in various environments in Lebanon. IEEE Internet Things J. 2019;6(2):2366–78.
- [19] Popoola SI, Atayero AA, Arausi OD, Matthews VO. Path loss dataset for modeling radio wave propagation in smart campus environment. Data Br [Internet]. 2018;17:1062–73. Available from: <http://dx.doi.org/10.1016/j.dib.2018.02.026>

Calculation of RMS Values for Variable Frequency Sinusoidal Signals, Using Phasors and Digital SAL and CAL Filters

Yonatan AGUIRRE^a, Flabio GUTIERREZ^{b,1}, Richard ABRAMONTE^b, Nilthon ARCE^c and Antenor ALIAGA^a

^a*Department of Electronics, Universidad Nacional de Piura, Piura, Peru*

^b*Department of Mathematics, Universidad Nacional de Piura, Piura, Peru*

^c*Department of Civil Engineering, Universidad Nacional de Jaen, Cajamarca, Peru*

Abstract. The accurate measurement of RMS values of voltage and current is crucial for the monitoring and protection of power systems and in general for electrical power distribution systems. The authors in this paper have developed an algorithm to calculate the RMS value of sinusoidal signals of varying frequency, by applying phasors obtained from digital filters SAL and CAL. In conditions in which the frequency of the grid varies, a phase shift is presented in the phasor of the grid voltage which is used for the design of a regulator that allows to obtain accurate RMS voltage values. The choice of the SAL and CAL digital filters is due to their low computational requirement, so they can be implemented in general-purpose microcontrollers.

Keywords. RMS, phasors, digital filters, regulators

1. Introduction

The accurate measurement of RMS voltage and current values is crucial for the monitoring and protection of power systems and power distribution systems in general. One of the applications in which it is necessary to measure RMS voltage (V_{RMS}) in a simple and accurate way is smart grids, which are a reliable supply of energy [1]. The applications include, in addition to transmission lines, the load monitoring in smart homes and industries ([2], [3]) and power quality estimation [4].

Among the alternating current (AC) measurement techniques, the digital sampling technique is the most widely used. The input signal is digitized using an analog-to-digital converter (ADC) and the RMS value of the signal is calculated using a digital signal processor (DSP). The achievable accuracy for any given bandwidth varies with the accuracy of the ADC and the sampling rate [5].

The implementation of the algorithm based on the mathematical definition of RMS in general-purpose microcontrollers requires of complex operations that imply a high

¹ Corresponding Author: Flabio Gutierrez, Department of Mathematics, Universidad Nacional de Piura, Campus Universitario, Urb. Miraflores s/n, Castilla, Piura, Perú; E-mail: flabio@unp.edu.pe.

computational cost. Although some microcontrollers already integrate hardware to perform floating-point multiplication and division operations, in many cases other platforms are selected to implement mathematical tools to improve the performance of these operations ([6], [7]).

In real-time systems that require measuring RMS values as in ([8] - [10]), it is necessary to explore other algorithms that allow accurate measurements and have: low complexity, low computational cost and do not require specialized hardware; one of them is described in [11], which presents an algorithm to accurately obtain the V_{RMS} of the electrical grid, using phasors obtained by applying the digital signals processing techniques, specifically, using digital filters SAL and CAL.

The accuracy of the V_{RMS} values obtained by applying these filters requires that the selected sampling frequency is four times the frequency of the power grid. In this context, in environments where the frequency of the network presents variations and maintaining a constant sampling frequency, the above-mentioned requirement is not fulfilled, therefore, the V_{RMS} values obtained are not accurate.

This paper presents a new algorithm that uses the voltage phasor phase shift; because of the variations of the grid frequency, and a regulator to adjust the sampling frequency to four times the grid frequency, to fulfill the condition mentioned in the previous paragraph and to get accurate V_{RMS} values. Once the sampling time is adjusted, the actual grid frequency can be estimated.

Section two defines the SAL and CAL filters and the calculation of the V_{RMS} value. Section three covers the analysis of the accuracy of V_{RMS} values due to mains frequency variation and the relationship between phasor phase variation and line frequency variation. Section four describes the design of the regulator for the sampling frequency adjustment and finally in section five we show the results of the obtained V_{RMS} values as well as the estimated network frequency.

2. Preliminaries

2.1. Definition of RMS value

The mathematical definition of RMS of a signal $x(t)$ is given by equation (1)

$$X_{RMS} = \sqrt{\frac{1}{T} \int_0^T x^2(t) dt} \quad (1)$$

Where T is the period of the signal.

2.2. Phasor of a sinusoidal signal

Given a sinusoidal signal represented by equation (2)

$$v(t) = V_m \sin(\omega * t + \varphi) \quad (2)$$

Their phasor expressed in Cartesian form is given by the equation (3)

$$V = V_m \cos(\varphi) + jV_m \sin(\varphi) \quad (3)$$

And its representation in polar form is given by equation (4)

$$V = V_m \angle \varphi \quad (4)$$

Where:

V_m : maximum value of the sinusoidal signal.

φ : phase of the phasor for $t = 0$

2.3. Calculation of the phasor of $v(t)$ using SAL and CAL filters

Given a sinusoidal signal of amplitude V_m , with line frequency F_L and phase φ , defined by equation (5)

$$v(t) = V_m \sin(2 * \pi * F_L * t + \varphi) \quad (5)$$

Discretizing (5) by $t = k * T_s$, $\forall k \in \mathbb{Z}$, where $T_s = 1/F_s$ is the sampling time. For a sampling frequency F_s defined by equation (6), we obtain equation (7)

$$F_s = 4 * F_L \quad (6)$$

$$v(k) = V_m \sin\left(\frac{\pi}{2} * k + \varphi\right) \quad (7)$$

Using the finite impulse response filters (FIR), whose coefficients are defined by the second and third one-dimensional Walsh functions [12], designed SAL and CAL by Harmuth (1969) and used in [11], we have:

$$SAL = v(k) + v(k-1) - v(k-2) - v(k-3) \quad (8)$$

$$CAL = v(k) - v(k-1) - v(k-2) + v(k-3) \quad (9)$$

by replacing (7) in (8) and (9) we get

$$SAL = -2\sqrt{2} V_m \cos\left(\frac{\pi}{2} * k + \varphi + \frac{\pi}{4}\right) \quad (10)$$

$$CAL = 2\sqrt{2} V_m \sin\left(\frac{\pi}{2} * k + \varphi + \frac{\pi}{4}\right) \quad (11)$$

Making

$$\phi = \frac{\pi}{2} * k + \varphi + \frac{\pi}{4} \quad (12)$$

And knowing that $V_m = V_{RMS}\sqrt{2}$ and (12), (10) and (11) can be written as:

$$SAL = -4V_{RMS} \cos \phi \quad (13)$$

$$CAL = 4V_{RMS} \sin \phi \quad (14)$$

Using (13) and (14) we can express the phasor X by the following equation

$$X = -SAL + j(CAL) \quad (15)$$

and its representation in polar form is given by the equation (16)

$$X = 4V_{RMS} \angle \phi \quad (16)$$

From (16) we obtain the RMS value of the sinusoidal signal $v(t)$ defined in (2), given by equation (17)

$$V_{RMS} = \frac{|X|}{4} \quad (17)$$

3. Statement of the problem

This section analyzes the error of the algorithm to calculate the RMS voltage of the line, using the SAL and CAL filters due to a frequency variation in the grid, given by the following equation

$$v(t) = 220 \sqrt{2} * \sin(2\pi(F_L + \alpha)t) \quad (18)$$

Where $F_L = (60 + \alpha)$ Hz, is the frequency of the mains voltage with a nominal mains frequency equal to 60, a variation α and an RMS amplitude of 220V.

Using (18) and the procedure described above, we find the magnitude and phase variation of phasor X , for a fixed sampling frequency $F_s = 240$ Hz y $\alpha = \pm 1$ Hz, with respect to the number of frames, where a frame is defined by a set of four samples taken sequentially. The simulation results are shown in Fig. 1.a and 1.b respectively.

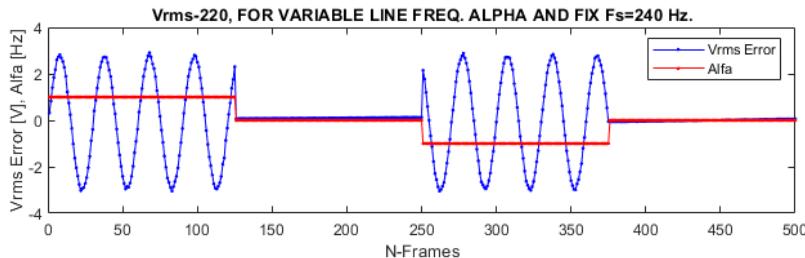


Figure 1.a Amplitude variation of RMS voltage

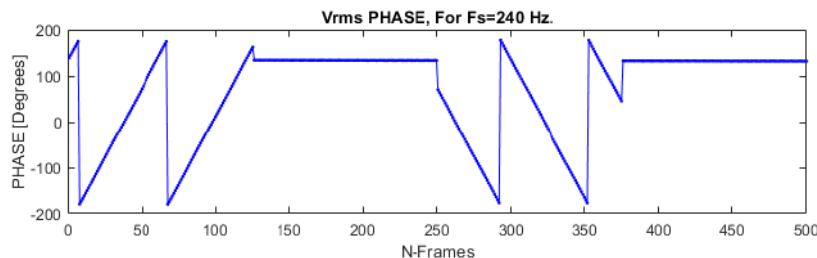


Figure 1.b Phase variation of RMS voltage

From Fig.1.a, we observe that the variation of the magnitude V_{RMS} of X is periodic, with a variation $\Delta v = \pm 3v$. Therefore, the RMS value is given by the following equation

$$V_{RMS} = 220 \pm 3v. \quad (19)$$

From Figure 1.b we observe that the variation of the phase of X is linear with respect to the variation of the frequency α and is defined by the following equation

$$M = \phi_n - \phi_{n-1} \quad (20)$$

The Table 1, shows the variations of V_{RMS} and M for values of $-1 \leq \alpha \leq +1$, which correspond to a range of grid frequency variations $59 \leq F_L \leq 61 \text{ Hz}$.

Table 1. Variations of V_{RMS} and M for values of $-1 \leq \alpha \leq +1$ and $F_S = 240 \text{ Hz}$

α	$F_L[\text{Hz}]$	Min Error [V_{RMS}]	Max Error [V_{RMS}]	$M [\text{deg/frame}]$
1.0	61.0	2.9728	2.7844	6.0
0.5	60.5	1.4627	1.4156	3.0
0.0	60.0	0.0	0.0	0.0
-0.5	59.5	1.4433	1.4142	-2.99
-1.0	59.0	2.9674	2.7791	-5.99

From the first and fourth columns of Table 1, a linear relationship between the phase variation M and the frequency variation α can be derived and is given by the following equation:

$$M = 6 * \alpha \quad (21)$$

Therefore, the objective of the regulator described in the next section is to achieve a slope $M=0$, thus satisfying (6).

4. Regulator design

This section presents the design of the regulator to obtain a phase variation of X equal to zero, which allows the RMS value to be calculated more accurately.

A Timer module (TMR1), implemented in all PIC microprocessors, is used to generate the sampling frequency, which generates an overflow interrupt every $T_S = 1/F_S$. The block diagram of this module is shown in Figure 2.0.

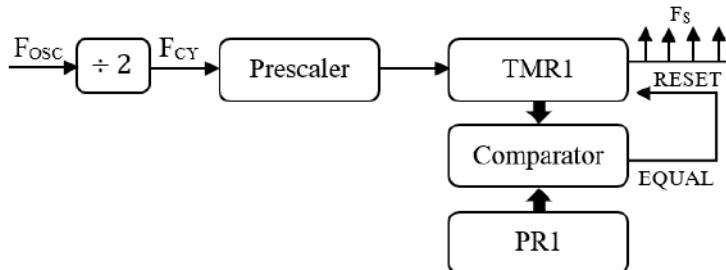


Figure 2. Simplified block diagram of the timer for sample time generation.

From Figure 2.0 the selected oscillator frequency is $F_{OSC} = 20\text{MHz}$ and a prescaler value of 1. Table 2, shows the PR1 register values for each line frequency with $PR1_0 = 41666$ being the value corresponding to a line frequency of 60 Hz and a sampling frequency of $F_s = 240\text{Hz}$.

Table 2. Values of register PR1 for $F_s = 4*F_L$

a	$F_L[\text{Hz}]$	$F_s[\text{Hz}]$	PR1	ΔR
1.0	61.0	244	40983	-683
0.0	60.0	240	41666	0
-1.0	59.0	236	42373	707

From Table 2, we observe that when there is a positive change of a it corresponds to a decrease in the value of PR1 of $\Delta R = -683$. Likewise, for a negative change of a corresponds to an increase of $\Delta R = +707$. These are the minimum and maximum values of the variations that we must use to regulate the sampling frequency.

The implementation of the regulator is shown in Fig.3.0, which "follows" the variations of F_L , to generate the correct sampling frequency F_s , which is used to change the sampling time of the SAL and CAL filters.

To achieve this objective the regulator; using a PI controller, generates the values of ΔR to modify the value of the PR1 register of the timer and thus obtain a new sampling frequency F_s that complies with the relation expressed in (6), taking the value of M to zero [13]. In this way F_s follows the variations of F_L , and using the SAL and CAL filters, we obtain accurate values of V_{RMS} .

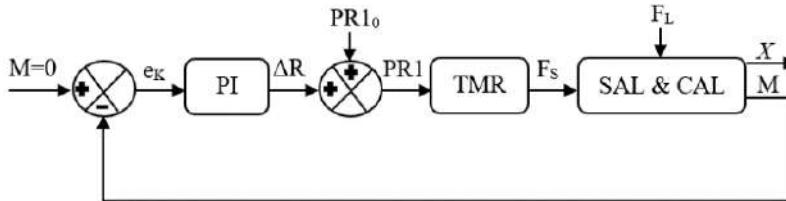


Figure 3. Block diagram of the regulator

The discrete PI controller is given by the equation (22)

$$C(z) = \frac{u(z)}{e(z)} = K * \frac{z}{z-1} \quad (22)$$

The difference equation for (22) is given by the following equation

$$u_k = u_{k-1} + K * e_k \quad (23)$$

5. Results

Simulations have been performed on a personal computer equipped with Intel (R) Core (TM) i5-8265U CPU @ 1.60 GHz 1.80 GHz processor with 8.00 GB RAM with

MATLAB scientific software version 2018th, for a value of $K = 50$ and line frequency variations $\alpha = \pm 1$, which corresponds to frequency variations between 61 and 59 Hz respectively. The results of this simulation are shown in Table 3 and Figure 4.0.

Table 3. Results of regulator implementation

α	$F_L[\text{Hz}]$	$V_{\text{RMS}}[\text{V}]$	ΔR	$F_{\text{EST}}[\text{Hz}]$
1.0	61.0	220.063	683.22	61.001
0.0	60.0	220.00	0.77	60.00
-1.0	59.0	220.063	706.73	59.002

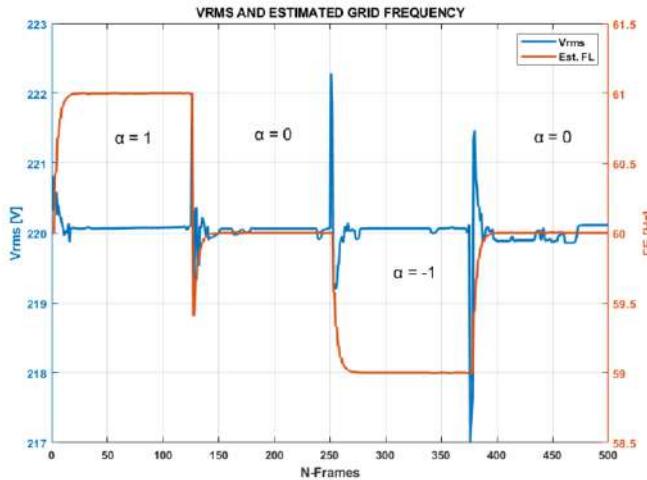


Fig.4. Estimated RMS voltage and line frequency

From the Table 3, we can obtain the percentage error obtained by the algorithm, for the RMS value and the frequency of the estimated grid, is given by

$$\text{error}_{\text{RMS}} = 0.03\%$$

$$\text{error}_{\text{frec}} = 0.003\%$$

6. Conclusions

The SAL and CAL filters used in this work are easy to implement since only additions and subtractions are required, i.e., they do not require a high computational load, therefore, they can be implemented in microcontrollers of any spectrum and low cost.

The SAL and CAL filters can be used to calculate the phasors of a sinusoidal signal and with them calculate the RMS value with high accuracy, when the frequency of the signal does not present variations. The algorithm presents errors in the calculation of the RMS value, in situations in which the frequency of the signal varies.

The addition of a regulator, to the algorithm that calculates the RMS value of the signal, allows accurate values to be obtained even when the signal frequency varies.

Since the period of the line is proportional to the values of the PR1 register of the timer, this tracking algorithm also allows us to estimate the frequency of the line.

References

- [1] Kang S, Yang S, Kim H. Non-intrusive voltage measurement of ac power lines for smart grid system based on electric field energy harvesting. *IET, Electronics Letters*. 2017. 53(3): 181-183.
- [2] Huang TD, Wang WS, Lian KJ. A New Power Signature for Nonintrusive Appliance Load Monitoring. *IEEE Trans. Smart Grid*. 2015. 6(4):1994-1995.
- [3] Yu L, Li H, Feng X, Duan J. Nonintrusive appliance load monitoring for smart homes: recent advances and future issues. *IEEE Instrum. Meas. Magazine*. 2016. 19(3):56-62.
- [4] Ruuskanen V., Koponen J, Kosonen A, Hehemann M, Keller R, Niemelä M, Ahola J. Power quality estimation of water electrolyzers based on current and voltage measurements. *Journal of Power Sources*, (2020). 450: 227603.
- [5] Mujumdar UB, Joshi JS. Microcontroller based true RMS current measurement under harmonic conditions. 2010. *IEEE International Conference on Sustainable Energy Technologies (ICSET)*: 1-5.
- [6] Eski A, Zavalani O, Komici D. RMS Value Measurement Using a Meter Based on ARM Cortex-M0 Microprocessor. *European Journal of Electrical Engineering and Computer Science*. 2017. 1(2).
- [7] Gómez EJ, Plazas DC, Restrepo M.F. Voltímetro true-rms sobre fpga basado en algoritmo cordic. *Tecnura: Tecnología y Cultura Afirmando el Conocimiento*. 2015. 19(2): 129-136.
- [8] M. M. Zaid, M. Nauman, M. Rahim, and M. U. Malik, "DSP BASED ENERGY MONITORING WITH ONLINE DAQ SYSTEM.," *Acta Tech. Corvininesis - Bull. Eng.*, vol. 11, no. 2, pp. 59–64, 2018.
- [9] J. Barros and E. Pérez, "Automatic detection and analysis of voltage events in power systems," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 5, pp. 1487–1493, Oct. 2006, doi: 10.1109/TIM.2006.881584.
- [10] F. Fortes and V. Fialho, "Smart energy meter for iot applications," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 11, pp. 2708–2712, Sep. 2019, doi: 10.35940/ijitee.K2168.0981119.
- [11] Schweitzer EO, Aliaga A. Digital programmable time-parameter relay offers versatility and accuracy. *IEEE Transactions on Power Apparatus and Systems*. 1980. 1: 152-157
- [12] Hazra, L., & Mukherjee, P. (2018). *Walsh Functions*. In *SpringerBriefs in Applied Sciences and Technology* (pp. 1–15). Springer Verlag. https://doi.org/10.1007/978-981-10-2809-0_1
- [13] Åström KJ, Hägglund T. The future of PID control. *Control engineering practice*. 2001. 9(11):1163-1175.

Optical Vortices Sharp Focusing by Silicon Ring Gratings Using High-Performance Computer Systems

Dmitry SAVELYEV^{a,b,1}

^a*Samara National Research University, Samara, Russia*

^b*IPSI RAS – Branch of the F SRC “Crystallography and Photonics” RAS, Samara, Russia*

Abstract. The diffraction of vortex laser beams with circular polarization (with different direction of polarization rotation) by silicon ring gratings was investigated in this paper. The silicon diffractive axicons with different numerical apertures (NA) were considered as such ring gratings. The considered diffractive axicons are compared with single silicon circular protrusion (cylinder). The finite difference time domain method was used for Light propagation (3D) through the proposed silicon ring gratings and silicon cylinder. The possibility of subwavelength focusing by varying the height of the elements is demonstrated. In particular, it is numerically shown that a silicon cylinder forms a light spot with the minimum size (intensity) of the longitudinal component of the electric field FWHM is 0.32λ .

Keywords. Silicon photonics, subwavelength ring gratings, optical vortices, FDTD

1. Introduction

The silicon photonics has attracted more attention from researchers in the past few decades [1-5]. This is due to a number of reasons, including the cheaper implementation of optical devices using standard semiconductor manufacturing technologies [2] and miniaturization of solutions for light transmission through optical fibers [3]. It is also possible to obtain a high integration density due to the large refractive index of silicon [1]. Photonics in general and silicon photonics in particular are becoming a promising alternative for the implementation of new accelerators for deep learning that can use light, both for communication and for computing, thanks to such advantages as high speed, wide bandwidth and parallelism [6,7].

A number of classical structures in the construction of crossing silicon waveguides are well known, such as subwavelength grating, shaped taper, holey subwavelength grating [3]. In particular, the methodology for designing silicon based waveguide Bragg gratings (WBGs) in detail discusses the paper [8]. And one of the ways to increase the number of channels for data transmission is the introduction of higher order modes in mode-division-multiplexed (MDM) systems, which also allows for

¹ Corresponding Author: Dmitry Savelyev, Department of Technical Cybernetics, Samara National Research University, 443086 Samara, Russia; E-mail: dmitrey.savelyev@yandex.ru.

increased device design flexibility [9]. One of the promising methods is channel thickening at modes with different orders of the vortex phase singularity [10,11].

It should be noted that for efficient light beams insertion into a small diameter optical fiber can be used sharp focusing [12]. The vortex modes are scale invariant at the exit from the fiber (using diffraction microstructures) and a decrease in the radiation divergence at the exit of an optical fiber can be obtained by structuring the exit end of a fiber with a ring structure [13]. Such a ring grating (for example, a diffractive axicon) with a shorter period (larger numerical aperture of the axicon) makes it possible to achieve sharper focusing than an aplanatic lens with the same numerical aperture of the system [14,15]. The diameter of the central spot at full width at half maximum intensity (FWHM) of such an element is 0.36 wavelength divided by the numerical aperture [16]. The possibility of manufacturing microaxicons expands the area of their application [15, 17] and allows them to be used to form structured light fields.

The introduction of vortex phase singularity into the incident beam makes it possible to enhance the longitudinal component of homogeneous-polarized laser beams on the optical axis in the focal region [18]. So, it which makes it possible to change the diffraction pattern due to the redistribution of energy between the components of the electromagnetic field [19,20]. This was shown earlier for focusing elements with a refractive index $n = 1.46$ [19, 20], but an increase in the contribution of the longitudinal component to the overall intensity pattern on the optical axis can be achieved [18] with an increasing of the refractive index.

The diffraction of vortex Gaussian beams by silicon ring gratings, representing diffraction axicons with different numerical apertures study in this paper. A separate protrusion in the form of a circular cylinder, close in size to the size of the central part of the diffractive axicons, was also considered. The effect of changing the height of the considered optical elements on subwavelength focusing is investigated. Numerical calculations of laser propagation (3D) were performed using high performance computer systems (computing cluster with a capacity of 800 Gflops) with the finite difference time domain (FDTD) method [21,22].

2. Investigation of the Laguerre-Gauss Modes (1,0) Diffraction by Silicon Cylinder

Modeling parameters: the size of the computational domain x, y, z in the range $[-5.8\lambda; 5.8\lambda]$, radiation wavelength $\lambda = 1.55 \mu\text{m}$, the spatial sampling step is $\lambda / 30$, the time sampling step is $\lambda / (60c)$, where c is the speed of light, and the thickness of the absorbing layer PML is 1.16λ (on all sides surrounding the computational domain). The source was located inside the substrate, which occupies the entire space below the relief and is partially embedded in the PML layer. The refractive index of the element n is 3.47. The Laguerre-Gauss mode (1,0) (optical vortex) was considered as the input laser radiation:

$$GL_{01}(r, \varphi, z) = \left(\frac{\sqrt{2}r}{\sigma(z)} \right) \exp[ikz - i2\eta(z)] \exp\left[\frac{i\pi r^2}{\lambda R(z)} \right] \exp\left[-\frac{r^2}{\sigma^2(z)} \right] \exp(i\varphi), \quad (1)$$

where $\varphi = \text{arctg}(y/x)$, $\eta(z) = \text{arctg}(z/z_0)$, $r^2 = x^2 + y^2$, $R(z) = z(1 + z_0^2/z^2)$ is the radius of curvature of the light field parabolic front, $z_0 = \pi\sigma_0^2/\lambda$ is the confocal parameter, $\sigma(z)$ is the effective beam radius, and λ is the wavelength.

The height h of the binary element relief (even if it is a separate protrusion), corresponding to the phase jump π radians, for the selected refractive index is as follows:

$$h = \frac{\pi}{k(n-1)} \approx 0.2\lambda, \quad (2)$$

where $k = 2\pi/\lambda$ – wave number, λ – wavelength of laser radiation, n – refractive index.

The FDTD method is currently one of the most common methods for solving Maxwell's equations written in differential form. The Maxwell's equations using the FDTD method are solved by discretizing using central differences in time and space, and then solving them numerically [23]. The Meep package (was developed by a research group of scientists from the Massachusetts University of Technology) was used for numerical simulation with FDTD method [24]. The Meep package uses standard Yee grid sampling.

The results of a study on the passage of the Laguerre-Gauss mode (1,0) through the silicon cylinder with a radius of $r = 2\lambda$ are shown in the Figure 1.

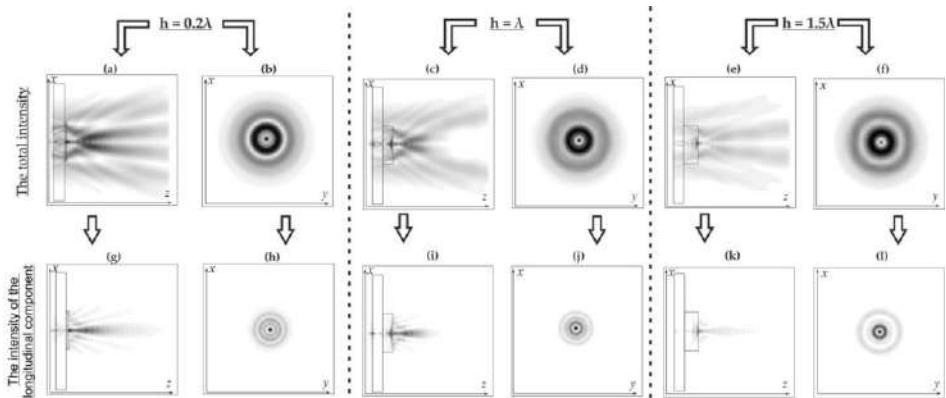


Figure 1. The optical vortex diffraction by circular protrusion ("—" - circular polarization): for total intensity of the electric field at $h = 0.2\lambda$ – (a), (b), at $h = \lambda$ – (c), (d), at $h = 1.5\lambda$ – (e), (f); and the intensity of the electric field longitudinal component at $h = 0.2\lambda$ – (g), (h), at $h = \lambda$ – (i), (j), at $h = 1.5\lambda$ – (k), (l).

The addition of an optical vortex significantly changes the focal pattern and the direction of rotation of circular polarization becomes important [18, 19]. At the second optical vortex order and higher for "—" - circular polarization (the sign of circular polarization is opposite to the sign of the introduced vortex phase singularity), a shadow round light spot was formed. So, the first order of an optical vortex in an incident beam considered in this paper. The polarization, the direction of which coincides with the direction of the vortex phase singularity, will be called "+" - circular polarization. As expected, a zero value (Figure 2) for a central focal spot indicates "+" - circular polarization, and "—" - circular polarization indicates by non-zero value [15].

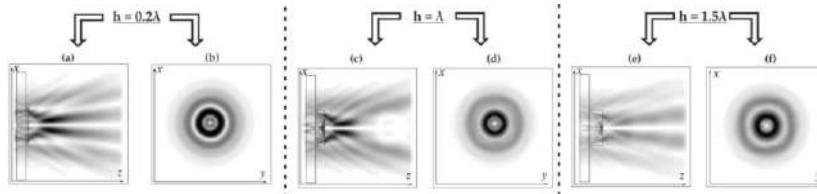


Figure 2. The optical vortex diffraction by circular protrusion ("+" - circular polarization), total intensity of the electric field at $h = 0.2\lambda$ - (a), (b), at $h = \lambda$ - (c), (d), at $h = 1.5\lambda$ - (e), (f).

The focal spot size on the optical axis was estimated at FWHM of the all intensity and FWHM of the intensity of the electric field longitudinal component (FWHMz). The distance from the edge of the microrelief to the section plane is denoted as d . The transverse diffraction patterns (xy plane) considered at a distance of $d = 0.5\lambda$. The cylinder height increasing resulted in better focusing: FWHM = 0.45λ at $h = 0.2\lambda$ (FWHMz = 0.37λ), FWHM = 0.37λ at $h = \lambda$ (FWHMz = 0.35λ) and FWHM = 0.36λ at $h = 1.5\lambda$ (FWHMz = 0.32λ).

3. Investigation of the Laguerre-Gauss Modes (1,0) Diffraction by Axicons

In this section, studies of the influence of a change in the relief height on the diffraction pattern in the near zone are continued. The simulation parameters are the same as in the previous section. The results of a study on the passage input beams through the diffractive axicons with NA = 0.25 (grating period 4λ) and NA = 0.95 (grating period 1.05λ) are shown in the Figure 3 and Figure 4.

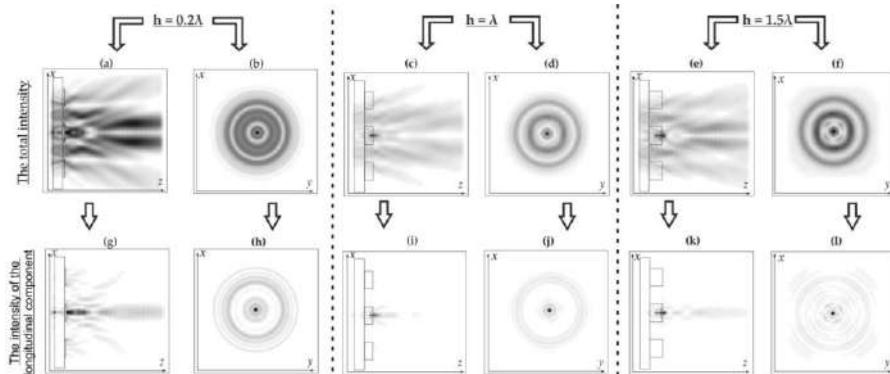


Figure 3. The optical vortex diffraction by an axicon with NA = 0.25 ("+" - circular polarization): for total intensity of the electric field at $h = 0.2\lambda$ - (a), (b), at $h = \lambda$ - (c), (d), at $h = 1.5\lambda$ - (e), (f); and the intensity of the electric field longitudinal component at $h = 0.2\lambda$ - (g), (h), at $h = \lambda$ - (i), (j), at $h = 1.5\lambda$ - (k), (l).

The transverse diffraction patterns (xy plane) considered at a distance of $d = 0.2\lambda$ for diffractive axicon with NA = 0.25 (and for case $h = 1.5\lambda$ for NA = 0.95) and $d = 0.5\lambda$ for diffractive axicon with NA = 0.95 ($h = 0.2\lambda$ and $h = \lambda$). The following results were obtained: for diffractive axicon with NA = 0.25, FWHM = 0.61λ at $h = 0.2\lambda$ (FWHMz = 0.46λ), FWHM = 0.48λ at $h = \lambda$ (FWHMz = 0.43λ) and FWHM = 0.86λ at $h = 1.5\lambda$ (FWHMz = 0.34λ); for diffractive axicon with NA = 0.95, FWHM = 0.48λ at $h = 0.2\lambda$ (FWHMz = 0.44λ), FWHM = 0.47λ at $h = \lambda$ (FWHMz = 0.43λ) and FWHM = 0.67λ at $h = 1.5\lambda$ (FWHMz = 0.36λ). The contribution of the longitudinal component of

the electric field to the total intensity decreases with an increase in the height of the aaxicon relief, despite a decrease in the size of the focal spot (only the longitudinal component). In addition, the size of the central focal spot begins to be influenced by the transverse components of the electric field, leading to its broadening.

It should also be noted that for the axicon with $NA = 0.95$, strong side lobes are formed for the case $h = 1.5\lambda$; however, at a lower height, it gives better focusing compared to a smaller numerical aperture. Also note that for a silicon cylinder with $r = 2\lambda$, a narrower focal spot on optical axis was obtained ($FWHM = 0.37\lambda$ at $h = \lambda$ and $FWHM = 0.36\lambda$ at $h = 1.5\lambda$), and at $h = 0.2\lambda$ the size of the focal spot is comparable with axicons result.

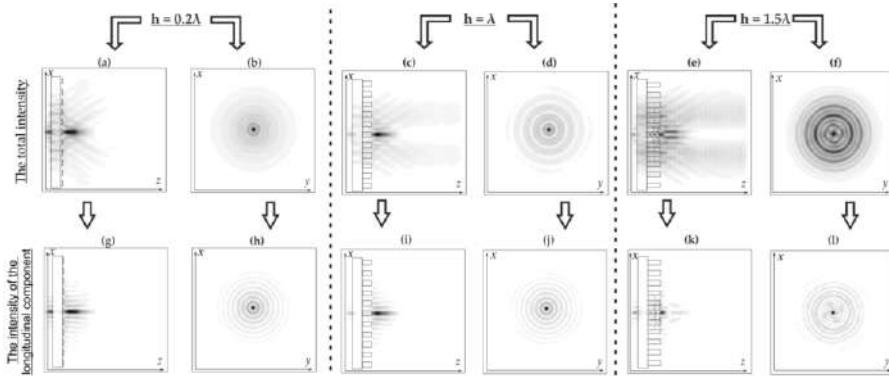


Figure 4. The optical vortex diffraction by an axicon with $NA = 0.95$ (" - circular polarization): for total intensity of the electric field at $h = 0.2\lambda$ – (a), (b), at $h = \lambda$ – (c), (d), at $h = 1.5\lambda$ – (e), (f); and the intensity of the electric field longitudinal component at $h = 0.2\lambda$ – (g), (h), at $h = \lambda$ – (i), (j), at $h = 1.5\lambda$ – (k), (l).

4. Conclusions

The diffraction of vortex laser beams with circular polarization by silicon ring gratings was investigated in this paper. The silicon cylinder focusing is compared with diffractive axicons with different grating periods and shown, that the result of focusing with a circular cylinder with a radius of $r = 2\lambda$ was better than focusing by a diffractive axicon with grating period 4λ ($NA = 0.25$) and comparable to focusing by a diffractive axicon with grating period 1.05λ ($NA = 0.95$) at low relief height.

The enhancement possibility of sub-wavelength focusing is demonstrated when the height of the elements is increasing. It is shown that in this case, the contribution of the electric field longitudinal component to the size of the central focal spot increases up to a certain height. Also shown that a silicon cylinder forms a light spot with the minimum size (intensity) of the electric field longitudinal component $FWHM$ is 0.32λ with $h = 1.5\lambda$.

Acknowledgement

This research was funded by the Russian Science Foundation (project No. 20-72-00051) in the parts «Investigation of the Laguerre-Gauss Modes (1,0) Diffraction by

Silicon Cylinder», «Investigation of the Laguerre-Gauss Modes (1,0) Diffraction by Axiicons»; the Ministry of Science and Higher Education and was performed under the «ERA.Net RUS plus» program and funded by RFBR, project number 20-52-76021» in the parts «Introduction», «Conclusions».

References

- [1] Mu X, Wu S, Cheng L, Fu HY. Edge couplers in silicon photonic integrated circuits: A review. *Applied Sciences* 2020; 10(4): 1538 (29 pp.).
- [2] Cheng L, Mao S, Li Z, Han Y, Fu HY. Grating couplers on silicon photonics: design principles, emerging trends and practical issues. *Micromachines* 2020; 11(7): 666 (25 pp.).
- [3] Wu S, Mu X, Cheng L, Mao S, Fu HY. State-of-the-art and perspectives on silicon waveguide crossings: A review. *Micromachines* 2020; 11(3): 326 (16 pp.).
- [4] Butt M, Khonina SN, Kazanskiy NL. Optical elements based on silicon photonics. *Computer Optics* 2019; 43(6): 1079-1083.
- [5] Leonardi AA, Faro MJL, Irrera A. Silicon Nanowires Synthesis by Metal-Assisted Chemical Etching: A Review. *Nanomaterials* 2021; 11(2): 383 (24 pp.).
- [6] Shastri BJ, et al. Photonics for artificial intelligence and neuromorphic computing. *Nature Photonics* 2021; 15(2): 102-114.
- [7] Xiang S, et al. A review: Photonics devices, architectures, and algorithms for optical neural computing. *Journal of Semiconductors* 2021; 42(2): 023105.
- [8] Kaushal S, et al. Optical signal processing based on silicon photonics waveguide Bragg gratings. *Frontiers of Optoelectronics* 2018; 11(2): 163-188.
- [9] Li C, Liu D, Dai D. Multimode silicon photonics. *Nanophotonics* 2018; 8(2): 227-247.
- [10] Bozinovic, N. et. al. Terabit-Scale Orbital Angular Momentum Mode Division Multiplexing in Fibers. *Science* 2013; 340(6140): 1545-1548.
- [11] Wu, H, et al. All-fiber second-order optical vortex generation based on strong modulated long-period grating in a four-mode fiber. *Optics Letters* 2017; 42(24): 5210-5213.
- [12] Savelyev DA, Khonina SN. Sharp focusing by means of binary relief at the end of the optical fiber. *Optical Technologies for Telecommunications* 2013; 2014; 9156: 915609 (6 pp.).
- [13] Liu Y, Xu H, Stief F, Zhitenev N, Yu M. Far-field superfocusing with an optical fiber based surface plasmonic lens made of nanoscale concentric annular slits. *Opt. Express* 2011; 19(21): 20233-20243.
- [14] Khonina SN, Savelyev DA, Kazanskiy NL. Analysis of polarisation states at sharp focusing. *Optik-International Journal for Light and Electron Optics* 2016; 127(6): 3372-3378.
- [15] Khonina SN, Savelyev DA, Kazanskiy NL. Vortex phase elements as detectors of polarization state. *Optics Express* 2015; 23(14): 17845-17859.
- [16] Kalosha VP, Golub I. Toward the subdiffraction focusing limit of optical superresolution. *Optics Letters* 2007; 32(24): 3540-3542.
- [17] Khonina S, Degtyarev S, Savelyev D, Ustinov A. Focused, evanescent, hollow, and collimated beams formed by microaxicons with different conical angles. *Optics Express* 2017; 25(16): 19052-19064.
- [18] Khonina SN, Savelyev DA. High-aperture binary axicons for the formation of the longitudinal electric field component on the optical axis for linear and circular polarizations of the illuminating beam. *Journal of Experimental and Theoretical Physics* 2013; 117(4): 623-630.
- [19] Savelyev DA, Khonina SN. Characteristics of sharp focusing of vortex Laguerre-Gaussian beams. *Computer Optics* 2015; 39(5): 654-662.
- [20] Khonina SN, et al. Experimental demonstration of the generation of the longitudinal E-field component on the optical axis with high-numerical-aperture binary axicons illuminated by linearly and circularly polarized beams. *Journal of Optics* 2013; 15(8): 085704.
- [21] Khonina S, Savelyev D. Optimization of the Optical Microelements Using High-Performance Computer Systems. *Radiophysics & Quantum Electronics* 2015; 57(8-9): 650-658.
- [22] Savelyev D. The investigation of the cylindrically polarized beams focusing by a diffractive axicon using high-performance computer systems. 2020 International Conference on Information Technology and Nanotechnology (ITNT), IEEE; 2020 May: 1-5.
- [23] Taflove A, Hagness SC. Computational electrodynamics: the finite-difference time-domain method. Norwood, MA: Artech House; 2005. 1038 p.
- [24] Oskooi AF, Roundy D, Ibanescu M, Bermel P, Joannopoulos JD, Johnson SG. MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method. *Computer Physics Communications* 2010; 181(3): 687-702.

An Explainable Convolutional Neural Networks for Automatic Segmentation of the Left Ventricle in Cardiac MRI

Jun LIU^{a,c}, Feng DENG^{b,1}, Geng YUAN^c, Xue LIN^c, Houbing SONG^d and Yanzhi WANG^c

^a *Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA*

^b *Beijing Key Laboratory of High Dynamic Navigation Technology, Beijing Information Science & Technology University, Beijing, China*

^c *Department of Electrical & Computer Engineering, College of Engineering, Northeastern University, Boston, MA, USA*

^d *Security and Optimization for Networked Globe Laboratory (SONG Lab), Embry-Riddle Aeronautical University, Daytona Beach, FL, USA*

Abstract. Recently, the study on model interpretability has become a hot topic in deep learning research area. Especially in the field of medical imaging, the requirements for safety are extremely high; Moreover, it is very important for the model to be able to explain. However, the existing solutions for left ventricular segmentation by convolutional neural networks are black boxes; explainable CNNs remains a challenge; explainable deep learning models has always been a task often overlooked in the entire data science lifecycle by data scientists or deep learning engineers. Because of very limited medical imaging data, most solutions currently use transfer learning methods to transfer the model which used on large-scale benchmark data sets (such as ImageNet) to fine tune medical imaging models. Consequently, a large amount of useless parameters are generated, resulting in further barrier for the model to provide a convincing explanation. This paper presents a novel method to automatically segment the Left Ventricle in Cardiac MRI by explainable convolutional neural networks with optimized size and parameters by our enhanced Deep Learning GPU Training System. It is very suitable for deployment on mobile devices. We simplify deep learning tasks on DIGITS systems, monitoring performance, and displaying the heat map of each layer of the network with advanced visualizations in real time. Our experiment results demonstrated that the proposed method is feasible and efficient.

Keywords. Explainable, image segmentation, left ventricle, interpretable, transfer learning, CNNs, mobile device

1. Introduction

The purpose of medical image segmentation is to segment parts with special significance in medical images, extract relevant features, provide reliable basis for

¹ Corresponding Author, Feng Deng, Beijing Key Laboratory of High Dynamic Navigation Technology, Beijing Information Science & Technology University, Beijing, No.35 North Fourth Ring Road, Chaoyang District, Beijing, China; E-mail: dengfeng@bistu.edu.cn.

clinical diagnosis and pathological research, and help doctors make more accurate diagnosis. Due to the complexity of medical images themselves, a series of problems such as nonuniformity and individual differences need to be solved in the segmentation process, so the general image segmentation methods are difficult to be directly applied to medical image segmentation.

Because medical imaging is ultimately an aid to doctors' clinical diagnosis, it is not enough to tell doctors whether a MRI image is ill or not. At the same time, for the classification and segmentation results given by the network, doctors also want to know why, that is to say, from the doctor's point of view, they want to know how interpretable medical AI is. To meet these needs our model is explainable and understandable for doctor it allows us to verify hypotheses and help us to improve model's performance.

2. Related Work

Although the high nonlinearity endows the multilayer neural network with high model representation ability, it can achieve very gratifying performance on many issues with some parameter adjustment techniques that can be called modern alchemy. Model designer more want to know what knowledge the model has learned from the data (expressed in a way that human beings can understand), which leads to the final decision. Whether it can help designer find some potential correlations, if designer want to develop an application based on deep learning model to help doctor judge patient risks. Besides the final judgment results, designer may also need to know what factors the model is based on. If a model can't be explained at all, its application in many fields will be limited because it can't give more reliable information. In this paper, we focus on solving the interpretability problem of deep learning.

Our major contributions of this paper are the following:

- This paper discusses the modeling stage from explainable methods before modeling,
- An explainable model and explained the model design is provided
- This paper use NVIDIA Deep Learning GPU Training System(Digits) provide visual analysis of typical convolution layer and de-convolution layer.
- This paper provides experiment methods and an evaluation method, it verified the importance of Dice metric as a measure of medical image segmentation.

3. Explainable Model design

3.1 Input Data Set

The data set we utilizing Sunnybrook cardiac images which is a series of cardiac images (specifically MRI short-axis (SAX) scans) that have been expertly labeled. See References [1, 2, 3] for full citation information. The main view for assessing ventricle size is the short- axis stack (PSAX), Contains images taken in a plane perpendicular to the long axis (PLAX) of the left ventricle.

The examples of the MRI data are shown below. This image is an instance of the data. On the left Figure.1a are the MRI images and the right Figure.1b is the expertly-annotated regions (often called contours). The portions of the images that are part of

the LV are denoted in white. Note that the size of LV varies from image to image, but the LV typically takes up a relatively small region of the entire image.

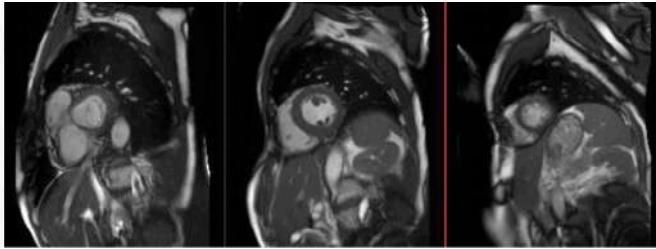


Figure.1a Original LV MRI image

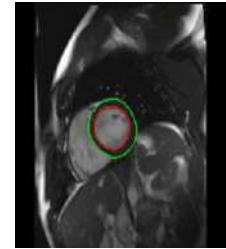


Figure.1b expert segmentation

The images themselves are originally 256 x 256 grayscale DICOM format, a common image format in medical imaging. The label is a tensor of size 256 x 256 x 2. The reason the last dimension is a 2 is that the pixel is in one of two classes so each pixel label has a vector of size 2 associated with it.

3.2 Explainable methods before modeling

Explainable methods before modeling also named Pre-modelling explainability is a collection of diverse methods with a common goal of gaining a better understanding of dataset used for model development [4]. The key of interpretable methods before modeling is to help designer quickly and comprehensively understand the characteristics of data distribution, so as to help they consider the possible problems in the modeling process and choose the most reasonable model to approach the optimal solution that can be achieved by the problem.

Compared with ordinary images, medical images have the characteristics of high complexity, large gray scale range and unclear boundary, and the internal structure of human body is relatively fixed. The distribution of segmentation targets in human body images is regular, and the semantics are not simple and clear(i.e. inter-annotator disagreements, poor segmentation reproducibility) so the network need combine low-resolution information during downsampling (providing the basis for object category recognition) and high-resolution information during upsampling (providing the basis for accurate segmentation and positioning).

Because the data collection of medical images is relatively difficult and the amount of data is small, if there are too many parameters in the model, it will easily lead to over-fitting, so the model with small size and few parameters is suitable.

3.3 Establish an explainable model

In order to capture small regions of interest, convolution layer is used to capture larger receptive fields.

This paper can accomplish this by using an image recognition neural network, and replacing the fully-connected layers (typically the last few layers) with deconvolution layers (arguably more accurately called transpose convolution layers).

Deconvolution [5-8] is an upsampling method that brings a smaller image data set back up to its original size for final pixel classification. It can be helpful to visualize

how the input data (in the case a tensor of size 256 x 256 x 1) "flows" through the graph, i.e., how the data is transformed via the different operations of convolution, pooling and such. The Figure.2 represents the transformations that our data will undergo in the next task.

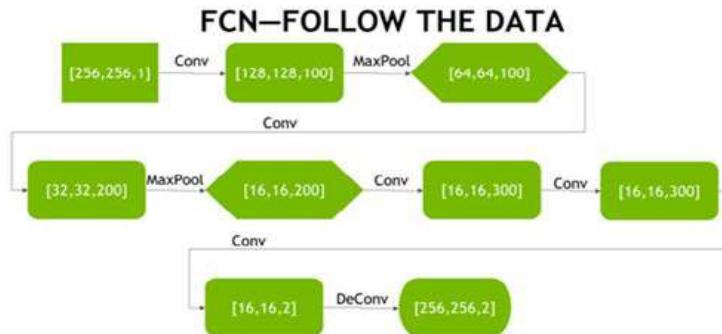


Figure.2 network structure

The network represented by the figure above is consisted convolution layers, pooling layers, and a final deconvolution layer, with the input image being transformed as indicated in the image.

1. Convolution1, 5 x 5 kernel, stride 2
2. Maxpooling1, 2 x 2 window, stride 2
3. Convolution2, 5 x 5 kernel, stride 2
4. Maxpooling2, 2 x 2 window, stride 2
5. Convolution3, 3 x 3 kernel, stride 1
6. Convolution4, 3 x 3 kernel, stride 1
7. Score_classes, 1x1 kernel, stride 1
8. Upscore (deconvolution), 31 x 31 kernel, stride 16

The original image is reduced to 1/4 after conv1 and pool1

The image is reduced to 1/16 after conv2 and pool2 for the second time

Conv3 and conv4 are convolution operation, and the number of feature Maps of the image changes, but the image size is still 1/16 of the original image. At this time, the image is no longer called feature map but heat Map.

The deconvolution layer is used to up-sample the feature map of the last convolution layer, so that it can be restored to the same size as the input image

For the operation of convolution, the input element matrix x and the output element matrix y , this process is described by matrix operation:

$$y = Cx \quad (1)$$

By deduction, we can get sparse matrix: C

The operation of deconvolution is to inverse the matrix operation:

$$x = C^T y \quad (2)$$

According to the matrix differential formula:

$$\frac{\partial Ax + b}{\partial x} = A^T \quad (3)$$

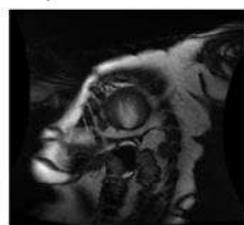
It can be deduced the gradient of the loss of the deconvolution to the input x :

$$\frac{\partial Loss}{\partial x_i} = \sum_i \frac{\partial Loss}{\partial y_i} \frac{\partial y_i}{\partial x_i} = \frac{\partial Loss}{\partial y_i} C_{i,j} = \frac{\partial Loss}{\partial y_i} C_{*,j} = C_{*,j}^T \frac{\partial Loss}{\partial y_i} \quad (4)$$

3.4 Explanatory after modeling

There are three interpretable methods for deep learning: 1. Hidden layer analysis, 2. Sensitivity analysis, and 3. Agent/substitute model. In this section, this paper mainly analysis the first method: hidden layer analysis.

The interpretable method after modeling is mainly aimed at the deep learning model with black box property. By evaluating the accuracy of the model, designer can get the change of the performance of the hidden layer in the whole training process and after the training [9]. This paper use our [Deep Learning GPU Training System](#) to train and analysis the model's performance. Visualize the result of convolution kernel after activation. Designer can see the result of image convolution, which helps designer to understand the function of convolution kernel. Through the heat map, designer can know which parts of the image play a key role in the image classification problem, and at the same time, designer can locate the position of objects in the image



Layer visualizations

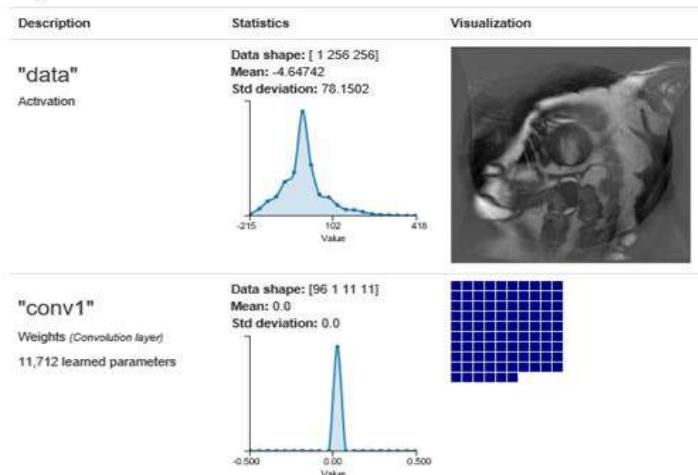


Figure.3 convolution layer1 Visualization

On our model based on MRI dataset training, after 5 cycles of training, In Figure3, the features of the image extracted by convolution layer are coarse, in the subsequent convolution2 layer feature map is obviously improved more detail.

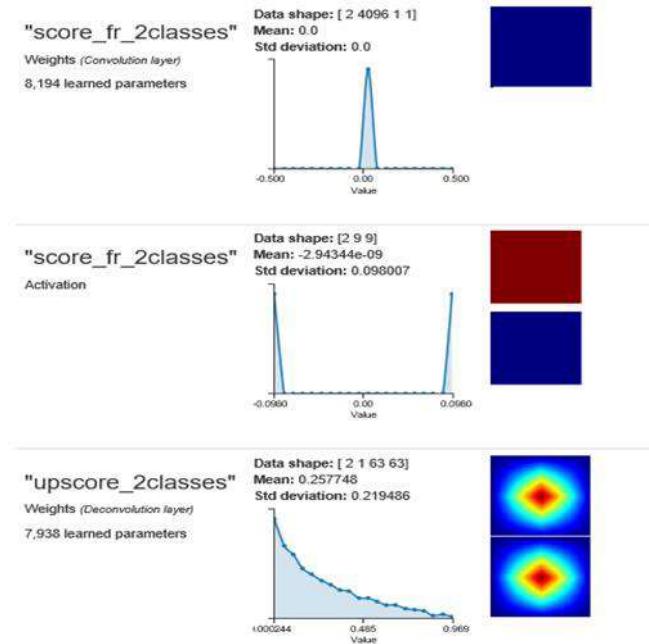


Figure.4 full convolution layer Visualization

In Figure4, full convolution layer has extracted the left ventricle and non-left ventricle parts feature in the picture into two part. In deconvolution layer left ventricle is highlighted in red, indicating that the network is looking at the correct position when making classification judgment. through the heat map, we can know which parts of the image play a key role in the image classification problem, and at the same time, we can locate the position of objects in the image.

4. Experiments with model

4.1 Workflow

- 1.Prepare input data--Input data can be Numpy arrays.
- 2.Build the Computation Graph--Create the graph of the neural network including specialized nodes such as inference, loss and training nodes.
- 3.Train the model--inject input data into the graph in a Session and loop over your input data. Customize your batch size, number of epochs, learning rate, etc.
- 4.Evaluate the model--run inference (using the same graph from training) on previously unseen data and evaluate the accuracy of your model based on a suitable metric.

4.2 Parameter search

At this point a neural network is created that it seems has the right structure to do a reasonably good job and an accuracy metric is used that correctly show how well the network is learning the segmentation task. But up to this point the evaluation accuracy

hasn't been as high as expected. The next thing to consider is try to search the parameter space a bit more. Up to now the number of epochs have been changed and there are a few more parameters can be tested that could push the accuracy score higher. These are:

- learning_rate: the initial learning rate
- decay_rate: the rate that the initial learning rate decays.,
e.g., 1.0 is no decay, 0.5 means cut the decay rate in half each step, etc.
- decay_steps: the number of steps to execute before changing the learning rate

The learning rate is the rate at which the weights are adjusted each time after run back propagation. If the learning rate is too large, it should be end up adjusting the weights by values that are too large and it should be end up oscillating around a correct solution instead of converging. If the learning rate is too small, the adjustments to the weights will be too small and it might take a very long time before we converge to a solution that we expect. One technique often utilized is a variable or adjustable learning rate. At the beginning of training, a larger learning rate will be used so that we make large adjustments to the weights and hopefully get in the neighborhood of a good solution. Then as we continue to train we'll successively decrease the learning rate so that we can begin to zero in on a solution. The three parameters listed above will help you control the learning rate, how much it changes, and how often it changes.

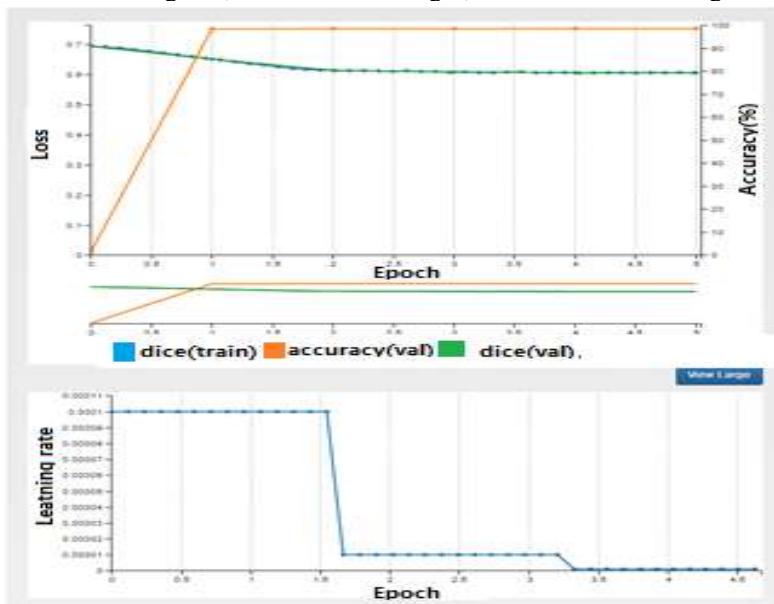


Figure 5 Predict Result high without Dice metric

From Figure 5 when set learning rate=0.0001 even through trained 1 epoch the model reaches an accuracy of over 98%, This means that 98% of pixels are correctly predicted as belonging to either the left ventricle or the background. This is very good!

5. Evaluation method

This paper considers fully what exactly we are computing when we check accuracy. The current accuracy metric is simply telling us how many pixels we are

computing correctly. So in the case above with 5 epochs, model is correctly predicting the value of a pixel 98 % of the time. However, notice from the images above that the region of LV is typically quite small compared to the entire image size. This leads to a problem called class imbalance [10], i.e., one class is much more probable than the other class. In this case, if simply designed a network to output the class not LV for every output pixel, result still have something like 95% accuracy. But that would be a seemingly useless network. This is meant to illustrate that high pixel accuracy doesn't always imply superior segmentation ability [11].

What the paper need is an accuracy metric that gives some indication of how well the network segments the left ventricle irrespective of the imbalance.

One metric can be used to more accurately determine how well the network is segmenting LV is called the Dice metric or Sorensen-Dice coefficient, among other names. This is a metric to compare the similarity of two samples. In this case it is used to compare the two areas of interest, i.e., the area of the expertly-labeled contour and the area of the predicted contour. The formula for computing the Dice metric is:

$$\text{Dice metric} = \frac{2A_{nl}}{A_n + A_l} \quad (5)$$

where A_n is the area of the contour predicted by this neural network, A_l is the area of the contour from the expertly-segmented label and A_{nl} is the intersection of the two, i.e., the area of the contour that is predicted correctly by the network. 1.0 means perfect score. More accurately compute how well predicting the contour against the label, We can just count pixels to the respective areas.

6. Experiments Result

When the training epoch is increased to 30 a significant increase can be seen in accuracy. In fact an accuracy of 98.3% is quite good.

There are a few more parameters can be tested that could push our accuracy score higher. These are:

- learning_rate 0.03
- decay_rate 0.75
- decay_steps 10000
- num_epochs 100

This model gest accuracy of 99.7% and Dice metric above 86.3%

Example of segmentation result

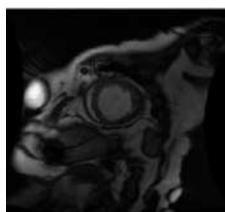


Figure.6a Original LV image

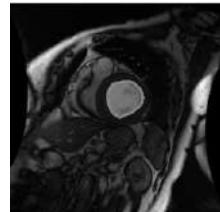


Figure.6b LV Validation

In Figure.6b the contour of the prediction is very smooth and the predicted results are almost consistent with Figure.6a the original left ventricular position

Compare with different explainable method on Sunnybrook datasets

Method	Dice metric(val)
ours	0.863

Xiaofeng Liu et al.(2021)[11]	0.860
Manuel Pérez-Pelegrí et al(2021) [12]	0.790
QiaoZheng et al.(2019[13]	0.760
Alexnet+FCN(base line)	0.622

7. Summary

This paper focused on design explainable convolutional neural network from the interpretability of deep learning methods for automatic Segmentation of the Left Ventricle in Cardiac MRI at a trade-off between accuracy and interpretability. The paper is different with traditional CNNs design, it present explainable methods before modeling, established an explainable model analyzed the hidden layer after modeling, the experimental steps and an evaluation method is provided. After fine tune this model obtained high accuracy and dice metric. This model is simply and lightweight, it is also very suitable to deploy on mobile devices to diagnose medical images.

Acknowledgment

NVIDIA Deep Learning Institute provided technical support and this work was supported in part by National Natural Science Foundation of China (41871348); Beijing Information Science & Technology University (2020KYNH224); and Beijing Key Laboratory of High Dynamic Navigation Technology (HND2019002).

References

- [1] cardiac images from earlier competition http://smial.sri.utoronto.ca/LV_Challenge/Data.html
- [2] "Sunnybrook Cardiac MR Database" is made available under the CC0 1.0 Universal license described above, and with more detail here: <http://creativecommons.org/publicdomain/zero/1.0/> .
- [3] Radau P, Lu Y, Connelly K, Paul G, Dick AJ, Wright GA. "Evaluation Framework for Algorithms Segmenting Short Axis Cardiac MRI." The MIDAS Journal -Cardiac MR Left Ventricle Segmentation Challenge, <http://hdl.handle.net/10380/3070>
- [4] Bahador Khaleghi, The How of Explainable AI: Pre-modelling Explainability, 2019-8. <https://towardsdatascience.com/the-how-of-explainable-ai-pre-modelling-explainability-699150495fe4>
- [5] Fully Convolutional Networks for Semantic Segmentation <http://fcn.berkeleyvision.org/>
- [6] Long, Shelhamer, Darrell; "Fully Convolutional Networks for Semantic Segmentation", CVPR 2015..
- [7] Zeiler, Krishnan, Taylor, Fergus; "Deconvolutional Networks", CVPR 2010.
- [8] Jonathan Long, Evan Shelhamer, Trevor Darrell Fully Convolutional Networks for Semantic Segmentation, CVPR2015
- [9] Guillaume Alain, Yoshua Bengio, Understanding intermediate layers using linear classifier probes, - arXiv preprint arXiv:1610.01644, 2016 - arxiv.org
- [10] Jason Brownlee on December 23, 2019 in Imbalanced Classification, A Gentle Introduction to Imbalanced Classification <https://machinelearningmastery.com/what-is-imbalanced-classification/>
- [11] Xiaofeng Liu, Segmentation of Cardiac Structures via Successive Subspace Learning with Saab Transform from Cine MRI, arXiv:2107.10718,2021
- [12] Manuel Pérez-Pelegrí et al., Automatic left ventricle volume calculation with explainability through a deep learning weak-supervision methodology, Computer Methods and Programs in Biomedicine, Volume 208, September 2021, 106275
- [13] Qiao zheng et al.Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow, Medical Image Analysis Volume 56, August 2019, Pages 80-95

Total Variance Approach on Commercial Vehicle Handling Dynamics

Yaohua GUO and Jinquan DING¹

Yutong Bus Co., Ltd, Zhengzhou, China

Abstract. The suspension roll motion can produce roll steer, which are functions of roll angle, thus producing extra lateral forces. This paper develops the total variance approach to analyze the effect of suspension roll on commercial vehicle handling dynamics. The side-slip angle unified transfer function, reaction time, transition time, damping ratio and total variance are introduced with the effect of suspension. The vehicle designers could use this approach to get optimization design parameters of vehicle without numerical calculation. For the two-axle commercial vehicles, the total variance approach is useful and could serve as an important tool for evaluating the effect of vehicle suspension roll on commercial vehicle dynamics.

Keywords. Suspension roll steer, total variance approach, side-slip angle unified transfer function, vehicle handling dynamics

1. Introduction

The suspension roll steer and camber thrust are import factor for vehicle handling characteristics. The usual evaluation indexes could only represent single performance, it needs to evaluate the comprehensive performance by using one comprehensive vehicle handling characteristics evaluation index.

Schunck and Riekert established the linear two degree freedom model of vehicle dynamics [1]. Segel [2] established the three degree freedom model vehicle dynamics. Abe [3] presented an equivalent approach to analyze the effect of suspension roll on two-axle vehicle. K. Guo [4] analyzed the natural frequency, damping ratio, reaction time, transition time and total variance of the two-axle vehicle. Ding J [5] presented an equivalent approach to analyze both the effect of body roll and n-axle handling on vehicle dynamics by the equivalent wheelbase and steady factor. Y. Zhao [6] studied the total variance approach for the vehicle closed-Loop control System. K. Guo [7] justified the use of total variance analysis approach.

This paper develops the total variance approach to analyze the effect of suspension on commercial vehicle handling dynamics. For the two-axle commercial vehicle, the

¹ Corresponding Author, Yutong Bus Co., Ltd, Zhengzhou 450004, China; E-mail:dingjqxs@163.com

mathematical handling concepts, such as side-slip angle unified transfer function, reaction time, transition time, damping ratio and total variance, are calculated to analyze the effect of suspension roll steer without numerical calculation. The vehicle designers could use this approach to get optimization design parameters of vehicle.

2. Commercial vehicle model inclusive of roll

According to Ding J [5], the equations 3DOF can be converted to

$$MV\dot{\beta} + 2(K_1 + K_2)\beta + [MV + \frac{2(x_1 K_1 + x_2 K_2)}{V}]r - M_s h \dot{p} - 2Y_\phi \phi = 2K_1 \delta \quad (2.1)$$

$$2(x_1 K_1 + x_2 K_2)\beta + I_z \dot{r} + \frac{2(x_1^2 K_1 + x_2^2 K_2)}{V}r - I_{xz} \dot{p} - 2N_\phi \phi = 2x_1 K_1 \delta \quad (2.2)$$

$$-M_s h V \dot{\beta} - I_{xz} \dot{r} - M_s h V r + I_\phi \dot{p} + C_\phi p + (K_\phi - M_s g h) \phi = 0 \quad (2.3)$$

where

$$Y_\phi = \left(\frac{\partial \alpha_1}{\partial \phi} K_1 + \frac{\partial \alpha_2}{\partial \phi} K_2 \right) - \left(K_{c1} \frac{\partial \phi_1}{\partial \phi} + K_{c2} \frac{\partial \phi_2}{\partial \phi} \right)$$

$$N_\phi = \left(\frac{\partial \alpha_1}{\partial \phi} x_1 K_1 + \frac{\partial \alpha_2}{\partial \phi} x_2 K_2 \right) - \left(K_{c1} x_1 \frac{\partial \phi_1}{\partial \phi} + K_{c2} x_2 \frac{\partial \phi_2}{\partial \phi} \right)$$

By assuming the steady-state roll angle as

$$\phi = \left(M_s h V r + M_s h V \dot{\beta} + I_{xz} \dot{r} - I_\phi \dot{p} - C_\phi p \right) / (K_\phi - M_s g h) \approx \frac{M_s h V r}{(K_\phi - M_s g h)} \quad (2.4)$$

The equivalent 2DOF equations of motion with roll effect can be expressed as follows.

$$MV(\dot{\beta} + r) = 2K_1[\delta - \beta - \frac{x'_1}{V}r] + 2K_2[-\beta - \frac{x'_2}{V}r] \quad (2.5)$$

$$I_z \dot{r} = 2K_1 x_1 [\delta - \beta - \frac{x'_1}{V}r] + 2K_2 x_2 [-\beta - \frac{x'_2}{V}r] \quad (2.6)$$

where

$$x'_i = x_i (1 + E_i V^2) \quad (2.7)$$

$$E_i = -\frac{M_s h \left(\frac{\partial \alpha_i}{\partial \phi} - \frac{K_{ci}}{K_i} \frac{\partial \gamma_i}{\partial \phi} \right)}{x_i (K_\phi - M_s g h)} \quad (2.8)$$

The side-slip angle and yaw rate unified transfer function were

$$\frac{\beta}{\delta}(s) = Z_\beta \frac{\tau'_1 s + 1}{Z_2 s^2 + Z_1 s + 1} \quad (2.9)$$

$$\frac{r}{\delta}(s) = Z_r \frac{\tau_1 s + 1}{Z_2 s^2 + Z_1 s + 1} \quad (2.10)$$

where, Z_β is the steady-state side-slip angle gain, Z_r is the steady-state yaw rate gain.

The steady-state side-slip angle gain Z_β for two-axle vehicle is

$$Z_\beta = \frac{\beta}{\delta}(s)_{|s=0} = \frac{\frac{4K_1 K_2 x'_2 (x_2 - x_1)}{V} - 2x_1 K_1 M V}{-2M V (K_1 x_1 + K_2 x_2) + \frac{4K_1 K_2 (x_1 - x_2)(x'_1 - x'_2)}{V}} \quad (2.11)$$

$$\tau'_1 = \frac{2K_1 I_z}{\frac{4K_1 K_2 x'_2 (x_2 - x_1)}{V} - 2x_1 K_1 M V} \quad (2.12)$$

$$Z_1 = \frac{[I_z(K_1 + K_2) + M(x_1 x'_1 K_1 + x_2 x'_2 K_2)]V}{2[(K_1 + K_2)(x_1 x'_1 K_1 + x_2 x'_2 K_2) - (x_1 K_1 + x_2 K_2)(x'_1 K_1 + x'_2 K_2)] - M(x_1 K_1 + x_2 K_2)V^2} \quad (2.13)$$

$$Z_2 = \frac{M I_z V^2}{4[(K_1 + K_2)(x_1 x'_1 K_1 + x_2 x'_2 K_2) - (x_1 K_1 + x_2 K_2)(x'_1 K_1 + x'_2 K_2)] - 2M(x_1 K_1 + x_2 K_2)V^2} \quad (2.14)$$

3. Effect of suspension roll on two-axle Commercial vehicle dynamics

The effect of suspension roll on two-axle vehicle was compared and analyzed with suspension roll steer values (A/B/C/D) in Table 1 as follows.

Table 1. Suspension roll steer values

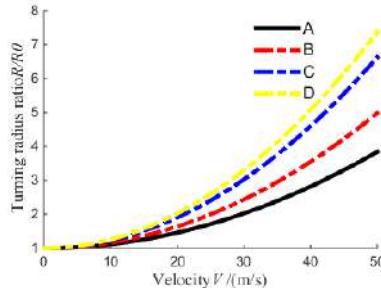
	$\partial\alpha_1 / \partial\phi$	$\partial\alpha_2 / \partial\phi$
A	0	0
B	-0.121	0.031
C	-0.341	0.031
D	-0.341	0.131

(1) The turning radius ratio inclusive of roll steer effect is increased, and the increasing is more with the velocity increasing. By sign convention, positive roll steer value is an oversteering effect on the front axle and an understeering effect on the second axle. Figure 1 shows the relationship between the turning radius ratio and the travelling speed with/without influence of suspension roll steer.

$$R / R_0 = (1 + U_2 V^2) \quad (3.1)$$

where

$$U_2 = \frac{2[(K_1 + K_2)(x_1^2 E_1 K_1 + x_2^2 E_2 K_2) - (x_1 K_1 + x_2 K_2)(x_1 E_1 K_1 + x_2 E_2 K_2)] - M(x_1 K_1 + x_2 K_2)}{2K_1 K_2 (x_1 - x_2)^2} \quad (3.2)$$

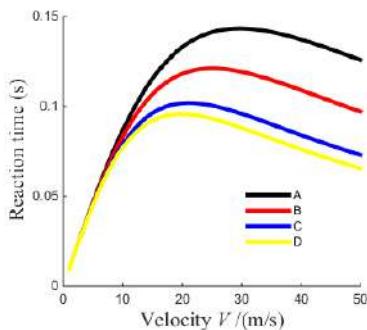
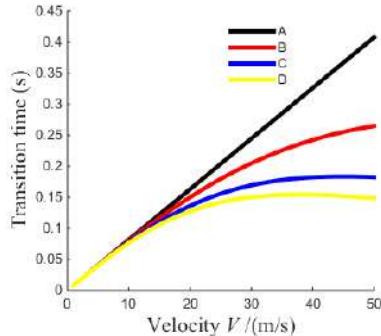
Figure 1. Relation of R / R_0 and V for the two-axle vehicle

The reaction time Z_f was shown in Figure 2 with the effect of vehicle suspension roll. Positive roll steer at the second suspension, or negative roll steer at the front suspension, could reduce the reaction time.

$$Z_f = \frac{I_z V [x_1 K_2 - x_2 K_1]}{2x_1 [(K_1 + K_2)(x_1 x_1' K_1 + x_2 x_2' K_2) - (x_1 K_1 + x_2 K_2)(x_1' K_1 + x_2' K_2)] - Mx_1(x_1 K_1 + x_2 K_2)V^2} \quad (3.3)$$

The transition time Z_c was shown in Figure 3 with the effect of vehicle suspension roll. Positive roll steer at the second suspension, or negative roll steer at the front suspension, could reduce the transition time. The effect of higher values of roll steer at the second suspension lead to decrease the transition time after reaching certain travelling speed.

$$Z_c = \frac{M I_z V}{[I_z (K_1 + K_2) + M(x_1 x_1' K_1 + x_2 x_2' K_2)]} \quad (3.4)$$

Figure 2. Relation of reaction time Z_f and V Figure 3. Relation of transition time Z_c and V

(2) The damping ratio ξ was shown in Figure 4, which was increased with positive roll steer at the second suspension, or negative roll steer at the front suspension. The configurations C and D show higher values of roll steer at the second suspension lead to increase the damping ratio after reaching certain travelling speed.

$$\xi = \frac{[I_z(K_1 + K_2) + M(x_1 x_1' K_1 + x_2 x_2' K_2)]V}{2[(K_1 + K_2)(x_1 x_1' K_1 + x_2 x_2' K_2) - (x_1 K_1 + x_2 K_2)(x_1' K_1 + x_2' K_2)] - M(x_1 K_1 + x_2 K_2)V^2} \sqrt{\frac{M I_z V^2}{4[(K_1 + K_2)(x_1 x_1' K_1 + x_2 x_2' K_2) - (x_1 K_1 + x_2 K_2)(x_1' K_1 + x_2' K_2)] - 2M(x_1 K_1 + x_2 K_2)V^2}} \quad (3.5)$$

(3) According to the K. Guo [7], the calculated **total variance** equation for the two-axle vehicle is

$$E_\delta = \frac{Z_2 + (Z_1 - \tau_1)^2}{2Z_1} \quad (3.6)$$

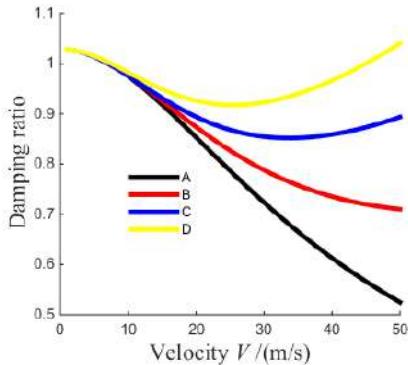
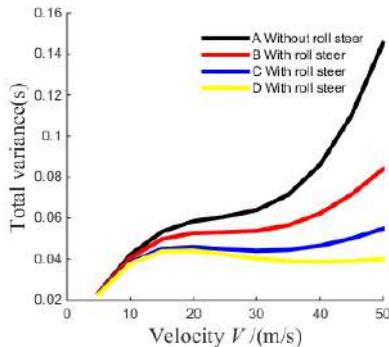
Figure 4. Relation of damping ratio ξ and V Figure 5. **Total variance** for the two-axle vehicle

Figure 5 shows the relationship between the **total variance** and the travelling speed with/without influence of suspension roll steer. Positive roll steer at the second suspension, or negative roll steer at the front suspension, could reduce the **total**

variance. So the vehicle response degree following the step steer input is better. According to equation(3.6), it could get the optimization design parameters for the two-axle vehicle by minimizing the **total variance** E_δ .

4. Conclusions

In conclusion, with positive roll steer at the second suspension, or negative roll steer at the first suspension, the effect of suspension roll was analyzed. The turning radius ratio inclusive of roll steer effect is increased, and the increasing is more with the velocity increasing, changing the vehicle steer characteristics toward understeer. The reaction time and transition time are reduced with the effect of vehicle suspension roll. The damping ratio is increased with the effect of vehicle suspension roll. Positive roll steer at the second suspension, or negative roll steer at the front suspension, could reduce the total variance. The vehicle response degree following the step steer input is better. It could get the optimization design parameters for the two-axle vehicle by minimizing the total variance.

References

- [1] Riekert P, Schunck T E. Zur Fahrmechanik des gummibereiften Kraftfahrzeugs[J]. Archive of Applied Mechanics, 1941, 12(1):70-70.
- [2] Segel L. Theoretical prediction and experimental substantiation of the response of the automobile to steering control[J]. Proceedings of the Institution of Mechanical Engineers: Automobile Division, 1956, 10(1): 310-330.
- [3] Masato Abe, Vehicle Handling Dynamics Theory and Application, Japan (2009), pp. 166-196.
- [4] K. Guo, Vehicle Handling Dynamics, China (1991), pp.245-251.
- [5] Ding J, Guo K. Development of a generalised equivalent estimation approach for multi-axle vehicle handling dynamics[J]. Vehicle System Dynamics, 2016:1-38.
- [6] Youqun Zhao, Guangde Zou, Baohu Liu. Active Safety Index and Calculation Method of Man-Vehicle Road Closed-Loop Control System[J]. Journal of Shandong University of Technology (Natural Science Edition), 1997(2): 7-10.
- [7] K. Guo, D. Wang. Research on evaluation method of vehicle handling stability test at high speed[J]. Automobile Technology, 1980(02):21-30.

Compressive Video Sensing Based on Intra-Inter-Frame Constraints and Genetic Algorithm

Yuchen YUE^a, Hua LI^{a,1} and Jianhua LUO^b

^a*Department of Armament and Control, Army Academy of Armored Forces, Beijing, China*

^b*Center of Maneuver and Training, Army Academy of Armored Forces, Beijing, China*

Abstract. Establishing structured reconstruction models and efficient reconstruction algorithms according to practical engineering needs is of great concern in the applied research of Compressed Sensing (CS) theory. Targeting problems during high-speed video capture, the paper proposes a set of video CS scheme based on intra-frame and inter-frame constraints and Genetic Algorithm (GA). Firstly, it employs the intra-frame and inter-frame correlation of the video signals as the priori information, creating a video CS reconstruction model on the basis of temporal and spatial similarity constraints. Then it utilizes overcomplete dictionary of Ridgelet to divide the video frames into three structures, smooth, single-oriented, or multijointed. Video frames cluster according to the structure using Affinity Propagation (AP) algorithm, and finally clusters are reconstructed using evolutionary algorithm. It is proved efficient in terms of reconstruction result in the experiment.

Keywords. Compressed sensing, genetic algorithm, AP cluster, Ridgelet, overcomplete dictionary

1. Introduction

Compressed sensing theory [1-2] is a new information acquisition theory proposed by D. Donoho, E. Candès, and Chinese scientists T. Tao, which states that as long as a signal is compressible or sparse in some domain of mutation, it can be projected into a low-dimensional space by a measurement matrix, and the measurement data can be reconstructed by solving an optimal model. The measurement data can be reconstructed into the original signal by solving an optimization model. Currently, compressed perception is widely used in the field of video processing [3], and the existing video compressed perception reconstruction algorithms [4-7] are mainly divided into two categories: frame-by-frame reconstruction, i.e., distributed video compressive sensing algorithm, divides the video sequence into group of pictures (GOP), in which the first frame of each GOP is the Key frame and the rest are the CS frames. Key frames are sampled with high sampling rate to obtain better reconstruction quality through intra-frame reconstruction, while CS frames are sampled with lower sampling rate and

¹Corresponding author: Hua Li, Department of Armament and control, Army Academy of armored Forces, Feng tai, Beijing; E-mail: Gonzo123@sina.com.

reconstructed with the aid of Side Information generated from Key frames through motion compensation to obtain high quality reconstructed video, but the disadvantage of this method is that when there are frames with abrupt scene changes in the video, the fixed-length GOP does not produce better results when there are frames with abrupt scene changes in the video, in addition, the algorithm does not make use of the structural prior information of the video. The overall reconstruction class algorithm does not distinguish between Key frames and CS frames at the coding end, but uses different methods for different types of measurements for reconstruction. The reference [8] points out that the structural a priori information of the signal plays a key role in reducing the number of measurements required for accurate signal reconstruction and reducing the uncertainty of the reconstruction model, and for high-speed videos with drastic scene changes, to ensure the efficiency of video acquisition, a structured compressed sensing framework to ensure the efficiency of video acquisition, the sparse representation of the video signal is achieved by constructing a structured perfect dictionary, and a structured measurement scheme suitable for it is formed by using the structural prior information of the video signal, and the reconstruction of high-speed video is achieved based on the structural prior.

In this paper, a genetic algorithm video compression-aware reconstruction algorithm based on video frame motion classification is proposed using the characteristic of relatively single block structure information [9-10]. The algorithm first divides the video sequence into a Data Group (DG) consisting of several consecutive video frames. Then divides all video frames into blocks, and determined the blocks as change class blocks and non-change class blocks based on the judgment of whether the blocks change within adjacent video frames, and then, constructs a Ridgelet over-complete dictionary to determine the block structure type of all video frames in a single DG, retaining only the non-change class block measurements in the previous frame of two adjacent video frames, and retaining all change class block measurements in two adjacent frames. Next, all the retained block measurements are treated as a Data Unit (DU), based on which the DU is clustered using the AP clustering algorithm to make full use of the intra-frame correlation and inter-frame correlation of the video sequence, and the clusters are reconstructed using the genetic evolution algorithm. Finally, the video reconstruction method is trained by video sequence samples, and the reconstruction performance of the proposed algorithm is further improved by training.

2. Block structure classification based on Ridgelet overcomplete dictionary

Video frames have self-similarity. After the non-overlapping blocking operation of a single video frame using the blocking strategy, the obtained blocks contain only a limited number of different structural classes, and the measurement vectors of structurally similar blocks also have similarity. Therefore, the similarity of the measurement vectors of the blocks can be used for metrics and classification can be achieved based on the structural properties of the blocks. Each class of blocks can be represented by the same set of atoms in the overcomplete dictionary. This imposes two requirements on the overcomplete dictionary. First, the overcomplete dictionary should have a redundant orientation structure in order to be able to adaptively and sparsely represent texture and edge information in an image with arbitrary orientation. Second, the overcomplete dictionary should have an efficient search and optimization algorithm that can find the

appropriate combination of atoms and corresponding combination coefficients for the sparse representation of the signal to be reconstructed in the shortest possible time.

The overcomplete dictionary $\mathbf{D} \in \mathbb{C}^{B \times N}$ used to represent the block is noted as $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N)$, $i = 1, 2, \dots, N$, where \mathbf{d}_i is the atom in dictionary \mathbf{D} with ordinal number i , generated in the manner of equation (1).

$$\mathbf{d}_i(\mathbf{z}) = \frac{1}{W} \left[e^{-(a_i \mathbf{u}_i^T \mathbf{z} - b_i)^2/2} - \frac{1}{2} e^{-(a_i \mathbf{u}_i^T \mathbf{z} - b_i)^2/8} \right] \quad (1)$$

Where $\mathbf{d}_i(\mathbf{z}) \in \mathbb{C}^{\sqrt{B} \times \sqrt{B}}$ is the atom with the same size as the block, and $\mathbf{d}_i \in \mathbb{C}^B$ is the vectorized atom. $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2) \in [0, 1, 2, \dots, \sqrt{B} - 1]^2$ is the position vector of the atom, and $1/W$ is the normalization factor of the atom. The atom $\mathbf{d}_i(\mathbf{z})$ in the dictionary \mathbf{D} corresponds to the parameter set $\gamma_i = (\theta_i, a_i, b_i)$, where a_i is the scale parameter, b_i is the displacement parameter, θ_i is the orientation parameter, $\mathbf{u}_i = (\cos \theta_i, \sin \theta_i)^T$.

After the prototype atoms are selected, the atomic size of the dictionary and the sparse representation capability of the video frames also depend on the range of values of a_i , b_i and θ_i , their respective discrete intervals. Following the Ridgelet dictionary construction method of the reference [11], the parameter space is set as equation (2).

$$\Gamma_\gamma = \{(\theta, a, b) \mid \theta \in [0, \pi], a \in [0, 3], b \in \Gamma_b\} \quad (2)$$

Where the displacement parameter b_i is taken in relation to the orientation parameter θ_i , as shown in equation (3).

$$\Gamma_b = \begin{cases} \left[0, \sqrt{B}(\sin \theta + \cos \theta) \right], & \text{if } \theta \in [0, \pi/2) \\ \left[\sqrt{B} \cos \theta, \sqrt{B} \sin \theta \right], & \text{otherwise} \end{cases} \quad (3)$$

When using an overcomplete dictionary for the sparse representation of video frames, the individual atoms in the dictionary respond only to those blocks that are consistent in shape, scale, position and orientation information. Meanwhile, the subsequent algorithm wants to determine the block structure characteristics and classify them based on the characteristics of the sub-dictionaries. Therefore, the discrete interval of the orientation parameters must be small enough when discretizing the Ridgelet overcomplete dictionary.

Since two block sizes, 8×8 and 16×16 , are used in the experiment, two Ridgelet overcomplete dictionaries are constructed. Setting $a \in [0, 3]$, $\theta \in [0, 179]$, scale interval and displacement interval to 0.2 and 1, respectively, and directional interval to 5. Therefore, the constructed Ridgelet overcomplete dictionary for 8×8 blocks have 6271 atoms, and the Ridgelet overcomplete dictionary for 16×16 blocks have 12032 atoms. Ridgelet overcomplete dictionary for 16×16 blocks have a total of 12032 atoms.

Smoothed blocks are insensitive to the orientation parameter, so the atoms of each direction have good reconstruction quality for this type of blocks, and only the sub-

ictionaries of individual directions have high reconstruction quality for single-direction blocks and multi-direction blocks. The process of reconstructing blocks undoubtedly increases the difficulty of block structure feature determination, so the analytical formula-based method proposed in reference [12]. Firstly, we determine whether the block is smooth by the relationship between the variance of the block measurements and the mean variance of the block, and then we find the 10 most relevant atoms in the directional sub-dictionary over the complete dictionary along the same lines and solve the block measurement residuals by the least squares' method. If the block has the largest measurement value residual on a single direction sub-dictionary, the corresponding block is marked as a single-direction block, and vice versa as a multi-direction block. To ensure the fast convergence of the algorithm and to facilitate the optimal direction statistics of multi-directional class blocks, only the three directions with the smallest residuals are marked for all blocks marked as multi-directional blocks with the number of directions greater than or equal to 3. This method is inaccurate in the determination of block structure features at low sampling rate, but the purpose of adopting this method in this paper is block structure feature classification, not exact reconstruction, so it can be used for determination.

Smoothed blocks are insensitive to the orientation parameter, and the atoms of each direction have good reconstruction quality for this type of blocks. Only the sub-dictionaries of individual directions have high reconstruction quality for single-direction blocks and multi-direction blocks. The process of reconstructing blocks undoubtedly increases the difficulty of block structure feature determination, so the analytical formula-based method proposed in the reference [12]. Firstly, we determine whether the block is smooth by the relationship between the variance of the block measurements and the mean variance of the block, and then we find the 10 most relevant atoms in the directional sub-dictionary over the complete dictionary along the same lines and solve the block measurement residuals by the least square's method. If the block has the largest measurement value residual on a single direction sub-dictionary, the corresponding block is marked as a single-direction block, and vice versa as a multi-direction block. To ensure the fast convergence of the algorithm and to facilitate the optimal direction statistics of multi-directional class blocks, only the three directions with the smallest residuals are marked for all blocks marked as multi-directional blocks with the number of directions more than or equal to 3. Even though this method is inaccurate in the determination of block structure features at low sampling rate, but the purpose of adopting this method in this paper is block structure feature classification, not exact reconstruction, so it can be used for determination.

3. Genetic Algorithm-based Compressive Video Sensing Algorithm

3.1. The overall framework of the algorithm

The flow of the genetic algorithm-based compressive video sensing algorithm is shown in Figure 1.

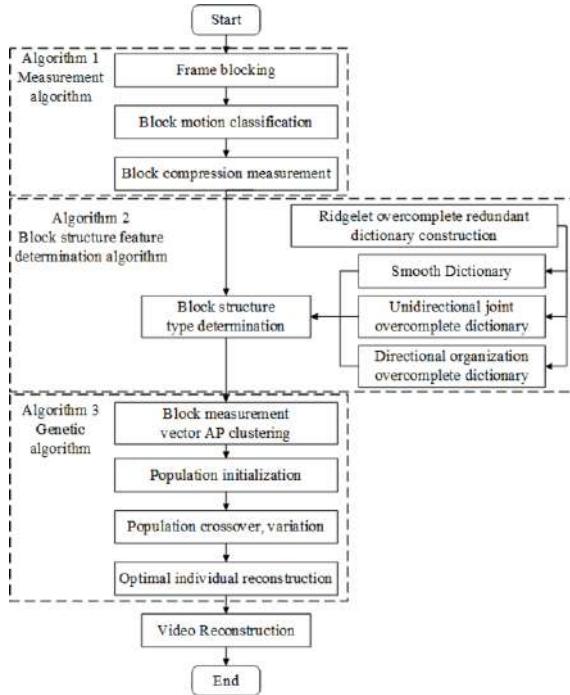


Figure 1. Genetic algorithm-based compressive video sensing algorithm flow block diagram.

The proposed genetic algorithm-based compressive video sensing algorithm consists of three parts: measurement algorithm, block structure type determination algorithm and genetic algorithm. Firstly, a blocking operation is implemented for each frame of the input video and a plurality of consecutive video frames are set as one DG, and all the blocks in a single DG are compressed and sampled to obtain measurement values. The blocks at the same position in two adjacent video frames are classified by motion through calculation and divided into change blocks and non-change blocks, and a single non-change block and all change blocks are retained. Then the measurements of all the retained blocks in DG are treated as a DU and subjected to AP clustering to maximize the intra- and inter-frame correlation of the video sequences, and the reconstruction is performed uniformly for the same clusters. The reconstruction algorithm uses Genetic Evolutionary Algorithm (GA).

3.2. Measurement Algorithm

A new video sampling scheme is proposed to make the proposed algorithm have better reconstruction quality in the changing part of the video. After the video frames are divided into multiple $B \times B$ -sized blocks, a Data Unit (DU) is set with n consecutive frames, and the l_2 -norm of the difference between the blocks at the same position in two adjacent frames in the DU is calculated, and a threshold σ is set. The blocks with the difference l_2 -norm value of the blocks at the same position in two adjacent frames less than the threshold σ are marked as non-changing blocks, and then the blocks at the same position in the previous video frame in two adjacent frames are kept and recorded; the blocks with the difference l_2 -norm value greater than the threshold σ are marked as changing blocks. The blocks in the same position in two adjacent frames are marked as

non-changing blocks, and then the blocks in the same position in the previous two frames are kept and recorded; the blocks with the difference l_2 -norm value greater than the threshold σ are marked as changing blocks, and all the changing blocks in the DU are recorded and kept.

3.3. Population Initialization Algorithm

After obtaining the measurement vectors of all the reserved blocks in the DU, the block structure classification is determined, which lays a solid foundation for the population initialization in the subsequent evolutionary algorithm solving. Different types of blocks adopt different population initialization strategies. The block structure category judgment method adopts the judgment method proposed in section 2:

①Smooth blocks

Dividing the Ridgelet over-complete dictionary into 16 sub-dictionaries according to the scale, and selecting the first 7 scale to form the sub-dictionary of the smoothing block, and using this to randomly generate the initial population corresponding to the smoothing block.

②Single directional blocks

Divide the Ridgelet overcomplete dictionary into 36 sub-dictionaries according to the direction, select the forward 2 directions and backward 2 directions including the dominant direction of the corresponding single direction block, a total of 5 directions to form the single direction sub-dictionary, and use this to randomly generate the initial population corresponding to the single direction block.

③Multi-directional blocks

In the process of judging multi-directional blocks, all the most relevant directions in the measurement vector of the block are first counted, and then the most dominant 3 directions are selected, and each direction is selected in the way of single-directional blocks in the forward 2 directions and the backward 2 directions, a total of 15 directions constitute the sub-dictionary of multi-directional blocks, and the initial populations corresponding to multi-directional blocks are randomly generated.

3.4. Reconstruction based on Genetic Algorithm

The reconstruction algorithm uses Genetic Algorithm (GA), and GA needs to consider the following five issues in the design phase.

①Individual Codes

An $N \times N$ -dimensional Ridgelet overcomplete dictionary is constructed, and a real number encoding is adopted to ensure that the atoms of this Ridgelet overcomplete dictionary represent the signal \mathbf{x} as optimally as possible, while each atom in the dictionary is numbered. Let the individual population be $P = [\rho_1, \rho_2, \dots, \rho_K]$, each gene position ρ_i ($1 \leq i \leq N$) in P is the number of the over-complete dictionary atoms, and K is the signal sparsity.

②Adaptation function

The fitness function is an index to search for feasible solutions, and its mathematical expression is shown in equation (4).

$$f(Y_m) = \frac{1}{\sum_{i=1}^n \|\mathbf{y}_i - \Phi \text{dec}(Y_m) \alpha_i\|_2^2} \quad (4)$$

where $f(Y_m)$ is the fitness of the same clusters, \mathbf{y}_i is the measurement vector of the signal, Y_m is the set of measurement vectors of the same clusters, $\text{dec}(Y_m)$ is the optimal set of atoms of Y_m , n is the capacity of Y_m , i is the number of individuals within Y_m , and Φ is the measurement matrix.

Because of the property that there are a large number of inter- and intra-frame similar blocks in the video. Using the AP clustering is used to cluster the measurement vectors of the blocks separately by the structural characteristics of the blocks Y_m , and the inverse of the residual sum of the block measurements and the clustered measurements is used as the fitness function. The sum of the fitness functions is found to take advantage of the constraints between similarly structured measurement data, while effectively avoiding possible deviations in the measurement vector search process, resulting in a blockier and more accurate search for the optimal combination atoms of Y_m .

③Decoding process

The GA uses a non-convex l_0 -model for compressed sensory reconstruction. First, search for the $\text{dec}(Y_m)$ that maximizes the $f(Y_m)$ in the reconstruction process. Then, solves the sparse representation coefficient α for each measurement corresponding to the dictionary atom using a least-squares algorithm. The mathematical expression is shown in equation (5).

$$\alpha = (\Phi \text{dec}(Y_m))^+ \mathbf{y} \quad (5)$$

Finally, using Ridgelet overcomplete dictionary \mathbf{D} to reconstruct the original signal. The mathematical expression is shown in equation (6).

$$\hat{\mathbf{x}} = \mathbf{D}\alpha \quad (6)$$

④Population initialization

According to Table 1 to initialization the population.

Table 1. Population initialization algorithm

Step 1 :	Determining the structural category of video blocks, marking smooth blocks and single/multi-directional video blocks respectively, with different population initialization generation methods for different categories of video blocks.
Step 2 :	The Ridgelet overcomplete dictionary is divided by scale, and the first seven scales are selected to form a smooth block sub-dictionary, from which the corresponding initial populations are randomly generated.
Step 3 :	The Ridgelet overcomplete dictionary is divided by direction, and after marking the dominant direction i of the unidirectional block, a total of five directions $\{i-2, i-1, i, i+1, i+2\}$ in the forward and backward adjacent directions are selected to form the unidirectional block sub-dictionary, and the population is initialized.
Step 4 :	Count the most dominant three directions $\{j, k, l\}$ among the measurement vectors marked as multi-directional blocks, and record only the most dominant three when there are more than 3. Each dominant direction (j) is selected as a single-directional block in the way of 5 directions $\{j-2, j-1, j, j+1, j+2\}$ respectively to form a single-directional block sub-dictionary and then combined into a multi-directional block sub-dictionary with a total of 15 directions.

Generating individuals for all directions (j, k, l) according to the above operation, and finally combining all individuals to become the initialized population of the multi-directional block class.

⑤Genetic operators

There are three types of genetic operators, crossover, mutation and selection, and the specific operators are designed as follows.

- Crossover operators

A single-point crossover operation is used, in which a gene locus is randomly selected in two randomly paired individuals, and the lower half of the two individuals are swapped at the midpoint of the gene locus to create two new individuals. Crossover occurs randomly according to a set crossover probability P_c .

- Mutation operator

Using mutation operator based on directional statistics, mutation occurs randomly according to a set probability of mutation P_m , all genetic positions in individuals smaller than P_m are mutated, and the mutation value used to replace the original value of the genetic position comes from the sub-dictionary of individuals with the least directional genetic position.

- Selection operator

The fitness of the two new individuals generated from the crossover of two paired parent individuals is solved by the fitness function respectively, and if the fitness of one of the two parent individuals is smaller than that of the individual with larger fitness in the new individual, this parent individual is removed from the parent population and the new individual is added, otherwise, no operation is performed.

Finally, the individual with the largest fitness within the population is computationally selected as the optimal atomic group for the reconstructed signal measurement vector clustering, and the inverse least squares operation is used to complete the signal reconstruction.

The pseudo-code of the genetic algorithm-based compressive video sensing reconstruction algorithm is shown in Table 2.

Table 2. Genetic algorithm-based compressive video sensing reconstruction algorithm

Input:	video Sequence \mathbf{X} , random Gaussian measurement matrix Φ , number of iterations, Ridgelet overcomplete dictionary \mathbf{D}
	Measurement of \mathbf{X} ; Determining the blocks as smooth, single directional, multi-directional; Extracting structurally identical blocks and AP clustering; Set the population size and sparsity. The population initialization is performed using the corresponding initialization scheme according to the clustering of the block measurement data;
Iteration:	Crossover and mutation operations on individuals in the population; Selecting the best individuals in the population for reconstruction $\hat{\mathbf{x}}$: $\alpha = (\Phi \text{dec}(\mathbf{x}_m))^+ \mathbf{y};$ $\hat{\mathbf{x}} = \mathbf{D} \alpha;$ Reconstruct the video sequence $\hat{\mathbf{x}}$ by recombining the estimated signal $\hat{\mathbf{X}}$ according to the chunking and coding strategy;
Output:	Reconstructed video $\hat{\mathbf{X}}$

4. Performance Analysis of Proposed Algorithm

In this section, several experiments are designed to verify the performance of the proposed evolutionary algorithm-based video compression perception algorithm and to optimize the parameter settings involved in the algorithm. The basic configuration of the experiments includes: the experimental platform is a laptop-type computer with Intel Core i7 2.8GHz CPU, 16GB RAM, and 64-bit Win7 operating system installed. All experiments were conducted in MATLAB R2017a.

4.1. Analysis of the impact of sampling rate on algorithm performance

The following experiments analyze the performance of the proposed algorithm at different sampling rates.

The experimental object is a 96-frame Traffic video sequence of 256×256 , with video frame chunk size $B = 8$. Construct an overcomplete dictionary based on Ridgelet, and set the direction range $[0, 179]$, direction interval as 5, total 36 directions, scale range $[0, 3]$, scale interval as 0.2, total 16 scales. The dictionary size as 6172. Set the population size of the genetic algorithm is 36, the individual size is 12. Set the crossover probability $P_c = 0.6$ and the mutation probability $P_m = 0.01$. Let 8 consecutive frames are used as a DG, and the number of smooth block iterations is set to 20, the number of single-direction block iterations to 50, and the number of multi-direction block iterations to 100. the sampling rates are set to 0.125, 0.2, and 0.25, respectively, and the algorithm performance evaluation criterion is the average PSNR of the reconstructed frames.

The comparison experiments are reconstructed using the Gaussian mixture model proposed in the reference [13], and to ensure fairness, the algorithm in the reference [13] is trained offline in advance to obtain a set of mean and variance as the initial parameters of the Gaussian mixture model. When the sampling rate is 0.125, the Gaussian mixture model reconstructs the sum of measurement data generated by making $\{0, 1\}$ measurements with probability 0.5 for 8 consecutive frames; when the sampling rate is 0.2, the Gaussian mixture model reconstructs the sum of measurement data generated by making $\{0, 1\}$ measurements with probability 0.5 for 5 consecutive frames; when the sampling rate is 0.25, the Gaussian mixture model reconstructs the sum of measurement data generated by making $\{0, 1\}$ measurements with probability 0.5 for 4 consecutive frames.

The first 8 frames of the Traffic sequence are used as an example for analysis. Figure 2 shows the consecutive images of the original sequence, and Figure 3 shows the enlarged image of the 3rd frame of the first 8 frames.



Figure 2. The first 8 original frames of Traffic sequence.

The proposed algorithm and the Gaussian mixture model proposed in the reference [13] are used to reconstruct the Traffic sequences respectively. When the sampling rate is 0.125, the reconstruction results of both for the first 8 frames of Traffic sequence are shown in Figure 4, where Figure 4(a) is the reconstruction result of the algorithm of reference [13], and Figure 4(b) is the reconstruction result of the proposed algorithm; when the sampling rate is 0.2, the reconstruction results of both for the first 8 frames of

Traffic sequence are shown in Figure 5, where Figure 5(a) is the reconstruction result of the algorithm of reference [13], and Figure 5(b) is the reconstruction result of the proposed algorithm. Figure 5(b) is the reconstruction result of the proposed algorithm; when the sampling rate is 0.25, the reconstruction results of both for the first 8 frames of the Traffic sequence are shown in Figure 6, where Figure 6(a) is the reconstruction result of the algorithm of the reference [13] and Figure 6(b) is the reconstruction result of the proposed algorithm.



Figure 3. The 3rd original frames of Traffic sequence.

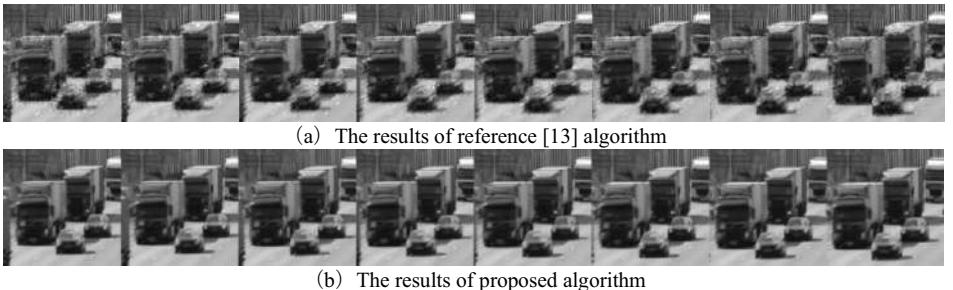


Figure 4. The results of first 8 reconstruction frames when sample rate is 0.125.

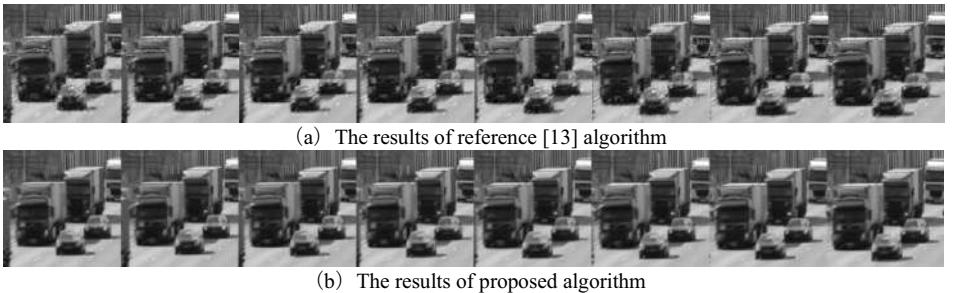


Figure 5. The results of first 8 reconstruction frames when sample rate is 0.2.

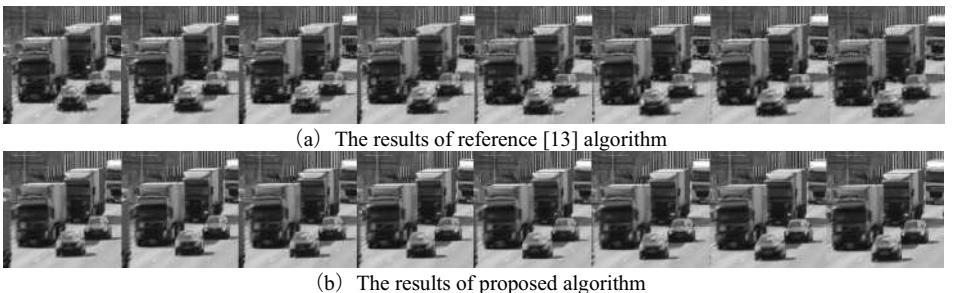


Figure 6. The results of first 8 reconstruction frames when sample rate is 0.25.

Figure 7 shows the reconstructed 3rd frame zoomed image of Traffic sequence for two algorithms with sampling rates of 0.125, 0.2 and 0.25 respectively. The PSNR of the

reconstructed Traffic sequences for the two algorithms at sampling rates of 0.125, 0.2 and 0.25 are respectively shown in Table 3.

Combining the results from Figure 4 to Figure 6 and Table 3, it can be seen that the average rate distortion performance of the proposed algorithm at all three sampling rates is better than that of the Gaussian mixture model proposed by reference [13].



Figure 7. Two algorithms to reconstruct the zoomed image of the 3rd frame of Traffic sequence at three sampling rate.

Table 3. The PSNR of two algorithms at sampling rates of 0.125, 0.2 and 0.25

DG Serial	Frame Serial	SR=0.125		SR=0.2		SR=0.25	
		Ref. [13]	Proposed	Ref. [13]	Proposed	Ref. [13]	Proposed
1	1-8	20.80	26.40	24.04	26.65	23.38	28.05
2	9-16	21.21	25.88	23.29	26.40	22.37	27.07
3	17-24	22.27	25.29	22.41	25.48	22.19	27.15
4	25-32	23.02	24.61	21.48	25.68	23.20	26.99
5	33-40	22.76	25.45	21.53	24.75	22.20	26.64
6	41-48	22.03	25.25	21.41	25.19	22.24	26.78
7	49-56	21.33	24.37	20.53	25.32	22.52	27.00
8	57-64	21.58	24.81	20.48	25.89	21.59	26.94
9	65-72	21.19	24.89	21.26	25.60	21.67	26.14
10	73-80	22.18	24.29	21.66	26.58	20.89	25.82
11	81-88	21.62	23.38	20.61	25.57	21.60	26.45
12	89-96	20.55	24.07	21.13	24.54	22.60	26.81
Average		21.71	24.89	21.65	25.64	22.20	26.82

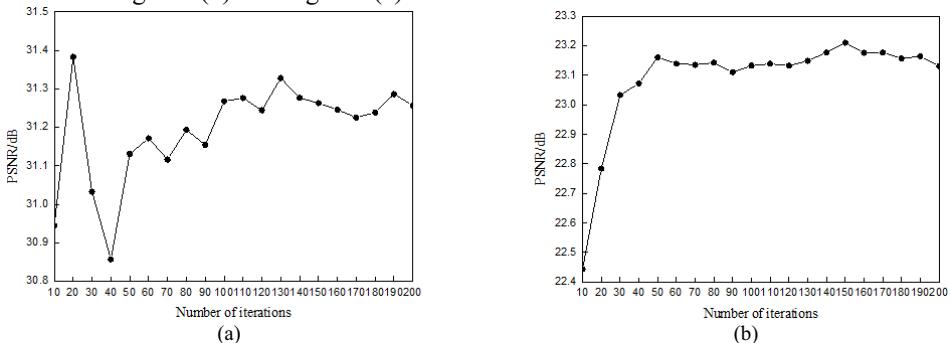
As can be seen from Figure 7, the proposed algorithm not only has a higher PSNR value when the sampling rate is 0.125, but also the reconstructed video frames from the subjective visual point of view have a good reconstruction effect in the edge part has a good reconstruction effect, while the reconstructed image by the Gaussian mixture model of reference [13] shows a large blur in the contour of the change region of the adjacent frames. When the sampling rate is 0.2, it can be seen from Figures 7 that the proposed algorithm demonstrates a better reconstruction detail effect, and the blurring phenomenon in the change region of the reconstructed video frames by the Gaussian

mixture model reconstruction algorithm proposed in the reference [13] is reduced, but the reduction effect is not obvious, and in general the blurring phenomenon in the edge of the change region still exists. When the sampling rate is 0.25, both algorithms improve in image details and visual effects, and the proposed algorithm still has better reconstruction performance.

4.2. Analysis of the Impact of Generation Number on Algorithm Performance

The number of evolutionary iterations set in section 4.1 are: the number of iterations for the smooth block is 20, the number of iterations for the unidirectional block is 50, and the number of iterations for the multi-directional block is 100. Next, the influence of generation number on the performance of the proposed algorithm is analyzed through experiments. In this experiment, the basic experimental parameters are the same as in section 4.1, and the overcomplete Ridgelet dictionary of 16×16 from Section 4.1 is used. The dictionary contains 12032 atoms, and the chunk size is 16×16 , with 8 frames as a DG. The experimental object is a 96-frame Traffic video sequence of size 256×256 .

First of all, there are the experiment of smoothing block evolution generations is performed. The limit number of iterations is set to 200, and the reconstruction data during the evolution process is saved. The average PSNR value of the reconstructed video frames is calculated every 10 iterations as the evaluation criterion of the reconstruction quality. The change of the average rate distortion performance of the smoothing block with the increase of the evolution number is shown in Figure 8(a). The horizontal coordinate in Figure 8 is the number of evolutionary generations, and the vertical coordinate is the average PSNR value of the smoothed block obtained by the proposed algorithm. From the experimental results, the average PSNR value of the reconstructed smoothing block reaches the peak when the evolution proceeds for 20 generations, and then the average PSNR of the reconstructed smoothing block obtains the minimum value when the evolution proceeds for 30 generations, and then the average PSNR value increases continuously with the increase of the evolution generations and eventually tends to level off. Therefore, the number of iterations of 20 is selected as the optimal evolutionary generation of the smoothing block. Similar experiments were also conducted to obtain the relationship curves between the number of iterations and the average PSNR value for the single-directional block and the multi-directional block as shown in Figure 8(b) and Figure 8(c).



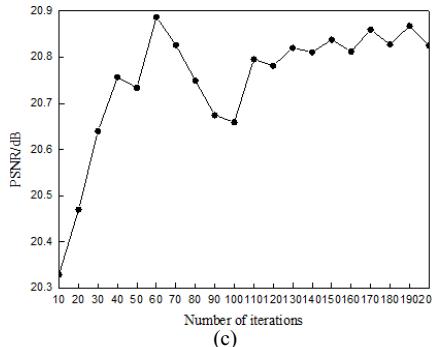


Figure 8. Number of iterations analysis diagram. (a) smooth blocks (b) single-directional blocks (c) multi-directional blocks.

According to Figure 8(b) and Figure 8(c), it can obtain the optimal evolutionary generations of 50 for the single-direction block and 100 for the multi-direction block. Therefore the values are taken as 20 generations for smooth blocks, 50 generations for single-direction blocks and 100 generations for multi-direction blocks, respectively.

5. Conclusion

This paper firstly describes the relevant technique used in the proposed video compressive sensing reconstruction algorithm based on intra-and-frame and evolutionary algorithm - the block structure type determination method based on Ridgelet over-complete dictionary, then describes in detail the implementation process of the proposed algorithm, designs a series of experiments, and analyzes different Then, a series of experiments are designed to analyze the effects of different sampling rates, data unit frames, block sizes and iterations on the performance of the algorithm, and determine the number of block evolution generations, block sizes and data unit frames that make the algorithm perform best.

References

- [1] Donoho D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4):1289-1306.
- [2] D L Donoho, Y Tsai. Extensions of Compressed Sensing [J]. *Signal Processing*. 2006, 86(3). 533-548.
- [3] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2481-2499, May 2012.
- [4] LIU Y and PADOS D A. Compressed-sensed-domain L1-PCA video surveillance [J]. *IEEE Transactions on Multimedia*, 2016, 18(3): 351-363.
- [5] GUO J, SONG B, and DU X. Significance evaluation of video data over media cloud based on compressed sensing[J]. *IEEE Transactions on Multimedia*, 2016, 18(7): 1297-1304.
- [6] REHMAN A U, SHAH G A, and TAHIR M. Compressed sensing based adaptive video coding for resource constrained devices[C]. *IEEE International Wireless Communications and Mobile Computing Conference*, Paphos, Cyprus, 2016: 170-175.
- [7] WANG J, GUPTA M, and SANKARANARAYANAN A C. LiSens A scalable architecture for video compressive sensing[C]. *IEEE International Conference on Computational Photography*, Houston, TX, 2015: 1-9.
- [8] Li Wan. Compressed Sensing Observation and Reconstruction Based on Structured Sparse and Convolutional Network [dissertation]. Xi'an: Xidian University, 2018.9.
- [9] M. Wang, A design and research on the compressed sensing measurement matrix, [dissertation] Xidian University, 2014.

- [10] Mun S, Fowler J E. Block compressed sensing of images using directional transforms[C]. International Conference on Image Processing, 2009: 3021-3024.
- [11] Do M N, Vetterli M. The Contourlet transform: an efficient directional multiresolution image representation[J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- [12] Xu Jinghuan. The Design of Ridgelet Redundant Dictionaries and Reconstruction Algorithm[D]. Xi'an: Xidian University, 2011.
- [13] Li X , Lan X , Yang M, et al. A new compressive sensing video coding framework based on Gaussian mixture model[J]. SIGNAL PROCESSING-IMAGE COMMUNICATION, 2017, 55:66-79.

Koch Snowflake Fractal Antenna Design in the Deep Space Bands for a Constellation of Cubesat Explorers

Orlando Francois GONZALES PALACIOS^{1,a,b}, Ricardo Erick DIAZ VARGAS^c,
Patrick H. STAKEM^d and Carlos Enrique ARELLANO RAMIREZ^b

^a*Beijing University of Aeronautics and Astronautics, Beijing, China*

^b*Universidad Nacional de Piura, Piura, Perú*

^c*Pontifícia Universidad Católica del Perú (PUCP), Lima, Perú*

^d*Office of STEM & Cubesat Inclusion GSFC (NASA), USA*

Abstract. This paper presents the design and simulation of a Koch curve fractal antenna, developed according to the second iteration of the Koch snowflake fractal for S-band, C-band, X-band and Ku-band. The simulated antenna shows good performance for the operating frequencies and desirable gain, bandwidth and VSWR parameters. Being a compact antenna, it has a size, geometry and characteristics that go in accord with the CubeSat's structure standards. The antenna was fabricated on a 1.5 mm thick FR-4 substrate. The VSWR achieved values are lower than 1.4 for the frequencies used (2.1 GHz to 2.4 GHz and 7.4 GHz to 8.9 GHz) with a simulated omnidirectional radiation pattern. A maximum gain of **6.8 dBi** was achieved. As this antenna works optimally in the S, C and X bands, it is adequate for deep space applications, especially in low-power consumption systems. This approach would be ideal for constellations of Cubesat explorers.

Keywords. Antenna, microstrip, fractal, S-band, C-band, X-band, Ku-band, deep space, cubesat, constellation, swarm, communication

1. Introduction

Aerospace technology in recent years has advanced exponentially, but it continues to present new challenges that stimulate human ingenuity. In the next decade, we will face the new challenge that is the colonization of the Moon and Mars. This boom in space exploration requires new design alternatives that meet the demands of increasingly complex aerospace standards. NASA is committed to landing American astronauts on the Moon by 2024, through the Artemis lunar exploration program [1], new technologies and innovative systems will be used for deep exploration of the Moon and deep space.

Possible applications on deep space exploration as the case of a “mothership” spacecraft carrying multiple Cubesats with multiple targets as an objective, such as an asteroid belt [2], or areas of phenomena of interest where multiple simultaneous observations are required are discussed in this paper. Within the communication group,

¹ Corresponding author: Orlando Francois GONZALES PALACIOS; E-mail: a20156013@pucp.edu.pe

a simple TCP/IP protocol could be used between the mothership and the various explorer Cubesats [2]. The mothership serves as an in-situ store-and-forward communication node.

Contributing to these technological innovations, this scientific article works on the design of a Koch snowflake fractal antenna that works in the S-band, C-band and X-band, which will be used in a constellation of CubeSats or small satellites to carry out scientific investigation in deep space. The communication capabilities (hardware and software) of the CubeSat satellite standard allow the transmission of data and telemetry for missions in LEO. However, interplanetary communications with 1 to 3 U Cubesats remain a challenge due to the limited size and power of these spacecrafts [3]. Recent advances in antenna and amplifier designs have expanded the communication possibilities for the Cubesat standard, although more research work is still needed on Cubesat communication systems. However, the simplicity and low cost of CubeSats make them attractive candidates for scientific missions to the Moon [4] and study experimental payloads such as this “Koch Snowflake Fractal Antenna” [5].

2. Deep Space Cubesat Constellation Application

2.1. Cubesat Satellites

A Cubesat is a small and affordable satellite that can be developed and launched by colleges, high schools, and even individuals. The specifications were developed by California Polytechnic State University (Cal Poly) and Stanford University in 1999. The basic structure is a 10-centimeter cube, weighing less than 1.33 kilograms [6]. This allows multiple of these standardized packages to be launched as secondary payloads on other missions [7].

The Cubesat design specification, developed by Cal Poly, defines the physical and interface specifications for Cubesats, and gives testing requirements for vibration, thermal-vacuum and shock tests, as well as safety tests [7]. Since a Cubesat flies along with other Cubesats in a deployment device, and with a primary payload, safety is a key concern.

2.2. Constellation Systems

Constellations are groups of satellites operating together to observe a single target. A constellation allows simultaneous observations of one target from multiple locations, or multiple targets simultaneously [8]. The step beyond constellation systems is a swarm architecture. In swarm robotics, the key issues are the communication between units, and cooperative behavior. The fractal antenna for this constellation of satellites is designed to primarily serve small satellites using UHF and VHF, as they are the most popular bands in nano and microsatellite standards. The capability of individual units does not matter much; what matters is the strength in numbers, and for this reason the given approach is applicable to groups of explorers arranged as constellations, clusters, and swarms.

2.3. Relay Constellation Orbital Design

A relay satellite constellation that orbits on a low altitude orbit will allow us to establish an autonomous communication network that provides connectivity between the surface nodes, satellite constellations and base stations on Earth. Some orbit options for this configuration are the following: Malapert station [9], L1 & L2 halo constellation [10], hybrid constellation, elliptical orbit constellation, polar circular orbit constellation and inclined circular orbit constellation [11]. These orbital alternatives are depicted on Figure 1.

2.4. Communication within collection of Cubesats (relay communication)

The Cubesat's small antennas, and relatively low power (1W) means we have to get clever with communications. One approach is to have the dispenser or mothership handle communications with Earth, and have short-range communication with the Cubesats using the store-and-forward technique [12]. A relay communication constellation system for low lunar orbits is proposed to be used in conjunction with the fractal antenna presented in this paper. As low-orbit relay satellites will pass over the surface sites relatively quickly, a high-orbiting mothership is needed to provide near-continuous communication support [13].

2.5. The Mothership

The mothership would also support point to multipoint communications between the various units of Cubesats or units on the surface. The protocol for communication between Cubesats and the mothership could be as simple as TCP/IP [12]. It would be ideal to consider the Autonomous Space Communications Technology (ASCoT) [14]. Since it nimbly executes a media access control (MAC) scheme, it simultaneously drives electronically directed arrays and switches communication links between multiple nodes separated in space or on the surface.

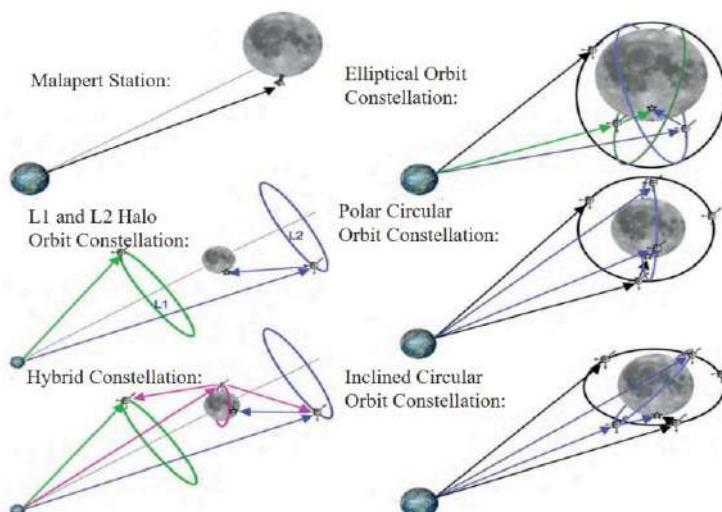


Figure 1. Relay Constellation Orbital Design configurations studied by the SCAWG [11]

3. Antenna Design

The proposed antenna must cover the S-band, C-band and X-band for deep-space communication standards [15]. 1U Cubesat standard states a satellite's dimension must be 10x10x10cm, therefore the antenna is constrained by this requirement to avoid the implementation of complex deployment systems. A fractal pattern antenna easily complies with these two requirements, due to its characteristic antenna miniaturization and broad bandwidth, multi band coverage. Since the antenna will be used for a low power application, the radiating pattern will be printed as a microstrip on a FR-4 substrate [16].

3.1. Fractal Design

Fractal patterns are self-contained and self-similar; thus, a fractal shape replicates itself as a fraction of the original shape with each iteration. As the fractal keeps self-replicating limited to a certain space, with a space-filling behavior, the Chu limit for small antennas is approached with higher iterations [17]. Among fractal patterns, Koch snowflake is generated using an iterative function system on an equilateral triangle [18]; thus, the perimeter of the Koch loop curve is [19]:

$$P_n = 3\sqrt{3} r \left(\frac{4}{3}\right)^n$$

Where r is the radio of the circle that circumscribes a Koch snowflake, and n is the number of iterations. Two iterations have been chosen as further iterations will resolve in smaller details which complicate the antenna fabrication [20]. The monopole and snowflake lengths have been calculated for the chosen fractal radius, as shown on Table 1.

Table 1. Monopole, snowflake and iteration lengths according to fractal radio.

Fractal radius (cm)	Iteration	Monopole length (cm)	Snowflake length (cm)	Segment length (cm)
1.767	0	3.061	9.18	3.06
1.767	1	4.081	12.24	1.02
1.767	2	5.442	16.33	0.34

3.2. Microstrip Antenna Parameters

This microstrip antenna is mounted on a 1.5mm FR-4 substrate and follows the electrical length correction procedures presented on [18]. The Koch snowflake fractal loop is imprinted as a microstrip line on the substrate, thus the dimensions shown on Table 1. have to be tuned to match substrate electromagnetic parameters. The results of these recalculations are shown on Table 2. where L is the pattern base length and L_n is the n -th iteration length ($n=2$), while $Loop_w$ is the pattern microstrip width. Antenna is fed by a 50 Ohms microstrip line with a $Feed_w$ width [5]. Final design of the antenna front and rear views along with its parameters are depicted in Figure 2.

Table 2. Antenna dimensions for bandwidth coverage optimization in the X-Band

Parameter	Dimensions (mm)	Parameter	Dimensions (mm)
L	30.61	$Subs_w$	44
L_n	3.4	$Subs_h$	42
$Feed_w$	2.9	Gr_w	6.00
$Loop_w$	6	Gr_h	26.30

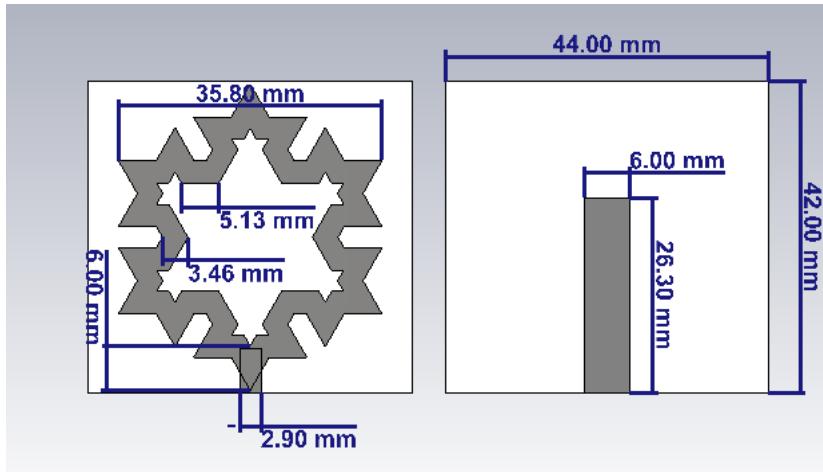


Figure 2. Koch snowflake fractal antenna dimensions

4. Simulation and Results

The dimensions of the substrate, fractal geometry and ground plane previously calculated are now optimized with computational electromagnetic finite element modeling. The proposed antenna modeling and X-band focused optimization have been carried out using EM CAD Microwave Studio by Computer Simulation Technology (CST®). The dimensions we obtained after the optimization are shown in Section 3.2 and the antenna performance evaluation in the sections below. This setting achieves a maximum gain of 6.8 dBi for the X-band uplink channel.

4.1. Bandwidth: S_{11} Parameter & VSWR

An antenna bandwidth is calculated according to its wave reflection parameters when matched to a load; these values can be presented as voltage standing wave ratio (VSWR) or as input reflection coefficient (S_{11}). Fractal pattern design allows the antenna to have a broad bandwidth and thus it has multiple working frequency bands: S-band working frequency is from 2.0 GHz to 2.5 GHz, C-band working frequency from 7.1 GHz to 8 GHz, X-band working frequency from 8 GHz to 9.3 GHz and Ku-band working frequency from 13.8 GHz to 18 GHz. The antenna's S_{11} and VSWR parameters when matched to a 50 Ohm load are shown on Figure 3 and Figure 4, these results cover the whole S-Band and X-Band Deep-Space communication standards frequencies and broaden the possible applications for this antenna.

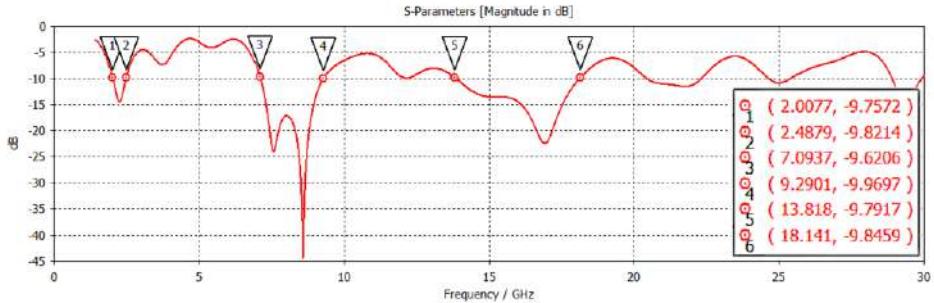


Figure 3. Antenna S11 parameters in S-Band, C-band, X-band and Ku-band (50 Ohms).

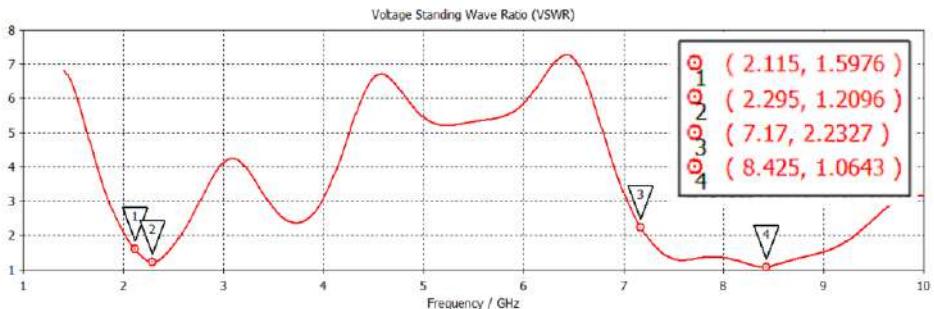


Figure 4. Voltage standing wave ratio values in S-Band, C-band and X-band (50 Ohms).

4.2. Radiation Pattern

Radiation pattern is expected to be that of an omnidirectional antenna, with its maximum along the fractal loop perpendicular axis. While effective aperture is low on S-band, due to the antenna's effective length, the antenna performs better for X-band [21]. Since Deep-space communications standards make no use of Ku-band, we'll present the antenna S, C and X bands gain radiation patterns in Figure 5 and Figure 6 respectively. X-Band uplink radiation pattern presents a 6.8 dBi gain, while C-Band downlink has the maximum gain of 4.37 dBi, this result is optimal as transmission requires more power than reception.

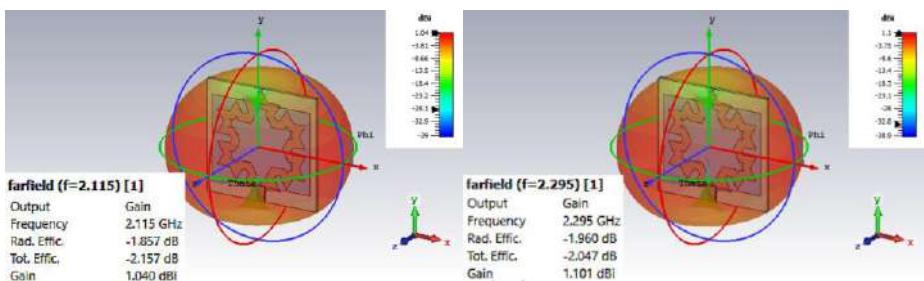


Figure 5. Deep space S-band 2110-2120 MHz up-link and 2290-2300 MHz down-link radiation patterns.

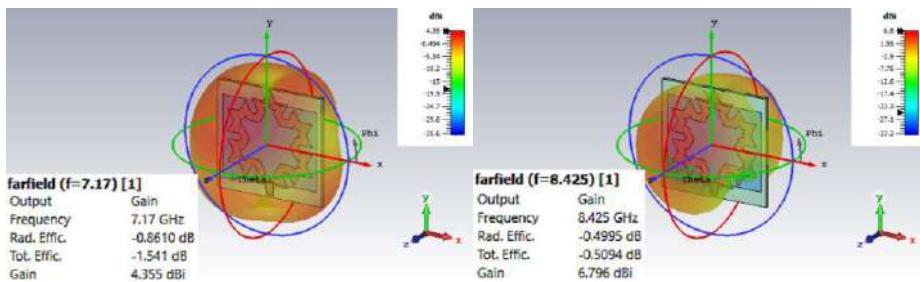


Figure 6. Deep space C-band 7145-7190 MHz up-link and X-band 8400-8450 MHz down-link radiation patterns.

Based on the previous radiation pattern figures and VSWR diagram, we can confirm the antenna optimized dimensions have been correctly tuned to the deep space bands working frequencies. The proposed antenna electromagnetic gain, S11 parameter and VSWR along each band are presented in Table 3.

Table 3. Antenna electromagnetic performance for S-band, C-Band, K-band and X-band

Center frequency (GHz)	Gain (dBi)	S11 (dB)	VSWR (-)
2.115	1.04	-12.765	1.598
2.295	1.1	-20.459	1.210
7.17	4.35	-8.374	2.233
8.425	6.8	-30.127	1.064

5. Conclusions and Future Works

The design of an antenna with a "Koch Snowflake" fractal geometry is introduced in this paper as a suitable candidate for deep space missions, as well as in various Cubesat applications which allow testing and developing technology at low cost.

The proposed antenna serves for dual-band communication since its working frequency band covers the entirety of S, C and X band up-link and down-link standard channels. The uplink gain is 6.8 dBi while the downlink gain is 4.37 dBi. The development of an omnidirectional radiation pattern allows CubeSats that use a passive magnetic attitude stabilization system to maintain a communication link with no need of an aiming system. This is ideal for low-power consumption space communication, as relay or store and forward systems, as a non-off-the-shelf alternative.

Applications such as the following could make use of this article: Deep-space missions such as asteroid exploration, investigation of the moons of Jupiter and Saturn.

We are currently working with the STK & Matlab Software for a complete simulation of the antenna constellation operation in deep space and need to further investigate the swarm system communication protocol and cooperative behavior.

References

- [1] Smith M, Douglas C, Hermann N, Mahoney E, Krezel J, McIntyre E, et al. The Artemis program: An overview of NASA's activities to return humans to the moon. AERO [Internet]. 2020 Aug, p. 1-10. Available from: <https://doi.org/10.1109/AERO47225.2020.9172323>.

- [2] Shao M, Turyshhev S, Zhai C. Big science in small packages: A constellation of cubesats to search for near earth asteroids. CLEO [Internet]. 2016 Jun. Available from: https://doi.org/10.1364/CLEO_AT.2016.JF1N.2.
- [3] Batista A, Gomez E, Qiao H, Schubert K. Constellation design of a lunar global positioning system using Cubesats and chip-scale atomic clocks. ESA [Internet]. 2012 Jul.
- [4] Kolbeck J, Fang X, Keidar M, Kang JS, Sanders M, Bakulinski N. Effects of metallic plasma thruster plume on solar cells. AERO [Internet]. 2018 Aug, p. 1-10. Available from: <https://doi.org/10.1109/AERO.2018.8396445>.
- [5] Yu Z, Yu J, Ran X. An improved koch snowflake fractal multiband antenna. PIMRC [Internet]. 2018 Feb. Available from: <https://doi.org/10.1109/PIMRC.2017.8292332>.
- [6] Swartwout M. The first one hundred CubeSats: A statistical look. JSSat. 2013 Dec; 2(2):213–233.
- [7] Toorian A, Diaz K, Lee S. The CubeSat approach to space access. AERO [Internet]. 2008 Mar. Available from: <https://doi.org/10.1109/AERO.2008.4526293>.
- [8] Stakem PH. Robotic exploration of the icy moons of the gas giants. 2020 Feb.
- [9] Sharpe BL, Schrunk DG. Malapert mountain: Gateway to the moon. ASR [Internet]. 2003 Jun. Available from: [https://doi.org/10.1016/S0273-1177\(03\)00535-0](https://doi.org/10.1016/S0273-1177(03)00535-0).
- [10] Circi C, Romagnoli D, Fumenti F. Halo orbit dynamics and properties for a lunar global positioning system design. MNRAS [Internet]. 2014 Jun. Available from: <https://doi.org/10.1093/mnras/stu1085>.
- [11] Schier J, Rush J, Vrotsos P, Williams W. Space communication architecture supporting exploration and science: plans & studies for 2010-2030. AIAA [Internet]. 2012 Dec. Available from: <https://doi.org/10.2514/6.2005-2517>.
- [12] Perea-Tamayo RG, Fuchs CM, Ergetu E, BingXuan L. Design and evaluation of a low-cost CubeSat communication relay constellation. LAMC [Internet]. 2019 Apr. Available from: <https://doi.org/10.1109/LAMC.2018.8699047>.
- [13] Bhasin K, Hackenberg A, Slywczak R, Bose P, Bergamo M, Hayden J. Lunar relay satellite network for space exploration: architecture, technologies and challenges. ICSSC [Internet]. 2006 Jun. Available from: <https://doi.org/10.2514/6.2006-5363>.
- [14] Gnawali O, Polyakovt M, Bose P, Govindan R. Data centric, position-based routing in space networks. IEEE [Internet]. 2005 Dec. Available from: <https://doi.org/10.1109/AERO.2005.1559422>.
- [15] Shin DK. Frequency and channel assignments. California Institute of Technology. 2015 Dic.
- [16] Siddiqui MG, Saroj A, Devesh, Agrawal N, Ansari JA. Design of isosceles-shaped fractal antenna for Ku and K-band applications. RAETCS [Internet]. 2018 Feb. Available from: <https://doi.org/10.1109/RAETCS.2018.8443921>.
- [17] Herrera MP, Inclan JC. Estudio y metodología de diseño de antenas utilizando geometría fractal (Antenas fractales). EPN [Internet]. 2014 Oct. Available from: <http://bibdigital.epn.edu.ec/handle/15000/11124>. Spanish
- [18] Gonzales OF, Diaz RE, Heraud JA, Correa SB. S-band Koch snowflake fractal antenna for Cubesats. IEEE [Internet]. 2016 Oct. Available from: <https://doi.org/10.1109/ANDESCON.2016.7836227>
- [19] Yassen MT, Hussan MR, Hammas HA, Al-Saedi H, Ali JK. A compact dual-band slot antenna based on Koch fractal snowflake annular ring. PIERS [Internet]. 2017 May. Available from: <https://doi.org/10.1109/PIERS.2017.8261826>.
- [20] Madray I, Kuemper R, Martin J, Madani M R. Design and COMSOL simulation of koch snowflake dipole fractal antennas. RFID-TA [Internet]. 2017 Sep. Available from: <https://doi.org/10.1109/RFID-TA.2017.8098902>.
- [21] Balanis CA. Antenna theory, analysis and design. 3rd edition. New York: John Wiley & Sons. Inc; 2015.

Searching and Rescuing Victims in Emergency: A Comprehensive Survey

Hubo Bi ^aand Erol Gelenbe ^{b,1}

^aBeijing Key Laboratory of Traffic Engineering, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China

^bInstitute of Theoretical & Applied Informatics, Polish Academy of Sciences (IITIS-PAN), Baltycka 5, 44-100 Gliwice, Poland

Abstract. The rapid progress in urban construction and informatization, especially the recent ubiquity of portable devices and intelligent infrastructures, has rapidly aggravated the concentration of pedestrians in built environments. These smart environments, which have been driven close to their limitation in daily operations, can be vulnerable during emergency situations and amplify the losses in lives and property. However, like every coin has two sides, the extra resources brought by smart equipments enlarge the optimisation space of an emergency search and rescue process and provide opportunities for new system frameworks and algorithms. In this paper, we review the current state and opportunities of the pedestrian search and rescue problem. key factors including system design, pedestrian dynamic modelling, and algorithm development are reviewed and summarised.

Keywords. Emergency management, Emergency Search and Rescue, Systemic Review

1. Introduction

As one of the most active research fields, emergency search and rescue planning, which features the widening gap between instant vast demands and restricted resources, has always been the favourite test field for cutting-edge technologies in computation, communication, and control. The basic aim of emergency search and rescue planning is to rescue victims out of dangerous areas while minimising the safety cost. However, this problem can be NP-hard due to the highly dynamic environments and complex interactive dynamics among pedestrian flows and information flows. Owing to its multidisciplinary nature, efforts have been made by scientists and sociologists in many domains to investigate this problem. In this paper, we focus on the computer-aided solutions of the emergency search and rescue planning problem.

The human history has seen a transformation in emergency management from a unorganized, reactive manner to an organised, proactive manner. Nowadays, an overall emergency management process can be divided into four aspects, emergency prevention, emergency preparedness, emergency response and emergency recovery. Emergency

¹Corresponding Author: Erol Gelenbe, Institute of Theoretical & Applied Informatics, Polish Academy of Sciences (IITIS-PAN); E-mail: gelenbe.erol@gmail.com.

search and rescue planning is the key research direction of the emergency response section. Originating from maritime search and rescue operations, search and rescue planning in emergency situations has motivated considerable research over the last several decades owing to the unfortunate increasing threat of both manmade and natural disasters. During a disaster-related emergency evacuation, evacuees may become immobilised and incapacitated due to injuries or obstacle contact. Therefore, to reduce the fatalities, various emergency management systems have been proposed to detect the location of incapacitated evacuees and dispatch rescuers to perform rescue operations. The main challenges of a rescue operation are threefold. The first challenge is how to efficiently search and locate injured evacuees or other objects in unknown environments, especially on how to coordinate the activities among various rescuers. The second is to design an appropriate rescuer assignment algorithm to allocate rescuers to injured evacuees in a real time and computationally efficient fashion under the highly dynamic hazardous environment. This is actually a NP-hard assignment problem [1], which aims to minimise the overall potential cost for the rescue operation. The third challenge is to search desired paths for rescuers and victims to fulfil their specific requirements, which is difficult since the “quality” of a path is affected by the spreading of the hazard, the dynamic congestion level, the movements and behaviours of evacuees and other rescuers on the path. In the following sections, we will summarise various systems and algorithms that have been proposed to meet the aforementioned challenges. A tree diagram of the reviewed research topics can be seen in Figure 1.

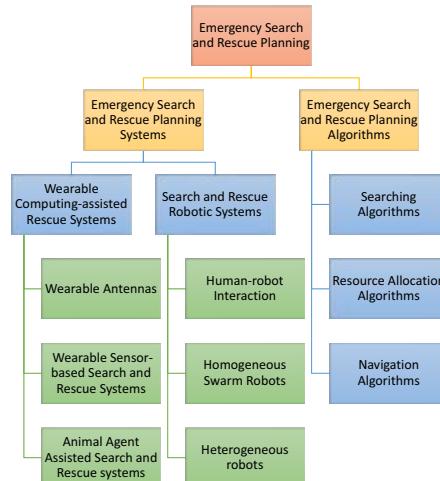


Figure 1. A tree diagram explaining the structure of the review.

2. Emergency Search and Rescue Planning Systems

Based on the facilities used, emergency search and rescue planning systems can be divided into two categories, wearable computing-assisted rescue systems and search & rescue robotic systems. Wearable computing-assisted rescue systems concentrate on providing

ing various enhancements for emergency personnel with the aid of wearable devices to increase the efficiency of rescue operations and improve the safety of emergency responders. On the other hand, disasters that create harsh environments with extreme temperature, toxic substances or various obstacles have exposed the unsafety and inefficiency of the human-centered search and rescue planning systems. These limitations have therefore inspired the development of the robot-centered search and rescue planning systems, which employ various mobile robots to conduct rescue operations.

2.1. *Wearable Computing-assisted Rescue Systems*

With the rising interest in body-centric wireless communications, which has also been standardised as a part of the fourth generation mobile communication systems (4G), considerable research has been dedicated to develop low cost, light-weight wearable antennas to maintain and improve emergency communications. The work in [2] implements an electrically-small wearable antenna, which is integrated into the sleeve of a jacket, to monitor the positions of emergency rescuers. The research in [3] presents an emergency rescue navigation system to direct firemen to “key corridors” to eliminate fire and congestion caused by evacuees or obstacles generated from hazard with the aid of personal digital assistants (PDAs) carried by firefighters. As an effort to decrease the likelihood of emergency events from the “prevention” aspect, the study in [4] presents a wearable personal healthcare and emergency aid system, namely “WAITER”, to monitor the health status of users with wearable vital signal sensors and alert the remote healthcare center when an emergency occurs.

Owing to the superior performance of animal agents during search and rescue operations in terms of mobility, energy utilisation efficiency, sensory acuity and intrinsic cognitive capacity, much effort has been dedicated to animal agent assisted search and rescue systems. The work in [5] employs cyber-enhanced working dogs to locate and reach survivors trapped under rubble in the aftermath of large-scale disasters; this system contains three components: a smart harness worn by a working dog to monitor the surrounding environment and the canine, a remote computer carried by the handler to analyse and control the behaviours of the canine, and mobile base stations such as unmanned vehicles to maintain the communication between the working dog and the handler.

2.2. *Search and Rescue Robotic Systems*

Since the human access to the victims in the aftermath of a disaster such as an earthquake or radioactive leakage can be time consuming and may induce further casualties, the use of robots that can be released rapidly to locate and rescue victims has drawn considerable attention in recent decades, and has gradually become a research domain after the Great Hanshin earthquake and the Oklahoma City bombing in 1995. Nowadays, research on “search and rescue robotic systems” has evolved as a major branch of “search and rescue”, especially for urban scale structural collapse environments. According to the different operating environments, search and rescue robots can be divided into autonomous underwater vehicles (AUVs), and urban search and rescue (USAR) vehicles, which contain various types of unmanned ground vehicles (UGVs) and unmanned aerial vehicles (UAVs). This literature review focuses on the research of USAR robotic systems, and the research on AUVs is excluded due to its distinctive emphases caused by the turbid underwater environment.

Due to the limitations on the mobility (tracks, wheels or combination of both) and intelligence, most of the current search and rescue robot systems are not advanced enough for fully autonomous operations. Hence, a research direction is to investigate the human-robot interaction during search and rescue operations. The first known use of USAR robots for an actual unstaged rescue mission occurs during the World Trade Center collapse, when various models of tele-operated robots from industry, military and academia were invited to assist the rescue operation.

One of the major challenges that teleoperated or semi-autonomous rescue robots face today is the high human-to-robot ratio (The human-to-robot ratio denotes the number of people needed to operate a robot), which reaches 2 : 1 or even 3 : 1 in the current state of practice. The high human-to-robot ratio restricts the scale of rescue robots to be deployed in the devastated region since too many operators can significantly increase the logistic burden and the training cost. One feasible solution to this problem is the use of homogeneous swarm robots, which are easier to control and understand. For instance, the work in [6] employs a team of homogeneous robots to search and explore several disaster sites; because of the possible environmental changes and robot failures in disaster sites, the robots are required to continuously redistribute to fulfill the desired population fraction of robots at each site. Inspired by the self-organised behavior of social insects, the study in [7] proposes a stochastic task allocation strategy to assign a swarm of homogeneous robots to various tasks during an search and rescue operation without inter-robot communications.

Owing to the diverse and complex tasks in rescue operations, most of the recently proposed USAR robot systems consist of heterogeneous robots with different capabilities instead of homogeneous robots. For instance, the study in [8] proposes a robot group which consists of three types of robots to rescue victims during nuclear plant accidents. The work in [9] presents a USAR robot system which contains aerial robots and a land robot to detect the potential survivors in high-destruction locations after an earthquake.

In addition, to verify the effectiveness of the proposed robotic systems, various real and simulated scenarios have been designed in the last two decades. The work in [10] designs a test arena with collapsed structures to evaluate the agility, planning or mapping algorithms and sensing ability of rescue robots. To mimic large scale disasters, a proving ground namely “Disaster City” with a size of $210,000m^2$ has been built by the US Federal Emergency Agency (FEMA) to conduct rescue drills for robots [11].

3. Emergency Search and Rescue Planning Algorithms

The state-of-the-art emergency search and rescue planning studies are originated and derived from research in maritime search and rescue activities, which has been a significant research topic since 1970s. Related algorithms in emergency search and rescue planning are commonly under the assumption that multiple utilities (rescuers, robots, etc.) are involved. The reason behind this is twofold. First, time efficiency is the dominant factor for the success of a search and rescue operation, and multi-agent based search and rescue operations are obviously more efficient than single-agent based search and rescue operations. Second, multi-agent systems are generally more cost-efficient and feasible than a single agent with all the capabilities [12]. In recent decades, related algorithms have aroused a new interest in not only searching victims in hazardous environments,

but also planetary exploration and de-mining [13]. The philosophy behind team-based search and rescue is to convert a complex problem into simpler sub-problems that are more efficient to solve. These algorithms are be classified into three types: first, searching algorithms, which are dedicated to search and locate injured evacuees in unknown environments, specifically concentrating on coordinating the search behaviours of rescuers to produce “swarm intelligence”; second, resource allocation algorithms, which aim to assign rescuers to injured evacuees in a desired manner when the locations of evacuees are revealed; third, navigation algorithms, which focus on discovering appropriate paths for rescuers when the locations of evacuees are known. In this section, the third category is excluded since all the on-line emergency navigation algorithms are applicable to this case, and the detailed review can be found in [14].

3.1. *Searching Algorithms in Emergency Search and Rescue Planning*

The work in [15] proposes a particle swarm optimisation (PSO) based multi-robot system to find targets by gradually optimising the pre-defined goal function; each robot is modeled as a particle whose velocity and position is determined by the neighbouring particles and the previous best position of this particle. Similarly, the study in [16] designs a PSO based multi-robot system to search targets and avoid obstacles in an unknown environment; specially, a relative coordinate system is used to avoid the dependence of the precise global location of robots. By employing a bio-inspired random search behaviour called lévy flight, the work in [17] presents an efficient multi-robot searching algorithm to search targets in an unknown environment.

3.2. *Resource Allocation Algorithms in Emergency Search and Rescue Planning*

The work in [1] presents a resource allocation algorithm on the basis of the Random Neural Network (RNN) with synchronised interactions [18] to assign rescuers to trapped victims in a fire-affected building; to optimise this task assignment process, several effect factors such as the cost of assigning a rescuer to a victim (e.g. the distance between them), the probability that a rescuer fails to save a victim, and the associated penalty that the rescuer fails to rescue the victim are taken into consideration, and are transformed into a goal function which reflects the trade-off between the total successful rescue costs and the failure penalties; the RNN is utilised as a fast optimisation algorithm to minimise the goal function with a gradient descent learning procedure; in the RNN, each possible rescuer-victim pair is considered as a neuron, and the neuron with the highest excitation probability after the training process is selected as the decision. The study in [19] proposes a task allocation algorithm for heterogeneous agents in hazard environments with respect to time, space and communication restrictions. As aforementioned in section 2.2, the work in [7] employs an ordinary differential equation model to redistribute rescue robots among various tasks in a decentralised manner without inter-robot communications.

4. Conclusion

In this paper, we provide a systemic review to challenges and solutions of the emergency search and rescue problem, we first review the history and evolution of this research area,

followed by the various frameworks of rescue systems along with different pedestrian models used in mimicking the movement patterns of a search or rescue process. After that, we summarise representative algorithms used to rescue victims. The work is a major supplement to our previous work in [14], they each cover one of the two main branches of the studies in emergency management.

References

- [1] Gelenbe E, Han Q. Near-Optimal Emergency Evacuation with Rescuer Allocation. In: In Proceedings of the 4th International Workshop on Pervasive Networks for Emergency Management (PerNEM'14). IEEE; 2014. p. 1-6.
- [2] Orefice M, Pirinoli P, Dassano G. Electrically-small wearable antennas for emergency services applications. In: Antenna Technology (iWAT), 2016 International Workshop on. IEEE; 2016. p. 131-4.
- [3] Li S, Zhan A, Wu X, Chen G. ERN: Emergence rescue navigation with wireless sensor networks. In: Parallel and Distributed Systems (ICPADS), 2009 15th International Conference on. IEEE; 2009. p. 361-8.
- [4] Wu W, Cao J, Zheng Y, Zheng YP. WAITER: A Wearable Personal Healthcare and Emergency Aid System. In: 2008 Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom); 2008. p. 680-5.
- [5] Bozkurt A, Roberts DL, Sherman BL, Brugarolas R, Mealin S, Majikes J, et al. Toward cyber-enhanced working dogs for search and rescue. IEEE Intelligent Systems. 2014;29(6):32-9.
- [6] Halász A, Hsieh MA, Berman S, Kumar V. Dynamic redistribution of a swarm of robots among multiple sites. In: Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on. IEEE; 2007. p. 2320-5.
- [7] Berman S, Halász Á, Hsieh MA, Kumar V. Optimized stochastic policies for task allocation in swarms of robots. IEEE Transactions on Robotics. 2009;25(4):927-37.
- [8] Iwano Y, Osuka K, Amano H. Proposal of a Rescue Robot System in Nuclear-Power Plants -Rescue Activity via Small Vehicle Robots -. In: Robotics and Biomimetics, 2004. ROBIO 2004. IEEE International Conference on; 2004. p. 227-32.
- [9] Ventura R, Lima PU. Search and rescue robots: The civil protection teams of the future. In: Emerging Security Technologies (EST), 2012 Third International Conference on. IEEE; 2012. p. 12-9.
- [10] Jacoff A, Messina E, Weiss B, Tadokoro S, Nakagawa Y, et al. Test arenas and performance metrics for urban search and rescue robots. In: Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on, vol. 4. IEEE; 2003. p. 3396-403.
- [11] Ventura R, Lima PU. Search and Rescue Robots: The Civil Protection Teams of the Future. In: Emerging Security Technologies (EST), 2012 Third International Conference on; 2012. p. 12-9.
- [12] Gautam A, Mohan S. A review of research in multi-robot systems. In: Industrial and Information Systems (ICIIS), 2012 7th IEEE International Conference on. IEEE; 2012. p. 1-5.
- [13] Gelenbe E, Cao Y. Autonomous search for mines. European Journal of Operational Research. 1998;108(2):319-33.
- [14] Bi H, Gelenbe E. A survey of algorithms and systems for evacuating people in confined spaces. Electronics. 2019;8(6):711.
- [15] Pugh J, Martinoli A. Inspiring and Modeling Multi-Robot Search with Particle Swarm Optimization. In: Swarm Intelligence Symposium, 2007. SIS 2007. IEEE; 2007. p. 332-9.
- [16] Zhu Q, Liang A, Guan H. A PSO-inspired multi-robot search algorithm independent of global information. In: Swarm Intelligence (SIS), 2011 IEEE Symposium on. IEEE; 2011. p. 1-7.
- [17] Sutantyo DK, Kernbach S, Levi P, Nepomnyashchikh V, et al. Multi-robot searching algorithm using lévy flight and artificial potential field. In: Safety Security and Rescue Robotics (SSRR), 2010 IEEE International Workshop on. IEEE; 2010. p. 1-6.
- [18] Gelenbe E, Timotheou S. Random neural networks with synchronized interactions. Neural Computation. 2008;20(9):2308-24.
- [19] Su X, Zhang M, Bai Q. Dynamic task allocation for heterogeneous agents in disaster environments under time, space and communication constraints. The Computer Journal. 2015;58(8):1776-91.

Design of Bluetooth Automatic Check in System Without Operation Based on Cross Mobile Platform

Yiqin BAO^{a1} and Zhengtang SUN^b

^aCollege of Information Engineering, Nanjing XiaoZhuang University, China

^bSchool of Computer Science and Technology, Heilongjiang University, China

Abstract. Because the traditional way of sign in is not to call the roll, or sign in on the paper list, or sign in through fingerprint recognition and face recognition, on the one hand, it is cumbersome and time-consuming, on the other hand, it is non-human. The update of roll call sign in is a hot topic at present, and many new schemes rush out. This paper introduces the automatic sign in system based on Bluetooth technology, which does not rely on the mobile phone operating system, does not need the mobile phone installation software, as long as you bring a mobile phone, you can realize automatic sign in. This paper designs and implements the automatic check-in system, and compares it with other ways, proving that it is an innovative check-in method.

Keywords. Automatic registration, bluetooth technology, JSON, SQLite

1. Introduction

At present, there are a total of more than 2000 colleges and universities in China. Many of them need to analyse and manage the data of the students who are always absent, so they need to count. Therefore, the attendance in a college is very important for the quality management. Now most of the colleges and universities have no difference with the past in the way of signing in, which is still a name called and the hands up, very troublesome and time-consuming. Sometimes teachers can according to their own experience, no teacher can do it in every class in usual, which leads a lot of students to produce fluky psychology. This will undoubtedly have a certain impact on the quality of teaching. To solve this problem, it is high time to change the traditional tedious way. There are a lot of new ways such as using a credit card, face recognition, fingerprint recognition, but either inconvenient, or high cost, most importantly is not humanized, so I need a new humanized way.

In recent years, many related technologies have been applied to check-in and roll call. Zhang et al. have studied the experimental check-in system based on RFID [1], Wang et al. have studied the mobile phone check-in system based on wireless router [2], and the roll call system based on face recognition [3-5]. However, due to these methods, on the one hand, the operation is cumbersome and time-consuming, on the other hand, it is not humanized and other weaknesses, so we hope to have a better way of roll call. Due to the popularity of mobile phones and the innovation of Bluetooth technology, it is feasible to use mobile phones to realize roll call through Bluetooth.

¹ Corresponding Author: Yiqin Bao; E-mail: baoyiqin@njxzc.edu.cn

2. Bluetooth Technology

Bluetooth technology[6-7] is a kind of radio technology which supports equipment with short-range (general 10m) communications. It can realize wireless information exchange between mobile phone and other equipment. Through the Bluetooth radio, it can be very convenient to get the device name, MAC address, and even the signal strength. Nowadays, the smartphones are all equipped with Bluetooth function, and have unified standard and interface.

3. The System Overall Structure

Automatic sign-in system based on Bluetooth technology, mainly consists of the server, the APP installed on teachers' mobile phone, the students' mobile phone with bluetooth opened, Bluetooth network and the Internet, as shown in Figure 1. At the first time, students should change the mobile phone name (with his own name), open the Bluetooth to release the radio signals; The server is responsible for inputting and management of student information; APP is responsible for carrying interface for the teachers' mobile phone connected to the server, and the receiving and processing of Bluetooth signal. When the teacher wants to call the roll, they need to download the student list from the server. Then choose class number to enter the main interface. Click start after, mobile phone to receive Bluetooth broadcast signal nearby. At the same time, students are asked to change the name of their mobile phone name and open Bluetooth. When the teacher gets one mobile phone Bluetooth broadcast, it automatically gets the name and MAC address. If there is a match between the device name and the one in the list. The corresponding field in the form will change. At the end, data can be submitted to the server when the 'submit' button is clicked.

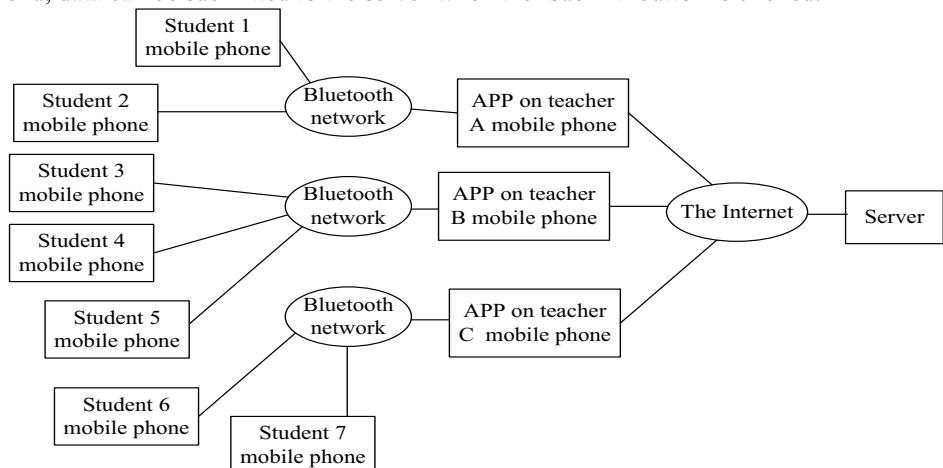


Figure 1. System architecture diagram

4. Design of the Server

4.1 Student Management Module

Student management module is responsible for the management of student information, not only can choose class, query, add, modify and delete student information, but can also import student information directly from XLS documents.

4.2 Interaction with the APP

In the implementation of the interaction between server and mobile phone terminal, the main communicating method is adapting HTTP wireless communication, using a kind of lightweight data interchange format, JSON. JSON format has clear and concise hierarchical structure which makes it easy to read and write, as well as easy for machine to parse and generate. The transmission efficiency is high.

The Servlet in server receives the requests APP sends, make a series of responses, including the resolution of the JSON statement, extracting the corresponding parameters, the operation of the database, when necessary, to call `JSONObject.put` (`Object key, Object value`) method to create a JSON statement. In the system, the statement mainly includes some attributes of students returned to the mobile phone terminal.

5. Design of the Server

The homepage of the APP contains class choose button, 'download' button, 'delete' button and 'uri configuration' button, as shown in Figure 2. Before the teachers use the APP function to call the roll automatically, they need to set the IP address and port number. Then let the students open the phone Bluetooth and changed its names to their own names. At the same time click on the "start" button. Thus the sign-in process will be completed automatically.



Figure 2. The homepage of the APP

5.1 Uri Configuration

Because the data transmission using the HTTP protocol, teachers need to set the server URI (Uniform Resource Identifier) first, IP address and PORT number are needed. In the system, once submitted, the data will be stored in the Shared Preferences, thus there is no need to set the IP and PORT each time. The configuration interface is shown in Figure 3.



Figure 3. Configuration Interface

5.2 Download the List

When the curriculum is started or the list is changed, the teacher need download the list of the corresponding class first. In this process, the app gets connection to the server by the URI and transfer parameters. Then it will obtain the JSON statement returned, analyse it into student objects and store them in SQLite database.

5.3 Choose Class

Teacher taught more than one classes, while the list to download contains all the students list, so it needs to realize the class selection function in APP. When designing student information database, the class number is included as a property. Therefore, the students can be divided into the corresponding class, which is the base of the system.

5.4 Call the Roll

After entering the class interface, a table will be displayed, including student number, name, and state (initially empty) of all the students in the class. A BroadCastReceiver is set in the APP MainActivity to receive Bluetooth broadcast sent by the devices around. When receiving a Bluetooth broadcast signal, it will get the name and match it with the list of the class. If matched, a 'yes' will be written into the 'State' column. Thus, we complete 'calling the roll' automatically.

In the design of Bluetooth technology, the Bluetooth radio receiving equipment on the natural need to use Bluetooth devices, so we must first add Bluetooth permissions in the AndroidManifest.xml file, and then define the Bluetooth adapter access to access the native Bluetooth device. If bluetooth devices is discovered in after scanning, it will instantiate the discovered device, and use the object to get its name and other required attributes.

Bluetooth permissions settings code as follows:

```
<uses-permission android:name="android.permission.BLUETOOTH" />
<uses-permission android:name="android.permission.BLUETOOTH_ADMIN" />
```

Bluetooth adapter instantiation code as follows:

```
private BluetoothAdapter mBluetoothAdapter;
mBluetoothAdapter = BluetoothAdapter.getDefaultAdapter();
if(mBluetoothAdapter !=null){
    if(!mBluetoothAdapter.isEnabled()){
        Intent intent = new Intent(BluetoothAdapter.ACTION_REQUEST_ENABLE);
        startActivityForResult(intent);
    }
    mBluetoothAdapter.startDiscovery();
}
else{
    textView.setText("Bluetooth device is no found! ");
}
if (action.equals(BluetoothDevice.ACTION_FOUND)) {
    BluetoothDevice
    device=intent.getParcelableExtra(BluetoothDevice.EXTRA_DEVICE);
    mTextView.append(device.getName() + " : " + device.getAddress()+"\n");
}
```

After clicking on the 'START' button, when discovering a matching device, corresponding fields in the table (and time) will be modified and saved in the SQLite, the name and address of all Bluetooth devices will also display under the table, as shown in Figure 4. Teachers can also clear the data for the next time to use it.

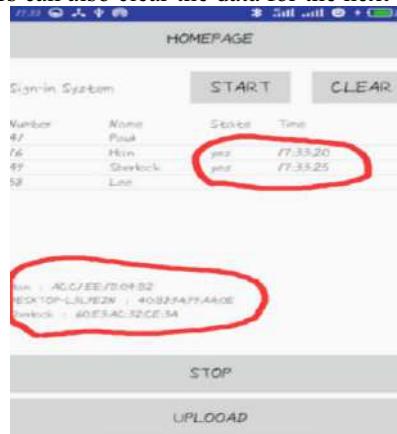


Figure 4. Automatic sign in interface

5.5 Upload

The way of uploading in this system is to get the students list from the SQLite to generate JSON format statement and transmit it to the server by HTTP protocol. The server will modify the database according to the JSON statement in, thus complete the uploading process.

6. The Comparison with Several Other Ways

In order to illustrate the feasibility of the sign in system based on Bluetooth adopted in this paper, we compare RFID swipe card sign in, wireless routing sign in, face recognition sign in, fingerprint recognition sign in, etc. the comparison results are shown in Table 1. The results show that the check-in method used in this paper has better advantages.

Table 1. Comparison of different kinds of sign-in modes

Solution	Advantages	Disadvantages
Credit card sign-in	Faster recognition speed and higher stability	The cost of the card is high and it might get damaged and lost easily. Data access can be affected by the environment.
Face recognition	No need of users' operation	The recognition speed is slow and the equipments are expensive.
Mobile phone sign-in based on wireless router	The step is simple, the operation is few and the accuracy is high.	Need to install WiFi to sign in, and there is a limit to the number of devices wireless router can connect; Frequency bands of WiFi generated by wireless router used to sign in are close to each other, which will interfere with each other, affecting the public network
Fingerprint recognition	Medium speed of recognition.	The equipments need maintenance and the cost is high.
Sign-in system based on Bluetooth	It's convenient and in addition to mobile phone without additional; user-friendly	Students need to modify the name of mobile phone with that of himself(herself).

7. Conclusions

This paper realizes the call sign based on Bluetooth technology and mobile Internet technology. It is a system of innovation with great advantages including clear train of thought, simple method and low cost. Due to its own factors, Bluetooth 4.0 has a range of just 10m. Sometimes the classroom is too large for the mobile phone to receive broadcast signals in time and the limitation will come into being. But with the development of Bluetooth 5.0 and the continuous progress of technology, the scope of application of Bluetooth technology will continue to expand, sign-in system based on Bluetooth will have great prospects for development.

Acknowledgement

This work is supported by Key topics of the '13th five-year plan' for Education Science in Jiangsu Province (B-b /2020/01/18).

References

- [1] Zhang Liang, Lingyun Feng, Wang Wen, Bi Tao. Design and implementation of experiment attendance system based on RFID [J]. Modern electronic technology, 2015 (1).
- [2] Wang Weihong, Chen Xiaoliang. Design and implementation of mobile phone system based on wireless router[J]. Journal of Zhejiang University of Technology, 2016 (6).
- [3] Jing Zhou, Quan Wei, Tang Jie, Yan Han. Sign-up system based on face recognition[J]. Software engineering, 2017 (5).
- [4] ZHUO Ya-qian, OU Bo. Face Anti-spoofing Algorithm for Noisy Environment[J]. Computer Science, 2021, 48(6A): 443-447.
- [5] WU Qing-hong, GAO Xiao-dong. Face Recognition in Non-ideal Environment Based on Sparse Representation and Support Vector Machine[J]. Computer Science, 2020, 47(6): 121-125.
- [6] Chen Xi, Xu Peng, Xu Xiaolei. Development of embedded industrial control system based on Bluetooth communication technology [J]. Instrument technology and sensor, 2015 (1): 74-76.
- [7] HONG Pengcheng, HUANG Yicai, YU Bin. Research on Bluetooth Encrypted Speech Synchronization for Mobile Communication Network. CEA, 2020, 56(13): 131-136.

Exploration and Research on Integrating Programming Education into Junior Middle School Mathematics Classroom

Yiqin BAO^{a1}, Gui WANG^a and Zhengtang SUN^b

^aCollege of Information Engineering, Nanjing XiaoZhuang University, China

^bSchool of Computer Science and Technology, Heilongjiang University, China

Abstract. We are in the era of big data, and it is particularly important to cultivate students' data processing ability. Python is an important auxiliary tool for data analysis, and its syntax is concise and clear, which is suitable for junior high school students to start learning code language. Making full use of modern information technology and Innovating Curriculum form is the new requirement of "China education modernization 2035". Integrating Python into junior high school mathematics classroom is not only the innovation of mathematics curriculum form, but also a good way to cultivate students' thinking ability and programming thought. Based on the understanding of Python software and data processing plate in junior high school mathematics curriculum, this paper analyzes the promotion of students' thinking ability by integrating Python into junior high school mathematics classroom, and explores the strategies of applying Python in teaching combined with cases.

Keywords. Python, Junior high school mathematics, data processing

1. Introduction

With the advent of the era of big data and artificial intelligence, as an important way to cultivate students' digital skills and programming thinking, programming education has become the starting point of educational reform and deep cultivation [1]. In recent years, various researches on the development of students' programming thinking have been carried out continuously and integrated into the school classroom, which has become the mainstream trend of the development of students' Programming Education [2]. Li et al. [3] engaged in experimental teaching research on Mu class. Tsai et al. [4] improved students' interest in learning through flipping class. He et al. [5] conducted research on teaching mode exploration and practice. However, Junior middle school mathematics teaching focuses on cultivating students' thinking ability, and most of these studies are the research of teaching mode. Therefore, it is worth exploring and studying to integrate programming education into mathematics teaching and python into junior middle school mathematics curriculum.

This paper first introduces the characteristics of programming language, then studies the role of programming language Python in enhancing thinking ability, adding teaching interest and expanding students' vision, and finally summarizes the full text.

¹ Corresponding Author: Yiqin Bao; E-mail: baoyiqin@njxzc.edu.cn

2. Powerful programming language-Python

Python is a popular programming language in recent years. It was developed by Guido van Rossum at the end of 1989. As a free and open source cross platform advanced dynamic programming language, Python supports command programming and functional programming, and fully supports object-oriented programming. Its syntax is simple and clear, and it provides a large number of built-in objects and functions, becoming a pattern is very consistent with human thinking methods and habits [6].

The features of Python programming language are as follows:

1) Powerful function and simple operation: it has a very wide range of applications. At present, it has penetrated into almost all majors and fields such as computer science and technology, statistical analysis, computational visualization, mobile terminal development, graphics and image processing, artificial intelligence, etc. [7].

2) Data analysis database has complete functions: its own database has complete functions, such as NumPy and SciPy for the development of computer science industry, Django for the development of computer web pages, etc. [8]. These databases include many typical data models, which are easy to use directly when processing data and provide work efficiency [9].

3) Cross platform operation is simple and low cost: at this stage, Python can be well compatible with Linux, Windows and other platforms, and can run well [10].

4) Meet the preferences of the audience: in many programming languages, Python stands out with its own advantages and becomes one of the most commonly used programming software for software designer.

The future era is the era of AI, and Python is the closest language to AI. Therefore, Python language course will become a trend of children's learning. At present, Python language is gradually replacing VB language into middle school information technology curriculum. The primary problem of Python Programming Teaching in junior high school is how to stimulate students' interest in learning, and then integrate with mathematics learning to help students improve their thinking ability, and finally achieve the training goal of students' comprehensive quality [11-12].

3. Introducing Python - Enhance thinking ability

Programming is a kind of activity to investigate students' thinking comprehensively. Bill Gates once said: "learning programming can exercise your thinking, help you think better, and create a way of thinking that I think is very useful in all fields". In the process of teaching, the introduction of Python language and the creation of teaching situation will become an effective way to realize the modernization of mathematics education. This can not only help students consolidate their mathematical knowledge and optimize their knowledge structure, but also stimulate their thinking ability and help them master the correct thinking process and method, so as to cultivate their thinking ability of image, divergence and intuition.

3.1 The influence of Python on the development of mathematical thinking

Mathematical thinking depends on solid basic knowledge and skills, which is the thinking activity form of mathematical thinking and problem solving. Junior high school students' thinking gradually from "experience type" to "theory type", which is the key period to cultivate students' mathematical thinking. At this moment, teachers

should grasp the critical period and correctly guide students to promote their better development.

In junior high school mathematics curriculum, students have begun to contact with simple data processing knowledge - "data around us", "data concentration", "data dispersion", "simple application of Statistics" and other mathematical knowledge plates are reflected. During this period of study, students initially form the ability of data processing, but these abilities only stay at the theoretical level. Therefore, optimizing students' existing mathematical cognitive structure is the premise of cultivating mathematical thinking. In the "future education" system, curriculum and teaching are the most important carrier of education. With the help of python, it will break the traditional course mode and enrich the classroom teaching content.

In the teaching of "3.4 variance" in the ninth grade mathematics volume I of Jiangsu Education Press, understanding the process of the generation and formation of the concept of variance is the key and difficult point of students' learning, and it is also the place where students are easily confused. Teachers in the classroom into Python can be a good solution to this problem. With the help of Python to draw the image of data, through the comparison of different groups of data graphs, we can feel the discrete degree of data in the gap between each group of data and its average value, so as to deeply understand the concept of variance and master mathematical knowledge. In this process, students actively participate in thinking and solving problems, which will gradually improve their mathematical thinking ability.

3.2 The influence of Python on the development of logical thinking

Logical thinking ability is a person's ability to think correctly and reasonably, that is, the ability to observe, compare, analyze, synthesize, abstract, generalize, judge and reason things, and the ability to use scientific logical methods to express his thinking process accurately and orderly. Relatively speaking, logical thinking ability can be more reflected in mathematics and other theoretical disciplines.

With the reform of the new curriculum, people attach great importance to the cultivation of students' logical thinking ability. Therefore, the cultivation of students' logical thinking ability has gradually become the top priority of teachers' Teaching. In the process of teaching, if teachers can take programming education to let students contact with programming education and gradually learn logical thinking, it will have a huge impact on students' lifelong development and social development, as well as on promoting the teaching progress of this subject.

Programming for logic rigor, thinking creativity is very high, its core is algorithm, and the ability of logical thinking in mathematics is very important. If Python is added to the teaching, students can gradually develop their logical thinking ability in code writing. For example, when "seeking mode" is implemented with Python, we need to understand the concept of mode and its formation process, that is, to accumulate the most frequent non repeated values in the data, then we can write correct and beautiful code.

3.3 The influence of Python on the development of programming thinking

With the rapid development of the times and driven by the trends of informatization, intelligence and integration, the cultivation of programming thinking becomes clear and important, and the core of programming thinking is computational thinking, that is, "understanding problems - the thinking process of finding paths". Programming ideas are influenced by computer language and computer operating environment to a certain

extent. Therefore, programming thinking is not a commonly recognized programming skill, but an efficient way of thinking to solve any problem.

The new generation of artificial intelligence development plan issued in 2017 in China pointed out that artificial intelligence related courses should be set up in primary and secondary schools, and programming education should be gradually promoted. Python language as the most in line with the way of human thinking and habits of programming language into the school classroom is a very meaningful move, plays a directional role, bearing the goal of education.

At present, the computer has gradually become a pan social form of existence that affects the way of thinking and ability of human beings. In the junior high school mathematics classroom integrated with Python, students constantly use programming language in the process of teacher guidance, self practice and team exploration, so that children can imperceptibly cultivate programming thinking and solve problems more efficiently.

4. Python into the classroom -- add fun and broaden horizons

4.1 Using Python to design statistical chart

In junior high school mathematics data processing problems, Jiangsu Education Press seventh grade mathematics volume II first began to contact the knowledge of frequency, the most common topic is to calculate the frequency distribution table and draw the frequency distribution map according to the data. If we learn Python, we can better grasp the situation of the data, at the same time, we can check the answers of the written calculation.

Example 1: input a string, convert it into uppercase, count the times of each letter, and draw the statistical histogram.

1) The Python program is as follows:

```
a = input("Please enter character string:")
b = list(a.upper())
c = {}
for i in b:
    c[i] = b.count(i)
x = list(c.keys())
y = list(c.values())
#Draw histogram
plt.bar(np.arange(0, len(x), 1), height=y)
plt.xticks(np.arange(0, len(x), 1), x)
plt.xlabel("Upper case letters")
plt.ylabel("Numbers")
plt.title('Statistical histogram')
plt.show()
```

2) The results are as follows

If input: ABACsDEFcJGcFScwSJfLwPEWRFZZ

Then count the number of times each letter appears: {'A': 2, 'B': 1, 'C': 4, 'S': 3, 'D': 1, 'E': 2, 'F': 4, 'J': 2, 'G': 1, 'W': 3, 'L': 1, 'P': 1, 'R': 1, 'Z': 2}

The statistical histogram is as follows:

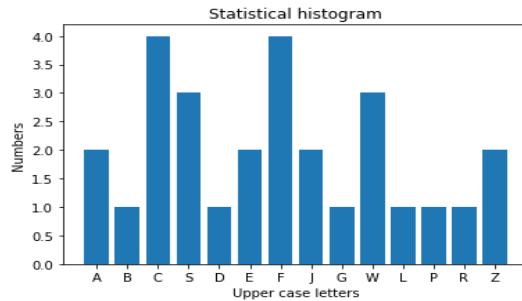


Figure 1. Histogram of letter occurrence times

4.2 Using Python to solve statistics

Mathematics teaching should be based on a solid foundation and impart knowledge through effective mathematics teaching methods. When junior high school students have a certain basic knowledge of probability, they will start to contact the knowledge of data, introduce basic statistics such as mean, median and variance, and require students to be able to solve them simply. If we use Python to help them learn with a good foundation of written calculation, it will stimulate students' desire to explore and improve their learning efficiency.

Example 2: Xiao Ming opened a grocery store. The following is his profit record in the past ten days (yuan): [85,98,76,98,126,138,100,98,106,117]. Request the average, standard deviation and median of this group.

1) The Python program is as follows:

```
#Raw Data
data = [85,98,76,98,126,138,100,98,106,117]
#Calculate average value
average = float(sum(data))/len(data)
#Calculate Standard deviation
total = 0
for value in data:
    total += (value - average) ** 2
stddev = math.sqrt(total/len(data))
#Calculate Median
data.sort()
mid = int(len(data) / 2)
if len(data) % 2 == 0:
    median = (data[mid - 1] + data[mid]) / 2.0
else:
    median = data[mid]
print(average, stddev, median)
```

2) The results are as follows

Average: 98, Standard deviation: 17.55448660599335, Median: 99.0

In the above solving process, we use Python to demonstrate examples and solve problems, stimulate students' interest in learning, strengthen students' understanding of the principle of statistics solving, so as to master how to use these statistics to solve problems in combination with practice.

5. Conclusions

In the teaching process of junior high school mathematics data processing plate, students may have a certain interest in contacting new knowledge at the initial stage, teachers can explain some examples and lead students to practice, which will stimulate students' interest in learning. This can strengthen students' understanding of the principle of statistical solution in the code writing, so as to achieve the effect of learning knowledge transfer, and then cultivate students' thinking ability.

Students can creatively use data information and data knowledge to solve problems, and find the beauty of mathematics in life in this process, so as to further improve students' interest in mathematics, and then potentially promote moral and moral education Attitude, emotion and other goals.

Acknowledgement

This work is supported by Key topics of the '13th five-year plan' for Education Science in Jiangsu Province (B-b /2020/01/18).

References

- [1] Zhang Jinbao, Ji Lingyan. The evolution of the orientation of information technology education in primary and secondary schools [J]. *Audio visual education research*, 2018 (5): 108-114.
- [2] Sun Lihui, Liu Siyuan, Li Manman. Action path of children's programming education in the era of artificial intelligence -- Based on the survey report of "necessary conditions for the development of children's programming education" in Japan [J]. *Audio visual education research*, 2019,40 (08): 114-120 + 128.
- [3] Xianyi Li, Senlin Tong, Xiaoying Liu. Experimental teaching preview system based on Mu class concept. *Laboratory science*, 2017, 20 (5): 161-163.
- [4] Tina Pingting Tsai, Jyhjong Lin, Chingsheng Hsu, "Using ePUB3 eBook-based SNS to Enhance Students ' Learning Effects in Flipped Classes," *Journal of Internet Technology*, 2020(12),21(7):1999-2008.
- [5] HE Shi-tian,TAN De-li.Exploration and practice of "principles of computer composition" teaching mode based on synergy inside and outside class[J]. *Computer Engineering & Science*. 2019(12),41(1):144-148.
- [6] Tao Wenling, Hou Dongqing. Discussion on Python application based on problem solving -- Taking Python batch processing excel file as an example [J]. *Journal of Ezhou University*, 2019,26 (01): 107-109 + 112.
- [7] Xiao Le, Cong Tianwei, Yan Wei. Web big data collection and data analysis based on Python [J]. *Computer knowledge and technology*, 2018,14 (22): 9-11.
- [8] Wang Liang, Zuo Wentao. Research on the application of Python programming language in big data collection and analysis [J]. *Computer products and circulation*, 2020 (01): 22.
- [9] Xu Dongsheng, Chen Hongxia. Research on using Python to promote agricultural product information efficiency and application willingness under the background of big data [J]. *Rural economy*, 2018 (05): 99-103.
- [10] Xia huosong, pan Xiaoting. Research on the relationship between big data academic research and talent demand based on Python mining [J]. *Journal of information resource management*, 2017,7 (01): 4-12.
- [11] Chen Xiaojuan. Learning by doing: a case study of Python program [J]. *China educational technology equipment*, 2019 (03): 98-99 + 102.
- [12] Qiu Yu, Shang Zhi Nian. Practice and thinking of Python program teaching in junior high school based on micro: Bi [J]. *China information technology education*, 2018 (Z3): 52-55.

RISC-V Based Safety System-on-Chip with Hardware Comparator

EIKE HAHN¹, DOMINIK KALINOWSKI, WALDEMAR MUELLER,
MOHAMED ABDELAWWAD and JOSEF BOERCSEOK

ICAS, Institute for Computer Architecture and System Programming, University of Kassel, Kassel, Germany

Abstract. In this paper, a Safety System-on-Chip based on the open-source RISC-V processor SweRV EH1 from Western Digital is presented. A hardware comparator concept is followed. The SSoC is implemented on a Xilinx FPGA system and extended with standard peripherals from the Xilinx IP library and from Cobham Gaisler, so that the overall system has an Ethernet interface in addition to GPIO and UART. The goal is to create a complete redundancy approach with a hardware fault tolerance of nearly 1 from input to output based on the freely available RISC-V instruction set and prove its feasibility.

Keywords. Functional Safety, Miniaturized Safety Systems, Safety SoC, RISC-V

1. Introduction

The ICAS of the University of Kassel has been working for many years on the development of novel structures for embedded systems related to functional safety applications, which are realized specifically for use on programmable hardware or as ASICs. The goal is always the conformity with existing safety standards like IEC 61508 [1], ISO26262 [2], ISO 13849 [3] or similar.

Nowadays, functional safety microcontroller systems are increasingly being designed as a completely integrated system, a so-called Safety System-on-Chip (SSoC). Depending on the application, a hardware fault tolerance (HFT) of at least 1 is targeted for safety-critical architectures. 1-out-of-2 (1oo2), or in general MooN, stands for the degree of redundancy and the tolerance to errors in the system. In the best case, a 1oo2 system achieves a HFT of 1. The overall system consists of two independent systems, one of which is required to perform the task. A 1oo3 system can thus achieve a HFT of 2, a 1oo4 system a HFT of 3 [4].

Current systems mostly contain redundant CPU structures, but the peripherals are singular implemented. This leads to lower HFT than 1 for a 1oo2 redundant system. For most applications this approach is sufficient but there are cases where full redundancy is needed. Examples for such systems are the SSoCs from Texas Instruments [5] or from

¹Corresponding Author: Research Engineer, Institute for Computer Architecture and System Programming, University of Kassel, Wilhelmshöher Allee 71, 34121 Kassel, Germany; E-mail: eike.hahn@uni-kassel.de.

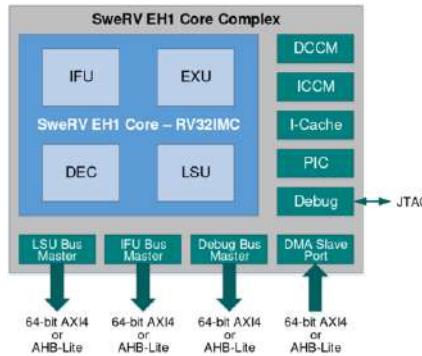


Figure 1. SweRV EH1 Structure [8, p. 1]

Infineon [6]. Also, the new RISC-V safety processor from Fraunhofer IPMS and CAST Inc. only provide a redundant processor architecture with singular peripherals [7].

In the development of functional safety microcontroller systems, microcontroller IPs are used, which are connected to form 1oo2, 1oo2D, 1oo4 or other structures depending on the application. Potentially applicable is a large number of available microcontrollers, which are based on a wide variety of instruction sets. The range extends from freely available and open source to highly optimized and therefore expensive architectures.

Especially for research and for applications with manageable quantities, the IP costs for the microcontrollers used are essential. Furthermore, in the area of functional safety, the use of proven-in-use hardware is important in order to be able to exclude structural errors in the hardware components as far as possible. The SweRV EH1 from Western Digital [8] (Figure 1) fulfils both requirements, since on the one hand it is open source and available under Apache 2.0 license and on the other hand it is already used by Western Digital in millions of hard disk controllers [9].

The paper at hand presents the implementation of a safety structure based on the mentioned SweRV EH1. The resulting ReSC-5 SSoC is based on a full 1oo2 architecture with HFT near to 1 and is implemented on a Xilinx AC701 [10] designed as FPGA. A hardware comparator is implemented as comparison unit. Finally, the SSoC is validated both in simulation and implementation on the FPGA, and the functionality of the hardware comparator is proven by fault injection.

2. Architectural Model of ReSC-5 SSoC

In the field of functional safety, several approaches exist describing how a required safety level can be achieved. In addition to the requirements for reliability, these approaches also include requirements for the availability of a system.

Conventional 1oo2 redundant structures, as used by common safety processors, are based on a hardware comparator principle. Here, the internal system buses of the processing units are permanently compared. This approach triggers a fault condition, and puts the system into a safe state, whenever a difference in the bus systems of the redundant channels occur.

In addition to the development of widely used hardware comparator systems [11–13], ICAS at the University of Kassel has been developing structures with software

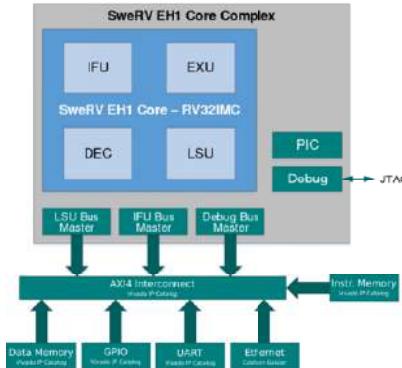


Figure 2. Modified SweRV EH1 Structure

comparators [14–16] for several years. This approach is based on the principle that the processing units of the channels independently compare important values and, in the case of a difference, independently set the system into a safe state. The advantage of this approach is the possibility of executing different software on the processing units, which, however, evaluate safety-critical information independently of each other and thus guarantee the safety of the overall system. In the event of a fault, it is also possible to continue operating the system with a reduced safety level rather than switching it off.

However, here the following architecture for a hardware comparator concept is used: A two-channel system is implemented, which consists of two SweRV EH1 cores. These cores are modified and extended with different peripheral modules, so that two independent and completely redundant processing units are created, each of them represents a complete 1oo1 system. The core of each 1oo1 system is based on original SweRV EH1 core but the directly connected data memory (DCCM), the instruction memory (ICCM) and the instruction cache (I-Cache) are removed. The interrupt controller (PIC) and a debug interface remain in the system. The SweRV EH1 core is extended with a data and an instruction cache from the Xilinx Vivado IP Catalog [17]. Moreover, a GPIO, a UART and an Ethernet module are connected to enable minimal communication. The GPIO and UART modules are also taken from Xilinx, while the Ethernet module from Cobham Gaisler's GRLIB [18] is used. All peripheral modules are connected via an AXI4 interconnect provided by the Vivado library. Figure 2 shows the overall structure of the modified system.

Based on these 1oo1 systems, a 1oo2 system is built up: The single-channel system is instantiated twice for this purpose. The system clock for both channels comes from a common clock source, also called lock-step mode. Since both processing units must execute the same code at the same time, monitored by a hardware comparator, they share a common instruction memory. However, to keep the possibility of common-cause-failures (CCF) low, each processing unit has its own data memory.

The UART interface connections and the GPIO pins are routed separately to the outside in order to establish complete redundancy for them as well. Likewise, each debug interface is routed out such that both processing units can be verified and debugged simultaneously. This redundancy implies nearly no difference in design in contrast to a singular implementation, but a singular implementation would be susceptible for CCF and is therefore avoided.

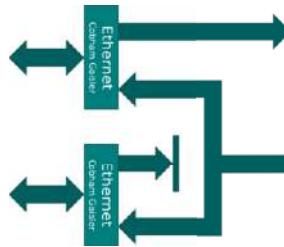


Figure 3. Single Ethernet Interface Concept

An Ethernet interface allows the system to be implemented in a higher-level structure and appears as a singular element. As this is implemented singular, special attention must be paid and is implemented as shown in Figure 3. One of the processing units realizes the communication to the outside while the second processing unit only listens passively for all incoming signals. The second unit does not notice that all its own outgoing signals are not connected, since the first unit sends identical data at the same time and therefore also receives the expected data at the expected moment. The principle represents a common method for the design of singular interfaces of redundant systems. However, this method is more critical for an Ethernet interface than for other standard interfaces, since the base frequency is orders of magnitude faster than, for example, the one of I²C, which is usually 400 kHz.

The major issue in the conceptual design of the overall system is the avoidance of common-cause-failures. The proposed method to avoid CCF foresees redundancy of the peripherals. A remaining drawback is, that a failure may affect both processing units at the same time. Such a failure can, for example, be caused by external radiation or by changing electro-magnetic fields, and can never be avoided.

To circumvent this, one of the two processing units is delayed by two system clock cycles to realize temporal independence between the two channels. Then similarly affecting failures result in different outcomes in both channels, as both units do not longer perform the same operation at the same time. In case of a bit-flip in both processing units caused by radiation, the comparator structure detects a difference between the processing units and puts the system into a safe state. An overview of the implemented SSoC structure can be seen in Figure 4.

Since the Ethernet interface is implemented singularly, and both processing units need to access it simultaneously, the signals from and to the delayed unit must also be delayed. The delay is realized by D-flipflops in the data path.

3. Implementation of the Hardware Comparator

Finally, a hardware comparator is integrated into the system. The basic task of this is to compare the system buses of the two processing units and, in the event of a difference, to set the system to a safe state. With the approach implemented here, the instructions on the instruction bus are compared, since this carries all CPU instructions. If there are differences, this is a strong indication that an error has occurred in one of the two processing units and the system must be put into safe state. Due to the two delaying clock

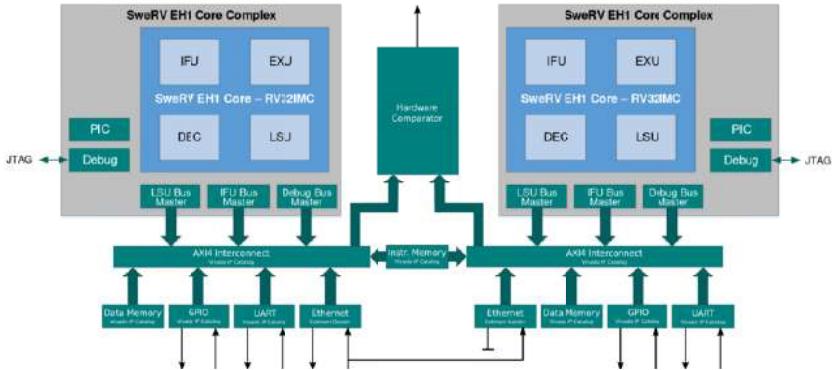


Figure 4. 1oo2 Safety Structure Concept of the ReSC-5 SSoC

cycles between the two processing units, the compared signal must also be delayed such that the comparator compares the same data with regards to content.

The structural design of the comparator can be described with the functionality of an XOR gate. The same signals of both processing units are XORed and in case of a difference a corresponding failure signal is set. This is brought out as a digital output in the design created here. In case of system integration, it can be connected to a special reset input of the SSoC. The reset state is defined as a safe state in most cases. A special reset controller can detect the reason of the last reset at system start-up and distinguish between a normal system start-up or a reset due to a fault condition. A subsequent software-based system diagnosis detects whether the failure had a transient or a permanent cause.

4. Validation of the RESC-5 SSoC

The implemented safety structure is first validated in simulation and then on a FPGA. The environment integrated in Xilinx Vivado is used as the simulation environment. Test programs, which are written in C, are loaded as a .coe file into the Vivado memory model of the instruction memory and executed there when the simulation is started.

First, the GPIO and the UART module are tested. The test program for the GPIO reads an input which is connected to an external push button and controls a LED connected to another pin of the GPIO module. Since both channels have separate GPIO modules, push buttons and LEDs are connected to both channels. Figure 5 shows the FPGA board with the expansion board and the debuggers.

Second, a program is developed, which validates the Ethernet interface. The free LWIP library is used to enable Ethernet communication. On a PC a command line application is used to send and receive single packets.

The communication is shown in Figure 6. It can be seen that the two channels of the safety structure receive a packet via Ethernet and return the data via UART. In addition, it can be seen in the second part of the figure that the tool receives data sent by the interfaces.

Finally, the hardware comparator is validated. For this purpose, a fault is injected intentionally. The wiring diagram for the concept can be seen in Figure 7. The bus signals



Figure 5. AC701 FPGA Eval Board

of the second processing unit are intentionally changed by an AND gate when a push button is pressed. The comparator applies a high level signal to the designated output directly after the key is pressed. This can be recognized by an illuminated LED.

5. Conclusion

The structure of a 1oo2 safety architecture with two SweRV RISC-V processor cores is presented in this paper. The comparator is based on the hardware comparator principle. The minimal system has been extended with different communication interfaces and represents a ready to use pin redundant safety system. Tests both in simulation and on FPGA have been successfully performed.

The presented method enables the possibility to realize this concept as an ASIC, for example through the Mini@sic program from Europractice [19], and to apply it in real applications – this is matter of future work of the authors. Furthermore, various structural specifications from the generic safety standard IEC 61508 were taken into account in the design generation, such that certification of the structure up to SIL3 is possible. However, a comprehensive calculation of the safety parameters and various qualitative and quantitative analyses, which are also necessary for certification, are still pending.

```

Start
Connected to ReSC-5 UDP Server
sent: "Test 0"
sent: "Test 1"
sent: "Test 2"
sent: "Test 3"
sent: "Test 4"
sent: "Test 5"

Start
Core A UDP Server
received: "Test 0"
received: "Test 1"
received: "Test 2"
received: "Test 3"
received: "Test 4"
received: "Test 5"

Start
Core B UDP Server
received: "Test 0"
received: "Test 1"
received: "Test 2"
received: "Test 3"
received: "Test 4"
received: "Test 5"

Start
Core B UDP Server
received: "Test 0"
received: "Test 1"
received: "Test 2"
received: "Test 3"
received: "Test 4"
received: "Test 5"

```

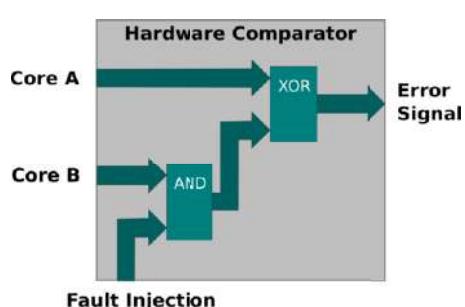


Figure 6. Ethernet Communication

Figure 7. Hardware Comparator Structure

References

- [1] Functional safety of electrical/electronic/programmable electronic safety-related systems, IEC 61508, IEC, Apr. 2010.
- [2] ISO 26262 Road Vehicles– Functional Safety, ISO, 2018.
- [3] ISO 13849 Safety of machinery, ISO, 2015.
- [4] J. Börzsök, Functional Safety: Basic Principles of Safety-related Systems, 1st ed. Heidelberg, Neckar: Hüthig Verlag, 2006.
- [5] Texas Instruments, Safety Manual for TMS570LS31x and TMS570LS21x Hercules™ and ARM® Safety Critical Microcontrollers: User’s Guide.
- [6] Infineon Technologies AG, Highly Integrated and Performance Optimized 32-bit Microcontrollers for Automotive and Industrial Applications. Neubiberg.
- [7] Fraunhofer Institute for Photonic Microsystems IPMS, EMSA5-FS – RISC-V Functional safety processor IP Core (accessed: Aug. 24 2021).
- [8] Western Digital Corporation, RISC-V SweRV™ EH1: Programmer’s Reference Manual. [Online]. Available: https://raw.githubusercontent.com/westerndigitalcorporation/swerv_eh1/master/docs/RISC-V_SweRV_EH1_PRM.pdf (accessed: Jul. 29 2021).
- [9] Zvonimir Z. Bandic, Robert Golla, SweRVcores Roadmap. Accessed: Jul. 29 2021. [Online]. Available: <https://riscv.org/wp-content/uploads/2019/12/12.11-14.20a3-Bandic-WD.SweRV.Cores.Roadmap.v4SCR.pdf>
- [10] Xilinx, Xilinx Artix-7 FPGA AC701 Evaluation Kit. [Online]. Available: <https://www.xilinx.com/products/boards-and-kits/ek-a7-ac701-g.html> (accessed: Jul. 29 2021).
- [11] A. Hayek, B. Machmür, M. Schreiber, J. Börzsök, S. Götz, and M. Epp, “HICore1: “Safety on a chip” turnkey solution for industrial control,” in 2014 IEEE 25th International Conference on Application-Specific Systems, Architectures and Processors, 2014, pp. 74–75.
- [12] M. Abdelawwad, A. Hayek, A. Alsuleiman, and J. Börzsök, “FPGA Implementation of a Safety System-on-Chip Based on 1004 Architecture Using LEON3 Processor,” in 2018 International Conference on Computer and Applications (ICCA), 2018, pp. 231–235.
- [13] A. Hayek and J. Börzsök, “On-chip safety system for embedded control applications,” in MELECON 2014 - 2014 17th IEEE Mediterranean Electrotechnical Conference, 2014, pp. 315–319.
- [14] Josef Börzsök, Waldemar Müller, Eike Hahn, Michael Schwarz, and Mohamed Abdelawwad, “Approach for a Safe-SoC for Cyber-physical Application according to IEC 61508,” International Journal of Computers, vol. 14, 2020, doi: 10.46300/9108.2020.14.12.
- [15] J. Börzsök, W. Müller, E. Hahn, M. Schwarz, and M. Abdelawwad, “Safe-System-on-Chip for Functional Safety,” in 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia, 2021, pp. 619–624. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9429321>
- [16] M. Abdelawwad, M. Drabesch, M. Schwarz, M. I. Hafiz, and J. Börzsök, “Communication SoC based on 1002D architecture for industrial human-robot-collaboration,” in 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia, 2021, pp. 954–959. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9429310>
- [17] Xilinx, Intellectual Property. [Online]. Available: <https://www.xilinx.com/products/intellectual-property.html>
- [18] Cobham Gaisler AB, SoC Library - Overview. [Online]. Available: <https://www.gaisler.com/index.php/products/ipcores/soclibrary/soclibrary-overview>
- [19] EUROPRACTICE, EUROPRACTICE — MPW & Mini@sic. [Online]. Available: <https://europractice-ic.com/mpw-prototyping/general/mpw-minisic/>

Classification of Knee Osteoarthritis Severity Using Modified Masks to Preprocess X-ray Images in a Deep Learning Model

Ching-Chung YANG¹

Department of Digital Media Design, Tatung Institute of Technology, Taiwan

Abstract. We propose a concise approach to facilitate the deep learning model for medical image classification of knee osteoarthritis severity. The characteristics of the input X-ray images are sharpened by a modified 5×5 mask before training and testing in this work. We compare the inference accuracies of two experiments using the same architecture with images sharpened and not sharpened respectively. And we find it tangible that the former performs much better than the latter. This technique could also be helpful when applied onto the edge devices for object detection and image segmentation.

Keywords. Image classification, deep learning, image sharpening

1. Introduction

Image classification has been inspiring the rapid development of deep learning these years. And it also becomes as important as to adapt the applications to various fields, such as medical treatment, traffic analysis, automated optical inspection, etc.

For medical image diagnosis employing artificial intelligence measures, the preparation of the dataset is sometimes demanding owing to the limited patient cases. In the meanwhile, the detected objects in the input images are subject to be confined within merely several classes. In consequence, there is no need to adopt large size neural network to fulfill the training and testing. A concise model with delicate structure might not only be adequate but easier to install in the edge devices.

However, the challenge is that the differences among the designated classes are usually quite small or even subtle. This inherent nature makes the inference accuracy of the medical deep learning schemes tend to dwell at an unacceptable level.

In this study, we would like to establish a compact model to achieve a sensible precision for classifying knee osteoarthritis severity. The datasets were originally downloaded from the website provided by Chen Ping-jun *et al.* [1]. We use a much simpler architecture to accomplish the similar accuracy by introducing a novel image sharpening technique.

¹ Corresponding author, Department of Digital Media Design, Tatung Institute of Technology, 253 Mi-Tuo Road, Chiayi City, Taiwan.; E-mail: yang10.cc@msa.hinet.net.

Most of the well-known contributions made to develop the deep neural networks focus on the construction of model layers [2-4]. While in this work we'll emphasize the importance at the data preprocessing stage.

The prevalent image preprocessing approaches consist of Gray-scale conversion, Gamma adjustment, Contrast-limited adaptive histogram equalization (CLAHE), etc. These methods mainly improve the global or local contrast of the original images to tell apart objects with distinct appearances. However, the processing of fine characteristics might help more when deciphering the classified objects with similar look. In this article, we demonstrate one unique image sharpening technique to uncover such importance. By applying our algorithm upon the input images for the neural network, the model precision is liable to be evidently escalated with less computing resource consumption.

The commonly used Laplacian is often considered a second-order derivative to sharpen an image's edges. This operator is regarded isotropic because of its rotation invariant for enhancement. But our previous work has proven that the mask filtering is actually a superposition of directional elements [5-13]. It seems promising to exploit this concept onto the image classification scope, in which a number of images had better be preprocessed in order to acquire sensible training weights.

2. Algorithm

It is common to use 3×3 masks to process image objects. While in this work, three masks derived from a 5×5 one is developed to flexibly enhance images.

For a 5×5 mask of $[0 \ -1 \ 0 \ -1 \ 0; \ -1 \ -1 \ -1 \ -1 \ -1; \ 0 \ -1 \ 16 \ -1 \ 0; \ -1 \ -1 \ -1 \ -1 \ -1; \ 0 \ -1 \ 0 \ -1 \ 0]$, there are sixteen coefficients environing the central pixel at coefficient 16. When sharpening an image $f(i, j)$, the second derivative is now formulated as the following:

$$\begin{aligned} \nabla^2 f(i, j) = & 16f(i, j) - [f(i+1, j) + f(i-1, j) + f(i, j+1) + f(i, j-1) \\ & + f(i+1, j+1) + f(i-1, j+1) + f(i+1, j-1) + f(i-1, j-1) \\ & + f(i+2, j+1) + f(i+1, j+2) + f(i-1, j+2) + f(i-2, j+1) \\ & + f(i-2, j-1) + f(i-1, j-2) + f(i+1, j-2) + f(i+2, j-1)] \end{aligned} \quad (1)$$

The result of Eq. (1) is supposed to be imposed on the original image to obtain an enhanced one $g(i, j)$ as the following:

$$g(i, j) = f(i, j) + \nabla^2 f(i, j) \quad (2)$$

referring to our previous works, it is reasonable to divide Eq. (1) into three parts as the following [14,15]:

$$\begin{aligned} \nabla^2 f(i, j) = & \{4f(i, j) - [f(i+1, j) + f(i-1, j) + f(i, j+1) + f(i, j-1)]\} \\ & + \{4f(i, j) - [f(i+1, j+1) + f(i-1, j+1) + f(i+1, j-1) + f(i-1, j-1)]\} \\ & + \{8f(i, j) - [f(i+2, j+1) + f(i+1, j+2) + f(i-1, j+2) + f(i-2, j+1) \\ & \quad + f(i-2, j-1) + f(i-1, j-2) + f(i+1, j-2) + f(i+2, j-1)]\} \\ = & \nabla^2 f_\alpha(i, j) + \nabla^2 f_\beta(i, j) + \nabla^2 f_\gamma(i, j) \end{aligned} \quad (3)$$

Eq. (3) means that a traditional 5×5 mask with 17 coefficients can now be dissembled into three masks, of which the second derivatives are just like summations of first derivatives along different directions as the following Eqs. (4), (5) and (6):

$$\begin{aligned}
\nabla^2 f_\alpha(i, j) &= 4f(i, j) - [f(i+1, j) + f(i-1, j) + f(i, j+1) + f(i, j-1)] \\
&= [f(i, j) - f(i+1, j)] + [f(i, j) - f(i-1, j)] \\
&\quad + [f(i, j) - f(i, j+1)] + [f(i, j) - f(i, j-1)]
\end{aligned} \tag{4}$$

$$\begin{aligned}
\nabla^2 f_\beta(i, j) &= 4f(i, j) - [f(i+1, j+1) + f(i-1, j+1) + f(i+1, j-1) + f(i-1, j-1)] \\
&= [f(i, j) - f(i+1, j+1)] + [f(i, j) - f(i-1, j+1)] \\
&\quad + [f(i, j) - f(i+1, j-1)] + [f(i, j) - f(i-1, j-1)]
\end{aligned} \tag{5}$$

$$\begin{aligned}
\nabla^2 f_\gamma(i, j) &= 8f(i, j) - [f(i+2, j+1) + f(i+1, j+2) + f(i-1, j+2) + f(i-2, j+1) \\
&\quad + f(i-2, j-1) + f(i-1, j-2) + f(i+1, j-2) + f(i+2, j-1)] \\
&= [f(i, j) - f(i+2, j+1)] + [f(i, j) - f(i+1, j+2)] \\
&\quad + [f(i, j) - f(i-1, j+2)] + [f(i, j) - f(i-2, j+1)] \\
&\quad + [f(i, j) - f(i-2, j-1)] + [f(i, j) - f(i-1, j-2)] \\
&\quad + [f(i, j) - f(i+1, j-2)] + [f(i, j) - f(i+2, j-1)]
\end{aligned} \tag{6}$$

Eq. (4) denotes the mask with four coefficients along the vertical and horizontal directions, Eq. (5) denotes the mask with four coefficients in the diagonal directions, and Eq. (6) is another one that has eight coefficients along other eight different directions. The three masks shown on the right-hand side of Fig. 1(a) essentially have different intrinsic properties due to the different distances from their coefficients to the target pixel, which is indicated in Fig. 1(b). Eq. (4) could extract the finest characteristics from the original image, Eq. (5) implies the higher spatial frequencies, and Eq. (6) contains more high-intermediate spatial frequencies. In order to flexibly adapt to different image compositions, Eq. (3) is more sensible to be modified as the following:

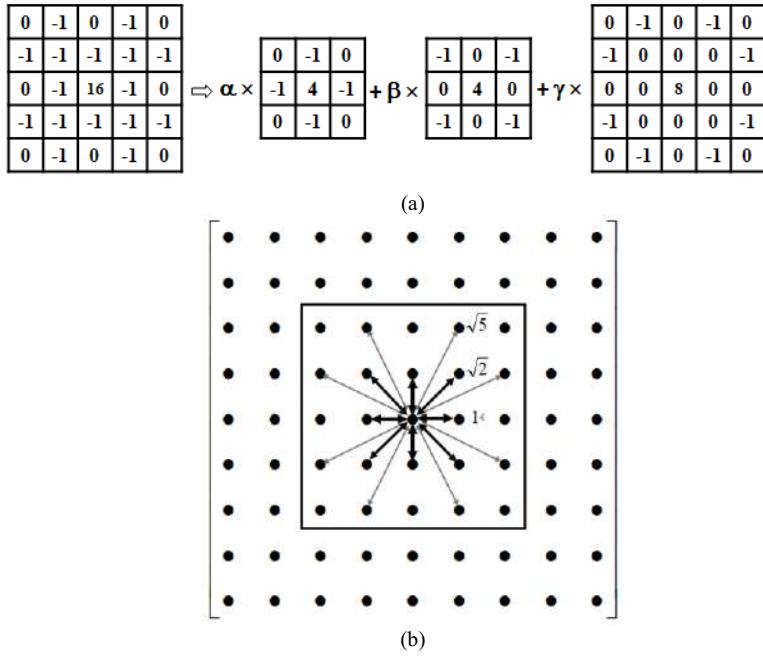


Figure 1. (a) The 5x5 mask divided into three masks with assigned weights. (b) The distances to the target pixel from different mask coefficients.

$$\nabla^2 f(i, j) \Rightarrow \alpha \times \nabla^2 f_\alpha(i, j) + \beta \times \nabla^2 f_\beta(i, j) + \lambda \times \nabla^2 f_\gamma(i, j) \quad (7)$$

And Eq. (2) could be rewritten as

$$g(i, j) = f(i, j) + \alpha \times \nabla^2 f_\alpha(i, j) + \beta \times \nabla^2 f_\beta(i, j) + \lambda \times \nabla^2 f_\gamma(i, j) \quad (8)$$

The above enhancement technique is developed here to preprocess the datasets used in this work. Then we subsequently build a neural network similar to Lenet-5 to illustrate the merit of our proposed method.

3. Experiment

Fig. 2 shows our entire neural network architecture. Images with and without sharpening are about to be input to the model respectively for comparison. Every two convolutions are followed by a maxpooling. The flattening converts the image matrix to a one-dimensional full connection layer. Afterward, a dropout processing is combined with the hidden layer to help reduce the over-fitting phenomenon. The output layer performs the classification of five classes.

The intended input data in Fig. 2 are all gray-scale images with a single channel. The first two convolutions containing 3×3 masks are with 64 and 32 output channels respectively. And the last two are both with 16 output channels. Because the two maxpoolings totally rescale an image to one sixteenth its original size, the flattened layer would contain the same pixels number as the input image of $128 \times 128 = 16384$.

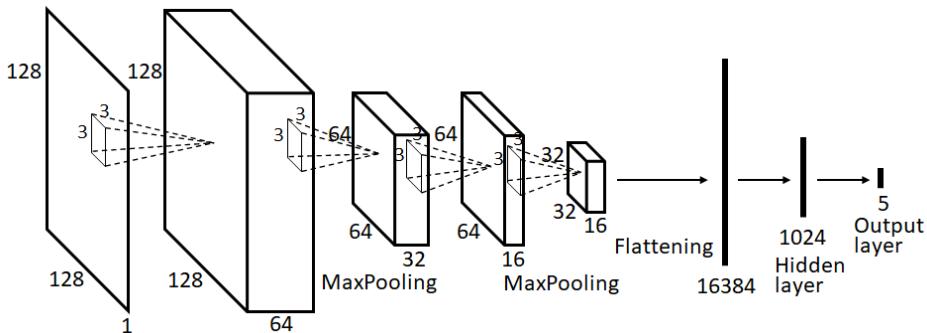


Figure 2. The CNN model and its parameters used in this work.

The images in the original datasets were all rescaled to be with the same size of 128×128 in this work. The severity of knee osteoarthritis is classified with five levels from 0 to 4. Images in the folders were rearranged so that every level contains 1500 images, with the exception that the folder labelled 4 only has 500 images. This is because that patients with most severe knee osteoarthritis are fewer than others. And the fewer images don't seem to degrade the experiment results too much.

Then we complete the desired datasets from all the labelled images. The training dataset includes 80% of them, the testing 10%, and the validation 10%. For comparison, we proceed twice the experiment. The first time is with datasets not sharpened, and the second time with datasets sharpened in advance by our method. Both of them exploit the same neural network model with identical parameters.

The constants used for sharpening images at the second time are $(\alpha, \beta, \gamma) = (0.4, 1.6, 0.15)$ in Eq. (8). In real practice, $f(i, j)$ in Eq. (8) can be multiplied by a constant near 1 to slightly adjust the brightness of the final image. We emphasize the weight of β as to magnify more higher spatial frequencies, which are considered important for revealing the characteristics of the knee osteoarthritis. All the constants are decided by empirical trial and error.

4. Result

Fig. 3(a) shows that the acquired inference accuracy is about 56% after 130 epochs of training at the first time. And it is about 71% at the second time. These outcomes were calculated by implementing Keras 2.4 and TensorFlow 1.14.0 in Python 3.7. Although the inference accuracy is possible to get higher after more epochs, Fig. 3(b) tells that the cross-entropy loss is also increasing in the meantime. In this situation, we have to take into account the cost of sacrificing the prediction confidence in case of running more epochs.

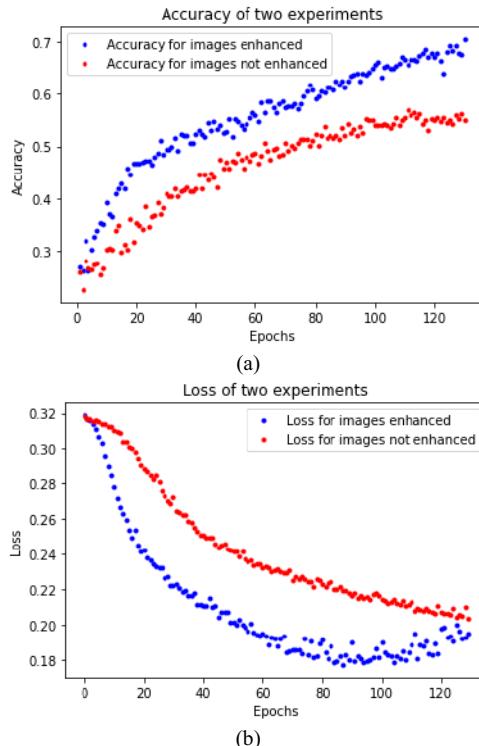


Figure 3. (a) The prediction accuracy diagram of the knee testing dataset for the two experiments. (b) The prediction loss diagram for the same two experiments.

At last, we load the trained weights trying to predict 50 labelled images in the testing dataset. Fig. 4 shows those images without enhancement and Fig. 5 shows the same images with enhancement. Table 1 elaborates the confusion matrices of these two experiments. It is conspicuous to observe the improvement in the matrix by our method.

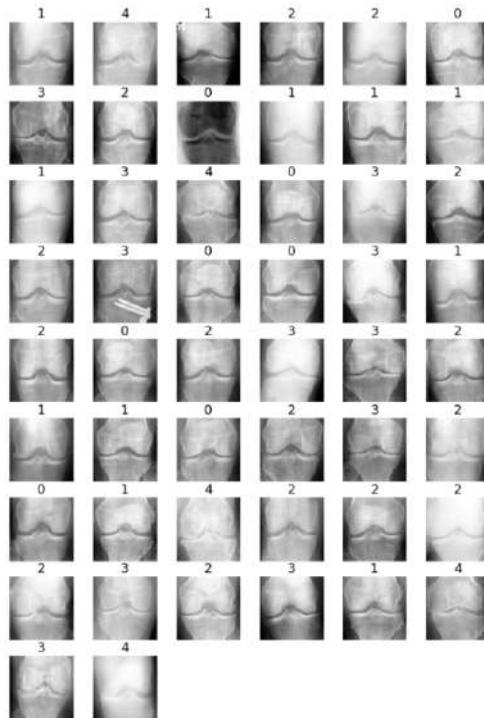


Figure 4. The 50 labelled knee testing images without enhancement.

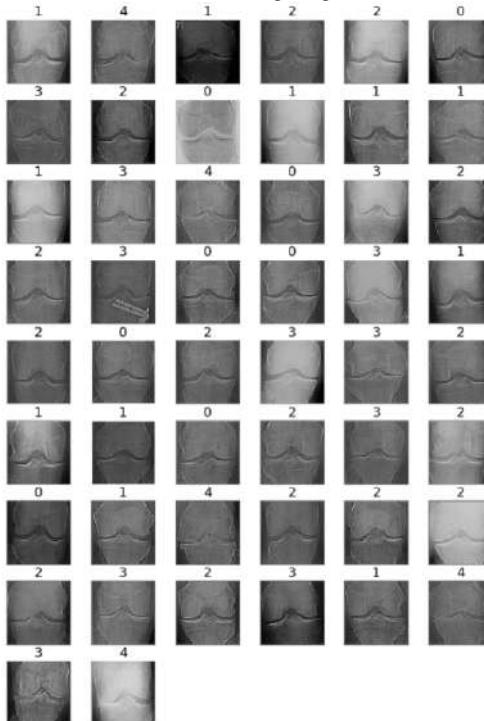


Figure 5. The same 50 labelled knee testing images with enhancement.

Table 1. The confusion matrices of the two experiments. (a). With images not enhanced. (b). With images enhanced.

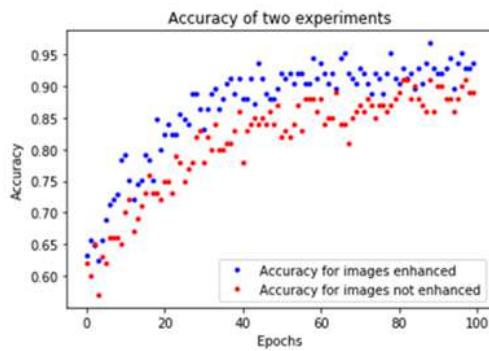
predicted	0	1	2	3	4
labelled					
0	5	1	2	0	0
1	3	2	5	1	0
2	2	6	7	0	0
3	0	0	1	10	0
4	0	0	0	2	3

(a)

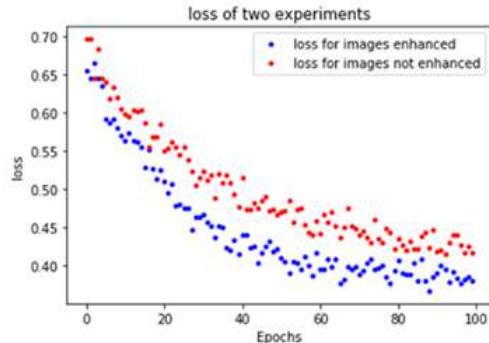
predicted	0	1	2	3	4
labelled					
0	6	0	2	0	0
1	2	5	4	0	0
2	3	3	9	0	0
3	0	0	1	10	0
4	0	0	0	0	5

(b)

5. Discussion and Conclusion



(a)



(b)

Figure 6. (a) The prediction accuracy diagram of the liver testing dataset for the two experiments. (b) The prediction loss diagram for the same two experiments.

To ensure the merit of our proposed method, we employ other datasets to train the model and also observe the improvement after the image enhancement. The practiced data are acquired from liver-cancer MRI scanning by Tungs' Taichung MetroHarbor Hospital in Taiwan. Livers inside patients' abdomens were scanned from top to bottom with image separation of several millimeters. The images had been with pseudonymization and de-identification when adopted in this work. We built the datasets by use of the separated scanned images, which are labelled 1 or 0 representing images with or without cancer focal area. Totally 664 images are labelled 0 and 336 images are labelled 1, for that the images with focal areas are fewer than those without.

Figure 6 shows the accuracy and loss curves for the training. The accuracy has been elevated from 90% to 95%. Figure 7 are the testing images without enhancement, and Figure 8 are the same images after enhancement. The constants used for Figure 8 are $(\alpha, \beta, \gamma) = (1.5, 0.1, 0.1)$. We emphasize the weight of α in this place because liver is composed of soft tissue with more sophisticated structures.

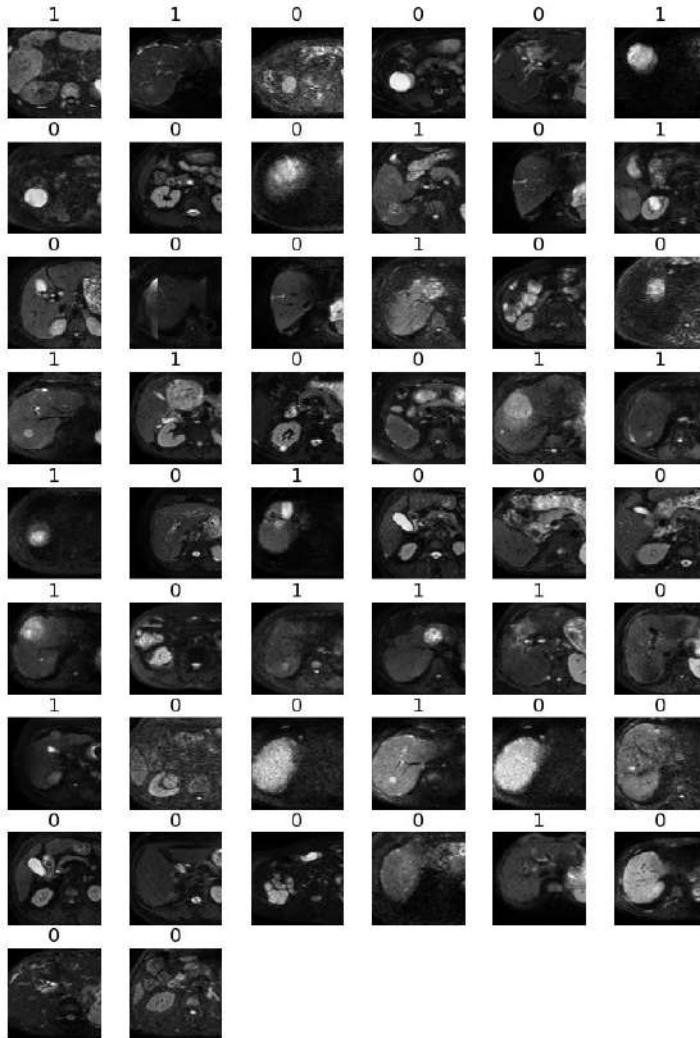


Figure 7. The 50 labelled liver testing images without enhancement.

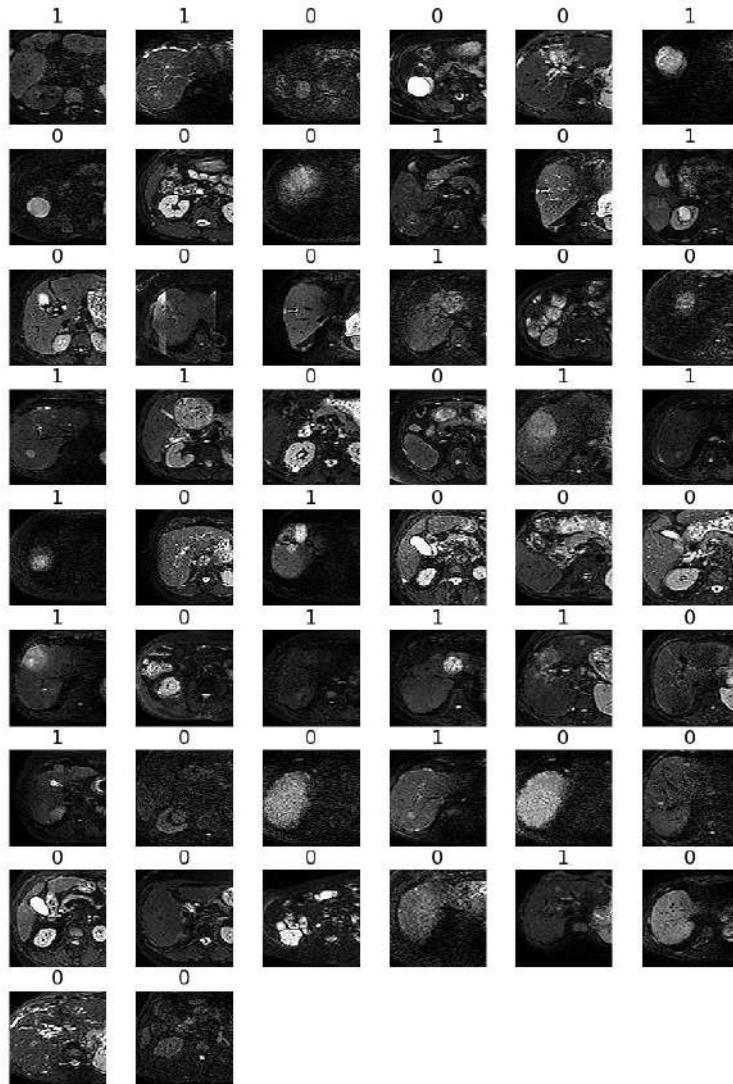


Figure 8. The same 50 labelled liver testing images with enhancement.

In conclusion, the image sharpening technique mentioned in this article is mainly based on mask-filtering approach, which has some similarity with the convolution operation in deep neural networks. Although one of them is with fixed coefficients and the other with trained coefficients in multi-channels, both of them extract features of the image.

Along with the emerging demand for edge computing software embedded in smaller devices, it seems promising that the recommended strategy serves as a preferential alternative. The proposed tactic not only reduces computing consumption, but provides the flexibility to accommodate to variant image compositions.

References

- [1] Chen PJ, Gao LL, Shi XS, Kyle A, Yang L. Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Computerized Medical Imaging and Graphics* 2019;75:84.
- [2] Yann L, Léon B, Yoshua B, Patrick H. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 1998;86:2278.
- [3] Krizhevsky A, Sutskever I, Geoffrey EH. ImageNet classification with deep convolutional neural networks. *Communications of the ACM* 2017;60:84.
- [4] Karen S, Andrew Z. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations* 2015 May.
- [5] Yang CC. Improving the sharpness of an image with non-uniform illumination. *Opt. Laser Technol.* 2005;37(3):235-8.
- [6] Yang CC. Image enhancement by modified contrast-stretching manipulation. *Opt. Laser Technol.* 2006;38(3):196-201.
- [7] Yang CC. Image enhancement by adjusting the contrast of spatial frequencies. *Optik* 2008;119(3):143-6.
- [8] Yang CC. Image enhancement by the modified high-pass filtering approach. *Optik* 2009;120(17):886-9.
- [9] Yang CC. A modification for the mask-filtering approach by superposing anisotropic derivatives in an image. *Optik* 2011;122(18):1684-7.
- [10] Yang CC. Color image enhancement by a modified mask-filtering approach. *Optik* 2012;123(19):1765-7.
- [11] Yang CC. Improving the overshooting of a sharpened image by employing nonlinear transfer functions in the mask-filtering approach. *Optik* 2013;124(17):2784-6.
- [12] Yang CC. Improving the sharpness of a non-uniformly illuminated image by an integral mask-filtering approach. *Optik* 2013;124(21):5049-51.
- [13] Yang CC. Finest image sharpening by use of the modified mask filter dealing with highest spatial frequencies. *Optik* 2014;125(8):1942-4.
- [14] Yang CC. Improving the sharpness of a color image with slowly-varying hues by undertaking its intensity and saturation factors. *Optik* 2014;125(22):6730-2.
- [15] Yang CC. Ultra-fine characteristics of a sharpened image rendered by decreasing the critical low-intermediate frequencies as well as increasing the critical higher frequencies. *Optik* 2015;126(23):3803-6.

Securing an IoT Medical System Using AI and a Unidirectional Network Device: Application to a Driver

Georges EL HAJAL^a, Roy ABI ZEID DAOU^{b, c¹}, Yves DUCQ^a and Josef BOERCQSOEK^d

^a Univ. Bordeaux, IMS laboratory, France

^b Biomedical Technologies Department, Lebanese German University, Lebanon

^c MART Learning, Education and Research Center, Chananir, Lebanon

^d University of Kassel, Institute for Computer Architecture and System Programming, Germany

Abstract. Security in systems and networks has always been a major issue for IT administrators. When it comes to medical applications, this concern is much more important due to the sensitivity of data and the risks that may be caused due to alteration or falsification of such critical information. The proposed paper presents a solution to assure the best security possible in such an environment. Thus, based on an application that monitors a driver's health while driving his car, a data diode will be implemented in order to assure security of the system by forcing unidirectional flow of network data to the healthcare provider side. Added to that, an AI-based program will be developed to verify the confidentiality, the integrity and the availability of the exchanged data and to check the patient health for abnormalities. Every sub-part of the system has been tested separately and results have shown that falsified data has been filtered out of the received end, e.g. the healthcare provider side.

Keywords. Cyber physical security, driver health, integrity, availability and confidentiality of data, unidirectional network device, AI algorithms

1. Introduction

The internet of things, or IoT, is a system of interrelated computing devices communicating through the internet allowing them to send and receive data. It allows the interaction between the physical world and the digital world. Sensors and actuators interact with the physical world and transform the data to the digital world. IoT technologies are contributing in the significant growth of generated data from medical appliances. Moreover, data security, confidentiality, integrity and availability are the challenges facing IoT applications and platforms due to the fast involvement of distributed diverse devices [1].

The eminence of IoT devices has been coupled with the increase of security vulnerabilities and exploitation. Most IoT systems utilize an architecture focusing on the connectivity with the cloud servers via gateways. Unfortunately, privacy and security risks are severe consequences of this architecture [2]. The communication between sensors and the IoT gateways might also be very susceptible to attack. Distributed

¹ Corresponding author, email: r.abizeiddaou@lgu.edu.lb

Denial-of-Service (DDoS) attacks can result in significant infrastructure collapse when targeting cloud servers. Moreover, a centralized server presents a risk to the entire system in case of a malware infection by generating a single point of failure.

In addition, IoT systems are not homogeneous in terms of security requirements, resource availability and networked devices. The devices operate in an open environment, which yields to increasing cyber physical risks and accessibility by adversaries.

The proposed application is integrated for drivers when doing their driving activity. Some sensors are deployed to measure their vital signs continuously and to inform healthcare providers whenever a medical issue is encountered. This is mainly applied for drivers having a medical history as heart failure, diabetes or even epilepsy. Thus, the main objective is to provide a secured transmission of data in order to make sure that the healthcare provider, or even the developed software, takes the right decision concerning the driver's health.

The proposed application is implemented for elderly people and being monitored remotely by healthcare providers; so, the proposed system relied on a telemedicine approach. Thus, the objective of this work is to make sure that the same data, sent from the patient side, is securely received for analysis by the healthcare provider to ensure the best follow up possible to this category of people.

Safety and security are highly recommended since this application deals with medical data. Thus, this paper will present a novel approach to ensure that the data, sent by the patient's device and processed by the healthcare provider tools, is authentic, confidential and available. The architecture of this systems consists of three main entities: the patient part where the sensors are placed, a gateway server that receives the data collected from the different patients to forward them to the required healthcare provider and the healthcare backend server that receives and analyses the data. So, to reach this objective, the proposed solution consists of two layers:

- a hardware layer of defense provided by installing a data diode (a unidirectional hardware device) between the gateway server and the backend server;
- a software layer of defense provided by an AI-based software installed on the gateway server. This software must identify, authenticate the receiver and check the data.

Thus, the novelty of this solution consists of presenting a solution combining both software and hardware to provide security for the overall system. This multi-layer security approach is not being adopted in almost all applications.

The rest of the paper is organized as follows. In Section 2, the applied tools and a description of the proposed solution will be presented. Section 3 shows the implementation of the proposed solution, also the analysis and the obtained results will be presented. Finally, Section 4 ends up with a conclusion and proposes some future works.

2. Proposed Solution

As already presented, the solution consists of two main levels: the unidirectional data diode and the AI-based software. Figure 1 shows the architecture of the system before and after implementing the solution. In more details, the figure to the left represents the three entities (*i.e.* the patients' sensors, the gateway server and the healthcare provider server) whereas the figure to the right represents the security add ins, mainly implemented in the intermediate layer.

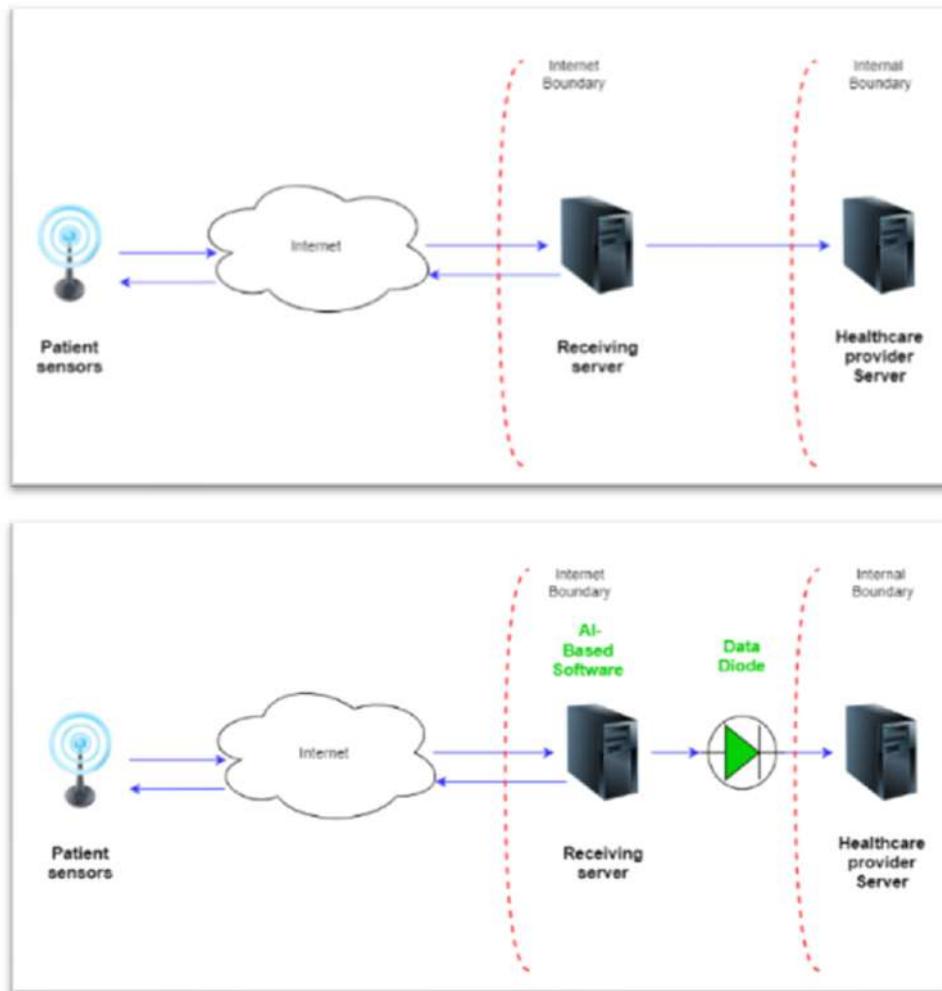


Figure 1. System architecture without (to the left) and with (to the right) the security modules/protocols
In more details, the system consists of three entities:

- A patient equipped by a heart rate sensor and a SpO₂ sensor to check his health while driving;
- A gateway server (IoT gateway) to receive the health data of multiple patients using this application. An AI-based software will monitor the incoming data to check for irregularities from the network and device perspectives, *i.e.* the IP, the time frame between sending of consecutive data, the formatting, etc... Also, it checks the health of the patient for abnormal measurements, and notifies the healthcare provider to reassess the patient health to contact him in case of a major problem;
- The healthcare provider's server that is used to monitor the patients' health.

Note that the unidirectional network device (Data Diode) will be installed between the IoT gateway and the healthcare provider's server to protect the internal system from malicious attacks [3].

Although not all vulnerabilities have been resolved, this work deals with most frequent and common ones according to OWASP API top ten security list [4].

Concerning the hardware layer, using an air gapped (isolated) network has been the most effective way to ensure that a network can't be compromised remotely. Air gapped systems ensure that, in case of a cyber security attack, no data can be communicated to the outside. The two most famous frameworks for cyber-attacks, the Lockheed Martin Cyber Kill Chain [5] and the MITRE ATT&CK [6] have described in their approaches the reconnaissance and the communication with the Command and Control (C&C) as vital for the success of any Cyber-attack. Data diodes provide security from these two phases because of the physical nature of unidirectional flow of data that doesn't allow the devices to interchange information.

As for the software layer, an AI-based software is installed on the input side of the exchange server to assess data transferred to the inside and to block any malicious action. Using Artificial Intelligence applications provide great security features, high transparency and enhance efficiency[7]. The development of edge computing permitted data generated by the IoT devices to be transferred to the edge gateways for further process and analysis before being forwarded to the servers that are used by the monitoring personnel.

Figure 2 represents a DFD to show the way data flows from the patient's side to the healthcare provider end considering all security measurements proposed in this solution.

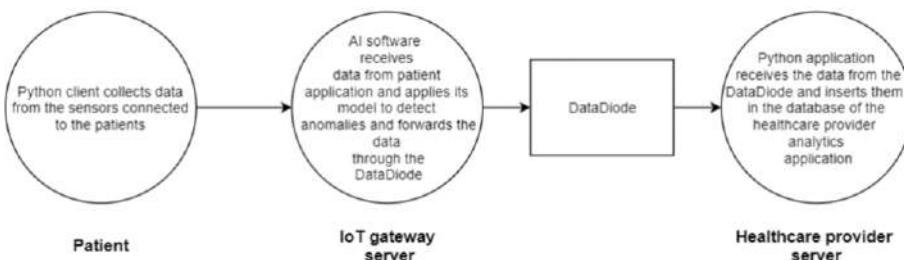


Figure 2. Scenario for the flow of data from the patient side to the healthcare provider backend

3. Implementation of the Proposed Solution

The solution divides the system into two segments: The first segment is the one between the patient (P) and the exchange server (the IoT gateway in this application) ($IoTg$), whereas the second segment is the one between the $IoTg$ and the server of the healthcare provider (HcP).

On the patient's side, a *raspberry pi zero* is equipped by an SPo2 sensor, heart rate sensor, and a GPS hat with 3G/4G & LTE Base HAT. The data of the patient is sent using a python script that collects all data from the sensors and send them to a Python Web application on the $IoTg$ server via an Internet connection. The $IoTg$ is connected to the input part of the Data Diode and the output part of the Data Diode is connected the server of the medical healthcare personnel. The AI application controls the sending of data from the input part and a python application handles data reception on the output part of the Data Diode to insert it in the database of the healthcare provider analytics application.

The preparation of the data for machine learning is essential in order to generate a model capable of detecting the anomalies later on. So, at first, the data was collected and

analysed. The anomalies detected and considered a priority to be handled by the system were:

- 1- Only the SPo2 value is different with more than 20% value deviation relative to the average of the patient.
 - 2- Only the heart rate value is different with more than 20% value deviation relative to the average of the patient.
 - 3- Attack on the API itself by a malicious actor.
 - 4- Traffic is coming from the same user agent but from the different IP.
 - 5- Traffic is coming from the different user agent but from the same IP.
 - 6- SQL injection on the Web Application.
 - 7- All the data is correct but different IP of original patient.
 - 8- The data sent by the user is not within the specified interval.

To train the AI/ML algorithms the following features were selected: user, spo2 value, heart rate value, GPS longitude value, GPS latitude value, date and time, IP, user-agent as shown in Figure 3 below.

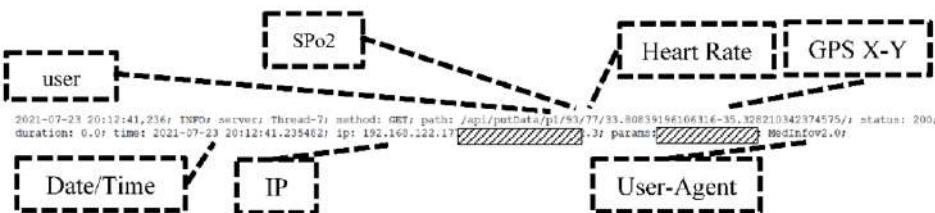


Figure 3. Features used for training AI/ML algorithms

The second step was to collect and label the data. We collected data values of normal traffic, then data values with medical issues, and at last, data values after simulating cyber-attacks using hping3 [8] to perform a DOS attack and Ettercap [9] for Man In the Middle attack from a KALI Linux machine. Our final dataset contained 27% normal traffic, 38% medical issue traffic and 35% cyber-attack related traffic.

Matlab R2018a was used to help us determine the best machine learning classifier that can achieve the highest accuracy. After applying 23 machine learning algorithms that are shown in Figure 4 below, some gave us a 100% accuracy while others gave us 37.9% accuracy. Added to that, we noticed that the KNN algorithm was the best for our data since it gave us 100% accuracy in all its variances. Based on these results, *KNeighborsClassifier* was used from “scikit-learn” [10] to be implemented in the python AI application.

KNN is a supervised classification algorithm. It generated new data points based on the closest data points or the k number. The latter should be large enough to limit noise in data, but also small enough in order not to interfere with the other classes [11]. So, different values of k were calculated in order to preserve the accuracy. For each value, the mean error was calculated. Figure 5 shows the accuracy value with respect to the value of k . Thus, one can find that any value below 19 is a good value for k .

1.1 ★ Tree	Accuracy: 100.0%
Last change: Fine Tree	8/8 features
1.2 ★ Tree	Accuracy: 100.0%
Last change: Medium Tree	8/8 features
1.3 ★ Tree	Accuracy: 96.4%
Last change: Coarse Tree	8/8 features
1.4 ★ Linear Discriminant	Accuracy: 83.2%
Last change: Linear Discriminant	8/8 features
1.5 ★ Quadratic Discriminant	Failed
Last change: Quadratic Discriminant	8/8 features
1.6 ★ SVM	Accuracy: 89.7%
Last change: Linear SVM	8/8 features
1.7 ★ SVM	Accuracy: 100.0%
Last change: Quadratic SVM	8/8 features
1.8 ★ SVM	Accuracy: 100.0%
Last change: Cubic SVM	8/8 features
1.9 ★ SVM	Accuracy: 100.0%
Last change: Fine Gaussian SVM	8/8 features
1.10 ★ SVM	Accuracy: 99.7%
Last change: Medium Gaussian SVM	8/8 features
1.11 ★ SVM	Accuracy: 96.5%
Last change: Coarse Gaussian SVM	8/8 features
1.12 ★ KNN	Accuracy: 100.0%
Last change: Fine KNN	8/8 features
1.13 ★ KNN	Accuracy: 100.0%
Last change: Medium KNN	8/8 features
1.14 ★ KNN	Accuracy: 100.0%
Last change: Coarse KNN	8/8 features
1.15 ★ KNN	Accuracy: 100.0%
Last change: Cosine KNN	8/8 features
1.16 ★ KNN	Accuracy: 100.0%
Last change: Cubic KNN	8/8 features
1.17 ★ KNN	Accuracy: 100.0%
Last change: Weighted KNN	8/8 features
1.18 ★ Ensemble	Accuracy: 37.9%
Last change: Boosted Trees	8/8 features
1.19 ★ Ensemble	Accuracy: 100.0%
Last change: Bagged Trees	8/8 features
1.20 ★ Ensemble	Accuracy: 79.7%
Last change: Subspace Discriminant	8/8 features
1.21 ★ Ensemble	Accuracy: 99.4%
Last change: Subspace KNN	8/8 features
1.22 ★ Ensemble	Accuracy: 37.9%
Last change: RUSBoosted Trees	8/8 features
2 ★ Quadratic Discriminant	Accuracy: 42.3%
Last change: 'Covariance structure' = 'Diagonal'	8/8 features

Figure 4 . Applied Machine learning algorithms and results

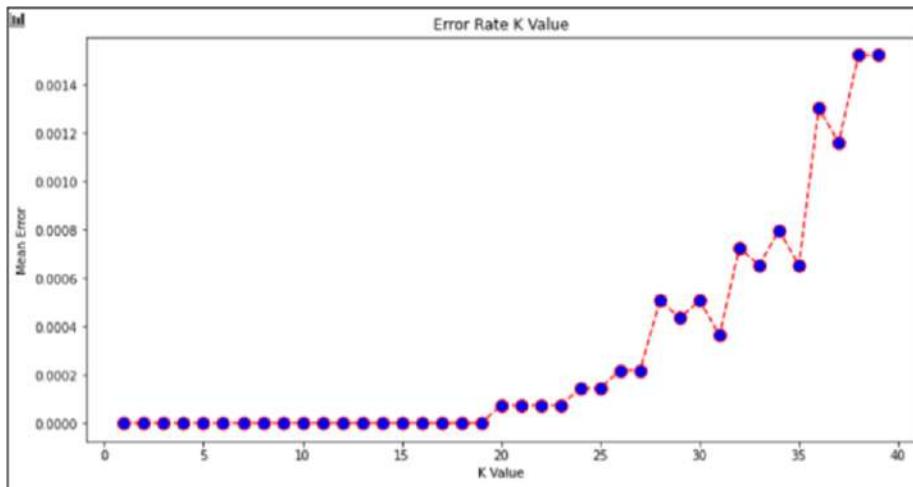


Figure 5. K-value vs Mean Error Value

Finally, the provided classifier was used in the AI application. The application will continuously read the received data and use the classifier to predict its nature and acts accordingly upon it. If it is normal traffic, the data is forwarded to the healthcare provider through the data diode. If it is a medical issue, the data will be labelled with a warning to notify the healthcare provider upon receipt. If it is classified as a cyber-attack, the application will block the IP of the sender and notifies the IT admin of the system of the issue. Figure 6 shows how the AI application is working flawlessly.

```

/api/putData/p2/95/98/33.28134249098785-35.60938853533984/ [Medical issue] SP02
/api/putData/p1/69/80/33.56298080389805-35.18545098031091 [Medical issue] Heart Rate
/api/putData/b2/7691041 [Cyber issue]
/api/putData/p1/94/75/33.721426622018754-35.44813912987872 [Normal]
/api/putData/p2/107/62/33.28270012174616-35.09866103994846 [Medical issue] SP02

```

Figure 6. AI application detecting normal or medical issues or Cyber issues.

From the Healthcare provider end, a python application will receive the data and forward them to the application used by the healthcare personnel to monitor the system.

After applying the AI algorithm, the hitting percentage was 100 for the KNN algorithm as 75% of the 54.000 packets were used for the training and 25% were used for testing. Added to that, neither the AI algorithm nor the data diode have caused any delay on data transmission between the patient device and the healthcare provider as time recordings were achieved first without implementing the security tools and it was then achieved after implementing the different peripherals and the time differences were very small, almost neglected.

4. Conclusions and future works

In the paper, an AI-based algorithm was implemented as a tool in the IoT gateway to ensure secure data transactions between the IoT distributed application. Added to that, to further enhance security, a unidirectional device known as Data Diode from our previous work [3] can provide additional layer of protection to the overall system.

The following main contributions are listed below:

1. Proposition of a novel architecture for IoT edge computing of medical devices to ensure security and privacy of transferred data. The proposed architecture uses mobile IoT technologies for communication;
2. Implementation of an AI-based application model to secure and authenticate framework for IoT devices in order to protect the sensor nodes' sensitive data from cyber-attacks. Furthermore, this solution provides automation of the threat detection and notification diffusion;
3. Experimental evaluation performance metrics of the proposed architecture regarding throughput, network usage, computational resources, transaction latency, and communication costs.

Building upon this work, our next step is to ameliorate our AI to secure the platform by classifying and prioritizing vulnerabilities based on aggregation of multiple sources of inputs such as CVSS score, the Exploit DB and other PoC providing websites.

References

- [1] J. Granjal, E. Monteiro and J. Sá Silva, "Security for the Internet of Things: A Survey of Existing Protocols and Open Research Issues," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1294-1312, 2015.
- [2] H. Garg and M. Dave, "Securing IoT Devices and Securely Connecting the Dots Using REST API and Middleware," in 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), 2019.
- [3] G. EL HAJAL, R. DAOU ABI ZEID, Y. DUCQ and J. Börcsök, "Designing and validating a cost effective safe network: application to a PACS system," in Fifth International Conference on Advances in Biomedical Engineering (ICABME), Tripoli, Lebanon, 2019.
- [4] OWASP, "OWASP API Security Project," [Online]. Available: <https://owasp.org/www-project-api-security/>. [Accessed 07 2021].

- [5] L. Martin, "the Cyber Kill Chain," 07 2021. [Online]. Available: <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>.
- [6] T. M. Corporation, "MITRE ATT&CK," 07 2021. [Online]. Available: <https://attack.mitre.org/>.
- [7] M. Dhingra, M. Jain and R. Jadon, "Role of artificial intelligence in enterprise information security: A review," in Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2016 .
- [8] "hping3," 2021. [Online]. Available: <https://tools.kali.org/information-gathering/hping3>.
- [9] "ETTERCAP," 2021. [Online]. Available: <https://www.ettercap-project.org/>.
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, "Scikit-learn: Machine Learning in {P}ython," *Journal of Machine Learning Research*, vol. 12, pp. 2825--2830, 2011.
- [11] K. Taunk, S. De, S. Verma and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," in International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019.

A Master Station and Terminal Data Exchange Method Based on Symmetric and Asymmetric Algorithms

Fan HE ^a, Zhengquan ANG ^b, Guanglun YANG ^a, Qingqin FU ^{a1}, Ling YI^c,
Pingjiang XU ^c, Jia LIU ^c, Zhaoqing LIANG ^c, Changsheng NIU ^c and Jianying CHEN^a

^a *ZhongGuanCun XinHaiZeYou Technology Co.Ltd, Beijing 100094, China*

^b *Beijing aerospace flight control center, Beijing 100094, China*

^c *Beijing Chip Microelectronics Technology Co., Ltd., Beijing 100192, China*

Abstract. This paper analyzes the disadvantages of the traditional method of data security between the master station and the terminal, and proposes a new method of data exchange between the master station and the terminal. This method improves the security of data interaction. In the process of using it, symmetric algorithm and asymmetric algorithm are combined, involving security mechanisms such as signature, certificate, MAC and symmetric encryption, so as to establish a secure link between the master station and the terminal, so as to protect the communication security between them.

Keyword. Master station, terminal, data interaction, symmetric algorithm, asymmetric algorithm

1. Introduction

In recent years, the State Grid Corporation has comprehensively promoted the construction of power consumption information collection systems. This system can achieve comprehensive coverage of all power users and gateways, realizing on-line monitoring of devices and real-time collection of important information such as user load, power, voltage, etc. Provide complete and accurate basic data for relevant systems, provide support for analysis and decision-making of all aspects of business management, and provide information foundation for intelligent two-way interactive services. This requires a large number of terminals to participate in the collection of data information. These power information's collecting devices can implement energy meter data collection, data management, data bidirectional transmission, and devices that forward or execute control commands.

To establish a comprehensive user power information collection system, it is necessary to construct a system main station, a transmission channel, an acquisition device, and an electronic energy meter. At present, most terminal data upload communication channels use GPRS communication. In order to ensure the safe and stable operation of the power user's electricity information collection system, to prevent the power supply interruption of the user, and to prevent a greater range of security

¹ Corresponding author, email: fujingqin@icrus.cn

risks caused by the public network and the user terminal invading the primary station, the primary station and the terminal need to perform authentication when performing data interaction with the terminal. And encryption, which requires a method and rules for specifying data transmission between the primary station and the terminal^[1-7].

2. Data exchange method between traditional primary station and terminal

The traditional method is: when the primary station interacts with the terminal data, the symmetric external authentication is performed first, and after the authentication is passed, the communication data is encrypted to ensure the secure transmission of the data.

The traditional method of interaction between the primary station and the terminal data, although there is an authentication operation before the data interaction, the symmetric algorithm is easy to be cracked and replayed, so the overall security is not high^[8-20].

3. Master station and terminal data exchange method based on combination of symmetric and asymmetric algorithms

In order to overcome the problems caused by the traditional methods, this paper proposes a new data exchange method for the primary station and terminal for power. This method divides the interaction between the primary station and the terminal into three steps:

1) The first step is the establishment of a session between the primary station and the terminal;

2) The second step is to use the session key negotiated in the session for data encryption operation;

3) The third step is to resume the session between the two if the communication reaches the limit of the number of times. After the session is restored, the second step can be repeated.

These three steps use a combination of symmetric and asymmetric algorithms in the process of use, involving security mechanisms such as signature, certificate, MAC, and symmetric encryption, so as to establish a secure link between the primary station and the terminal, thereby achieving the communication between the two is secure.

The specific implementation of this method is as follows.

3. 1. Master station and terminal session establishment

The specific process of session establishment is shown in Figure 1.

The specific method of session establishment is:

1. The primary station end group message 1, the specific organization message method is as follows:

- 1) Take the random number R1;
- 2) using a cipher machine to encrypt the random number R1 to obtain the ciphertext EKs1 (R1);

- 3) The version number represents the version of the protocol and the encryption algorithm used;
- 4) The session ID, initially initiated as 0, must be initialized to 0 once a day;
- 5) Take the master station certificate CM;
- 6) The MAC1 field is mainly used to use the last session key, the initial initiation, the field is filled with 0;
- 7) Digitally sign the version number, session ID, master certificate, MAC, and R1 to form signature S1.

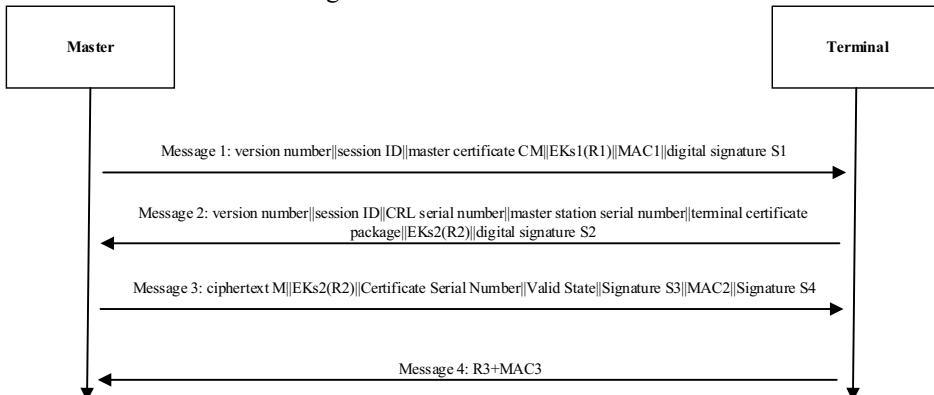


Figure 1. Master station and terminal session establishment process

2. The terminal receives the message 1 and performs the following operations:
 - 1) Verify the validity of the master certificate CM;
 - 2) decrypt EKs1 (R1) to get R1;
 - 3) Verify the legality of signature S1;
 - 4) The terminal ESAM chip generates a random number R2;
 - 5) Decrypting R2 with key EKs2 to generate EKs2(R2);
 - 6) Sign the S2 by using the terminal signature private key (version number, session ID, master certificate serial number, terminal certificate package, R2, R1).
3. The master station receives the message 2 and performs the following operations:
 - 1) Verify the validity of the terminal digital certificate;
 - 2) using the key EKs2 to solve the random number R2;
 - 3) verify the signature S2 using the terminal signature public key;
 - 4) The master station generates the master key K, and obtains the ciphertext M1 by using the terminal encryption public key encryption;
 - 5) Upload the EKs2 (R2) obtained in the message 2 and its serial number to the certificate server;
 - 6) using the master key K derived MAC key and initial vector, calculate the MAC for $\{(EKs2(R2) + \text{certificate serial number} + \text{valid state}) + \text{signature S3}\}$, generate MAC2;
 - 7) Sign the previous part using the master private key to form the signature S4.
4. The terminal receives the message 3 and performs the following operations:
 - 1) Verification signature S4

- 2) Verify its signature information, certificate serial number and generated EKs2(R2) to determine if the master certificate is correct;
 - 3) Using the terminal private key to unlock the ciphertext M1, and obtaining the master key K;
 - 4) calculating a master key, a data encryption key, a MAC key, and an initial vector according to a key derivation algorithm;
 - 5) Verify the correctness of MAC2;
 - 6) The terminal generates a random number R3 (for the MAC initial vector used for subsequent data interaction encryption);
 - 7) Calculating MAC3 with R3 using the MAC key and the initial vector;
 - 8) Send message 4 to the master station.
5. The primary station receives the message 4 and performs the following operations:
 - 1) Take R3 as the initial vector for subsequent data interaction;
 - 2) Verify the correctness of MAC3 using the MAC key and the initial vector.
 - 3) After the verification is passed, the session is successfully established, and the subsequent message uses the master key K, the data encryption key, the MAC key, the initial vector, and the random number R3 calculated in the interaction process.

3.2. Data Encryption Processing

In order to ensure the security of the session between the primary station and the terminal, the session key has a lifecycle limit when encrypting data using the negotiated session key (including: master key, data encryption key, MAC key, initial vector). When the session key expires after reaching a certain number of times, the session recovery process must be initiated by the primary station to renegotiate the new session key.

For example, if the session key lifetime is set to 10,000 times, the session key can be used up to 10,000 times. After 10,000 times, the session ID is set to 1. At this point, session recovery is required to communicate again, thus ensuring the timeliness and security of the session key.

3.3. Session Recovery

If the primary station side session ID flag is 1, the session recovery process is performed. The specific process and method of session recovery are as figure 2:

The specific method of session recovery is:

1. The primary station group message 1, the specific organization message method is as follows:

- 1) Take the random number R1;
- 2) using a cipher machine to encrypt the random number R1 to obtain the ciphertext EKs1 (R1);
- 3) The version number represents the version of the protocol and the encryption algorithm used;
- 4) The session ID is 1;
- 5) Take the master station certificate C M;

- 6) using the MAC key generated by the last session and the initial vector to perform MAC calculation on the session ID and EKs1 (R1) to obtain MAC1;
- 7) Digitally sign the version number, session ID, master certificate, MAC1, R1 to form signature S1.

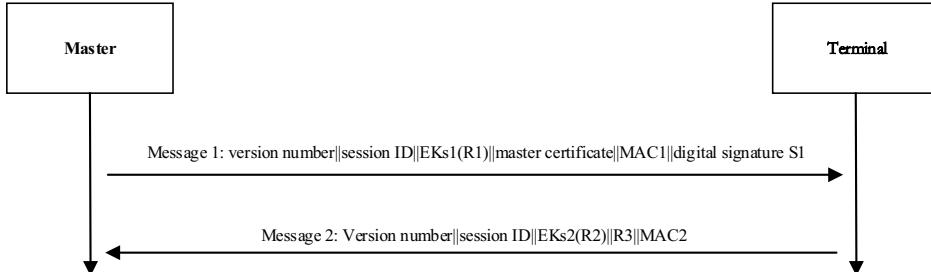


Figure 2. Master station and terminal session recovery process

2. The terminal receives the message 1 and performs the following operations:
 - 1) Take out the master station certificate CM in message 1, verify its validity, and expire;
 - 2) Verify the signature S1;
 - 3) Verify the MAC value in message 1;
 - 4) The terminal ESAM chip generates random numbers R2 and R3;
 - 5) using the preset key EKS2 to encrypt it to generate EKS2 (R2);
 - 6) Deriving a new master key, a data encryption key, a MAC key, and an initial vector using the new R1, R2, and the original master key K;
 - 7) performing MAC calculation on the session ID, EKS2 (R2) and R3 by using the newly generated MAC key and the initial vector to obtain MAC2;
 - 8) Send message 2.

3. The primary station receives the message 2 and performs the following operations:

- 1) decrypt EKS2 (R2) to get R2;
- 2) using the new R1, R2 and the original master key K to derive a new master key, data encryption key, MAC key and initial vector;
- 3) Verify the correctness of MAC2;
- 4) After the verification is passed, the registration is successful, and the subsequent message uses the newly calculated master key, data encryption key, MAC key, initial vector and random number R3.

3.4. Comparison of advantages between primary station and terminal data exchange methods

This paper proposes a new data exchange method for the primary station and terminal for power. This data exchange method has the following advantages:

- 1) This article expands the method of establishing the connection between the primary station and the terminal into two types, one is the new session establishment, and the other is the session recovery in the case that the session has been established.

- 2) After the master station and the terminal session are established, symmetric data and asymmetric algorithms can be used in subsequent data interaction, which improves the security of data interaction.
- 3) This paper proposes to limit the number of times the session negotiation key is used. When the number of communication between the two reaches a limited number of times, session recovery is required to ensure the security of the session.
- 4) The session key proposed in this paper has multiple key types such as master key, encryption key, MAC key, initial vector, etc., which are used in different occasions to meet the application requirements of different scenarios.
- 5) This paper proposes to use the negotiated master key to decentralize multiple keys such as encryption key, MAC key and initial vector, which ensures the randomness and diversity of key generation.

4. Conclusion

In this paper, the method of data exchange between the master station and the terminal based on the combination of symmetric and asymmetric algorithms is realized. The security of data interaction is improved because the security mechanism such as signature, certificate, MAC, and symmetric encryption is used in data interaction. At the same time, the interaction between the primary station and the terminal is divided into three steps: session establishment, data encryption, and session recovery. The user can select different steps according to the actual application situation to meet the application requirements of different scenarios and enhance the application diversity.

References

- [1] Li Tao, Zeng Ying. A New Remote Bidirectional Authentication Based on Dynamic Password[J]. Microcomputer Information, 2007, 11(3): 38-40.
- [2] Wu Jianwu. A remote password authentication scheme based on smart card[J]. Computer Engineering and Applications, 2007, 43(33): 158-160.
- [3] Kong Mengrong, Zhu Guohua. Remote authentication system based on smart card[J]. Computer Engineering and Design. 2008, 29-3: 606-608.
- [4] Wang Aiying. Smart Card Technology [M]. Beijing: Tsinghua University Press, 2009.
- [5] Xu Pingjiang. Design of smart card file system based on linked list method [J]. Microcomputer Information, 2011, 27 (11): 49-50.
- [6] Zhao Dongyan, Wang Yubo. Implementation of a smart card write protection mechanism [J]. Electronic Technology Application, 2014, 12: 32-34.
- [7] Lai Yuyang. A high security network data transmission implementation [J]. Information Security and Communication Confidentiality, 2016, 2: 109-112.
- [8] Xu Pingjiang. State machine based chip access control implementation [J]. Microelectronics and Computer, 2019, 36(7): 98-102.
- [9] Xu Pingjiang. Implementation of a high reliability read/write mechanism for security chips [J]. Computer Application Research, 2019 (increase): 280-282.
- [10] Du Shuwei, Zhao Dongyan. Smart grid chip technology and application [M]. Beijing: China Electric Power Press, 2019.
- [11] Fu Qingqin. Implementation of an improved smart card authentication method [J]. Computer Engineering and Science. 2014, 36 (1), 94-98.
- [12] Yuyang LAI. A grey lock method to support once pre-freezing mechanism in IC card[J]. Institute of Electrical and Electronics Engineers Inc. 2014, 1411-1414

- [13] Weibin LIN. A mechanism for patching ROM smart card[J]. Institute of Electrical and Electronics Engineers Inc. 2014, 1415-1417
- [14] Chen Huajun. An implement of the digital certificate on electrical IC card[J].CRC Press/Balkema, 2015,725-727
- [15] Fu Qingqin. A grey lock method to support multiple pre-freezing mechanism in IC card[J]. CRC Press/Balkema, 2015,1395-1400.
- [16] Qing-qin FU. [A Novel Power-down Protection Mechanism for Secure Chip Based on CRC Check](#)[A].EETA2017[C],2017,48-52.
- [17] Jia Liu. [Implementation of IC Card Authentication Method Based on Self-defined Algorithm](#)[A]. EETA2017[C],2017.
- [18] Fu Qingqin. A method for realizing secure chip multi-algorithm processing application[A]. Fuzzy Systems and Data Mining IV.2018,743-748.
- [19] Fu Qingqin. A new secure chip file access method based on security level information[A]. Fuzzy Systems and Data Mining IV.2018,749-756.
- [20] Pingjiang Xu. An Implementation of a Chip Security Mechanism[A]. Fuzzy Systems and Data Mining IV.2018,763-770.

Approach to Define the Reliability of Safety-Related Machine Learning Based Functions in Highly Automated Driving

Pengyu SI, Ossmane KRINI¹, Nadine MÜLLER and Aymen OUERTANI

Institute of Functional Safety, Cyber Security and Artificial Intelligence, Baden-Wuerttemberg Cooperative State University Lörrach, Germany

Abstract. Current standards cannot cover the safety requirements of machine learning based functions used in highly automated driving. Because of the opacity of neural networks, some self-driving functions cannot be developed following the V-model. These functions require the expansion of the standards. This paper focuses on this gap and defines functional reliability for such functions to help the future standards control the quality of machine learning based functions. As an example, reliability functions for pedestrian detection are built. Since the quality criteria in computer vision do not consider safety, new approaches for expression and evaluation of this reliability are designed.

Keywords. Functional safety, self-driving, HAD, reliability, pedestrian detection

1. Introduction

The development of semiconductor industry and artificial intelligence brings us closer to fully automated driving. The transition from hands-on driver assistance to hands-off automated driving, however, requires a huge change of system safety [1], which may rebuild the development process of safety systems in vehicles and redefine the legal responsibility between manufacturers and users. The utilization of artificial neural networks is seen as an effective approach for some or even all tasks in highly automated driving (HAD) [1,2], but the opacity of neural networks challenges the standard development process and functional safety considerations.

In this work, we consider the applicability of existing standards on machine learning based functions applied in HAD vehicles and propose a new kind of reliability to fill the gap of HAD safety requirement, which the current standards do not cover. We build functions to express this reliability mathematically in pedestrian detection and generalize the approaches that can improve this reliability.

Our paper is organized as follows. Section 2 introduces the insufficiency of current standards when applied to high-level self-driving vehicles. In section 3 we limit the scope of application and define the reliability functions for different purposes. Section 4 discusses the enhancement methods of reliability. We summarize our work in section 5.

¹ Corresponding author, Director of Institute of Functional Safety, Cyber Security and Artificial Intelligence located by Baden-Wuerttemberg Cooperative State University Lörrach, Hangstraße 46-50, D-79539 Lörrach, Germany; E-mail: krini@dhbw-loerrach.de.

2. The Conversion of the Functional Safety

Functional safety is an important part of the overall system safety in E/E/PE (Electrical/Electronic/Programmable Electronic) systems. It aims to protect people from unacceptable risk of physical injury or health damage, which is caused by human errors or system failures. In automobile industry, because of the high-level integration of electronic components and its inseparable relationship with safety, the functional safety must be considered in detail in the concept phase and completely validated in the test phase. Thus, the International Organization for Standardization published the ISO 26262: Road vehicles – Functional safety [3] to guide the automobile manufacturers. ISO 21448: Road vehicles – Safety Of The Intended Functionality [4], which is generally abbreviated as SOTIF, on the other hand, concentrates on how a target function is to be specified, developed, verified and validated so that it can be regarded as sufficiently safe.

Over the last decade, the development of neural networks (NN) and deep learning brings breakthroughs in computer vision and robotics, which deeply impacts the autonomous driving seen today both in academia and industry [2]. In automobile industry, the traditional standards for functional safety are no more suitable to guide the design and verification of NN-based safety functions [1,2]. The ISO 26262 follows the famous V-model, which splits the target function into sub-specifications and develops it under those specifications. Such a process is called white-box, the whole data processing is transparent and understandable. On the contrary, the NNs are generally black-box, the mapping between in- and output is a statistical approximation, which is not humanly understandable. The SOTIF aims to address such insufficiency of ISO 26262, but focuses mainly on driver assistance systems rather than HAD systems [1]. In driver assistance systems, the driver shall safely take over the vehicles when failure occurs. However, HAD system itself shall make sure a hazard-free status when failure occurs. Thus, the SOTIF cannot cover the safety requirements of HAD. The prospect of HAD demands more attention on the normalization of the development of NN-based functions in automobile. One of the core jobs for the normalization is bringing the functional safety requirements into HAD functions. Like the requirement of hardware reliabilities in ISO 26262, the reliabilities of functions in HAD system should be defined, so that corresponding ASIL (Automotive Safety Integrity Level) and safety methods can be determined.

The in ISO 26262 defined reliability, however, cannot describe the reliability of the HAD functionalities. In HAD systems, multiple sensors will be used for perception of the environment. Based on the environment modeling, the NNs will be performed to understand the surroundings (e.g. pedestrian, obstacle and parking space) and control the vehicle to drive safely. In this case, unlike traditional mechatronic systems in vehicles, one specific sensor takes responsibility for several tasks and multiple sensors work together for one specific task. Therefore, we propose to use the *functional reliability* to describe the quality of the specific HAD function. *Functional reliability* can be validated and calculated in development phase by using a validation dataset. In future standards, a reliability benchmark can be given to guide the automotive manufacturers. An example is given in the following sections to clarify this concept.

Similarly, the concept of software reliability exists in software engineering [5]. It can also not cover the HAD safety requirements, because different influences on safety of different functions are not considered. Imagine the consequence of a not detected person on the road and a not detected parking space. Different from software reliability, the *functional reliability* is function-oriented and services for safety evaluation.

3. The Reliability of Machine Learning Functions

In this chapter we use pedestrian detection as an example to clarify how the functional reliability of machine learning functions in HAD can be evaluated and how it describes the quality of the function regarding safety.

3.1. Function Boundary

In system safety approaches such as STAMP (Systems-Theoretic Accident Model and Processes), the safety constraints or boundaries must be clearly defined [6]. To analyze the intended function reliability, the function boundaries and use cases need to be clarified. Consequently, defined use cases are only part of the intended function [7].

There are two types of strategies in highly automated driving: End to End and perception-planning-action pipeline [2]. End to End means the algorithm maps the sensor data directly to driving commands without any middle steps, while the classical approach manually decomposes the process into the aforementioned pipeline. NVIDIA believes that handcrafted interfaces ultimately limit performance by restricting information flow through the system and researches continuously on End to End solution, they developed PilotNet to verify the strategy [8]. However, safety requirements for End to End learning approaches are more abstract, formulating and validating measurable performance criteria is significantly more difficult [1]. The decomposed process generally follows the perception and localization → high level path planning → behavior arbitration → motion control pipeline. In each part, several functions will be designed. Therefore, the function boundaries can be narrowly defined, so that the law responsibilities can be clarified when products are put on the market. Thus, considering the executability, the first boundary we formulate is: the reliability of machine learning based pedestrian detection is defined for decomposed HAD strategy.

The training technique is another factor that deeply influences the safety boundary. A neural network can be fully trained and then applied. During application the knowledge, abilities and performance of the network will not be changed. This is so-called offline training. Such training is centralized, all deployments have exactly the same performance. The opposite is online training, the neural network continuously learns from the inputs during application, for this, a mechanism to assess inputs is necessary. Obviously, each deployment of such networks evolves differently because of the individual diversity. This approach is often utilized in reinforcement learning. Clearly, decentralized, online training imposes currently great challenges for safety validation [1]. Thus, we define our second boundary: the considered reliability is aimed at centralized, offline trained neural networks.

Under these constraints we define the pedestrian detection function in HAD. First, we formulate the prioritized definition, which is safety-relevant and shall be verified. Hereafter comes the generalized definition.

- Prioritized definition: detection and tracking of all the pedestrians in front of the driving direction of the automated driving vehicles in urban areas (velocity usually under 60 km/h).
- Generalized definition: detection and tracking of all pedestrians in all directions around the vehicles.

Figure 1 shows a scenario of the use case and the corresponding definition zone.

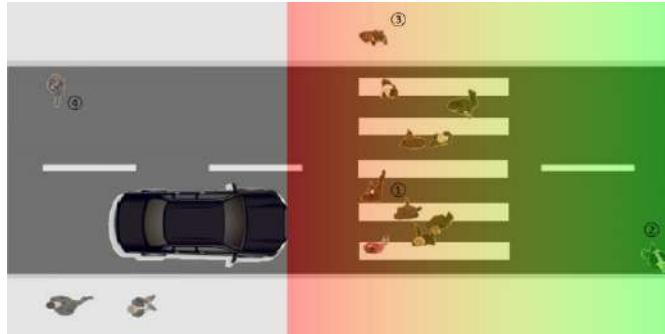


Figure 1. Driving scenario. The red-green colored area satisfies the prioritized definition. The red to green color gradient describes the urgency to take actions. Detection of people in ① demands the highest quality, while prediction ability attenuation in ② is allowed. Interaction with people like ③ belongs to researches in pedestrian behavior understanding [9,10]. Due to its safety irrelevance, detecting ④ is just a task for generalized definition.

3.2. Statistical Reliability of the Function

Camera is the main sensor that is generally used for pedestrian detection. CNN-based neural networks will be used for image processing. Unlike traditional sensor, which directly changes the measured quantity to an electrical signal, the output of NN is a statistical result based on the big training data. Traditional sensors give out a measured value with a confidence interval satisfying gauss distribution, which can be represented with measurement uncertainty. This representation is meaningful in quality management. The prediction from NN is also given confidence, however, this value reflects the goodness of fit of the sample to the fitted function, which cannot match the real probability of a right prediction. Thus, this description is not suitable for reliability.

In computer vision, the following concepts are widely utilized to measure object detection: tp – true positive, fp – false positive, fn – false negative, tn – true negative, Precision (P) – the ratio of tp compared to all detections, Recall (R) – the ratio of tn compared to all ground truths,

$$P = \frac{tp}{tp+fp} \text{ and } R = \frac{tp}{tp+fn} \quad (1)$$

and IOU – Intersection Over Union:

$$IOU = \frac{\mathcal{A}(D \cap T)}{\mathcal{A}(D \cup T)} \quad (2)$$

where \mathcal{A} denotes the area of a region in an image, D is the detected bounding box of the object, and T is the area of the bounding box of the matched ground-truth.

Since these concepts quantitatively describe the quality of single-object detection algorithms in different dimensions, we intend to expand them in defining reliability.

The tp and fp or P and R are not mathematically independent. The predefined $IOU_{threshold}$ determines the boundary between tp and fp . A detection is tp , when $IOU > IOU_{threshold}$, and vice versa. Thus, the $IOU_{threshold}$ should be uniformly defined for all pedestrian detection methods in HAD, so that the basic quality of single detection can be defined, and the P and R can be decoupled. Defining $IOU_{threshold}$

requires discussion and verification. A high $IOU_{threshold}$ leads acceptable detection seen as an error, while a low $IOU_{threshold}$ leads to insufficient precision on perception, the reported detection may deviate greatly from the ground truth.

Considering the reaction (e.g. brake, steer, decelerate) time of vehicles, the distance s between objects and the vehicle must be given for evaluating the reliability. The braking distances for the common speed limits of 30km/h and 50km/h in urban areas are 10m and 29m respectively [11]. We suggest using objects between 10 – 20m and 30 – 50m to evaluate the reliability. Other distances like 100m can be used complementarily for uncommon situations.

Precision P influences the passenger comfort of the vehicle. Algorithms with poor precision mark more positive detections than ground truths. The vehicle status will be more frequently adjusted, unnecessary emergency brakes might be operated. Recall R influences pedestrian safety. Undetected ground truths might lead to collisions and accidents. IOU represents the accuracy of detection. The higher the IOU reaches, the more accurately the velocity and distance can be predicted.

Based on the above arguments, we define the algebraic reliability \mathcal{R}_s :

$$\mathcal{R}_s = \sum k_i \cdot \mathcal{R}_i \quad (3)$$

where $s = 10m, 30m$, possibly 100m

$i = P, R$ and IOU , $\mathcal{R}_P = P$ etc.

k — influence factor, $k_P + k_R + k_{IOU} = 1$

The algebraic reliability describes the functional quality comprehensively. Reliabilities for different distances will be separately represented. Defining k_i is a task similar to defining $IOU_{threshold}$, a balance between comfort and safety must be considered. Regarding the importance of safety, k_R should dominate \mathcal{R}_s , followed by k_P .

Furthermore, we define the reliability vector \mathcal{R} :

$$\mathcal{R} = \begin{pmatrix} P \\ R \\ IOU \end{pmatrix} \quad (4)$$

The reliability vector \mathcal{R} enables the visualized representation. See [Figure 2](#).

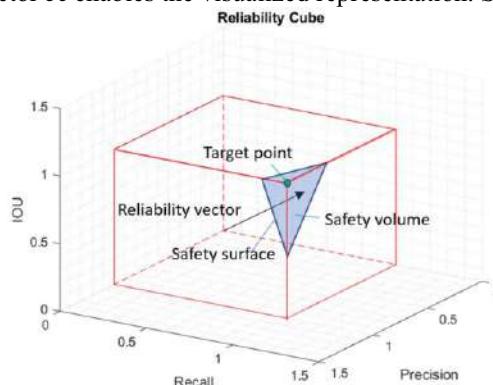


Figure 2. Visualization of reliability vector. A verified acceptable minimum reliability requirement can be shown as safety surface in the figure. The point (1,1,1) is the ideal function. The volume between them is defined as safety volume. Only the pedestrian detection algorithm, whose reliability vector points into this volume, can be seen as functional safe.

Regarding unevenness of above defined k_i , representing the reliability by the magnitude of the vector is not recommended.

Other researchers suggested the utilization of vertical and horizontal pixel deviation between ground truth and detection to judge the confidence level [7]. In this case, the direction influences of bias can be considered, but it complicates the reliability expression.

3.3. Dynamic Reliability of the Function

SSD (Single Shot MultiBox Detector) is nowadays one of the best object detection algorithms, it reaches an average precision (AP) of more than 84.5% for person detection in 512×512 images [12]. The improved algorithm FSSD reaches even 90.2% for the same inputs, with a single NVIDIA 1080Ti GPU, the speed reaches 65.8 FPS for 300×300 input size [13]. Such a score is impressive for normal object detection tasks. Nevertheless, in HAD systems, the missed detections must be handled to satisfy safety requirements. One of the simplest solutions is voting mechanism: Prediction bases on 3 frames, in which the object being in at least two frames as tp marked will be seen as positive. Suppose the AP of FSSD 90.2% is valid for each detection, feed it 3 independent frames, the final AP in this case can be improved significantly to 97.3%.

On the other hand, the pedestrian detection self has limited meaning without the above-mentioned high-level functions like velocity estimation and motion prediction. Such approaches belong to Multi-Object Tracking (MOT) in computer vision. Since HAD is a highly dynamic system, an effective prediction for other surrounded traffic participants can ameliorate both safety and comfort by predictive driving. The standard employed algorithm is tracking-by-detections [14], in which detections in continuous frames are generally basic inputs for tracking functions. The pedestrian detection we defined also includes pedestrian tracking. The deviation of the same object in different frames can influence the prediction quality.

Thus, we propose to validate the dynamic reliability of the function. In MOT there are two widely used metrics to evaluate the tracking quality: CLEAR MOT metrics and ID scores [14]. CLEAR MOT focuses on detection quality, which is similar to the criteria of statistical reliability, while the approach ID scores concentrates on matching cascade. The matching cascade, however, is not a safety-relevant factor in HAD systems. The switch of the IDs of two different pedestrians will not influence the vehicle's motion control. Therefore, both these two metrics are not suitable for dynamic reliability.

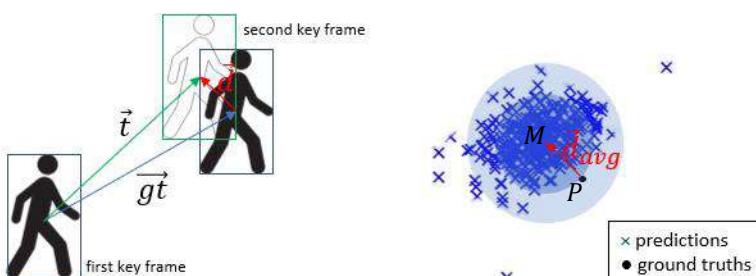


Figure 3. Left: \vec{gt} is the ground truth motion of the object, \vec{t} is the tracking motion of the object. \vec{t} will be fed to other functions. The key frames are the frames, where the object motion will be estimated. \vec{d} is the deviation vector we defined to describe the tracking quality. Right: Move all the endpoints of \vec{gt} to P , cross all the endpoints of \vec{d} , we see the distribution of all deviations of predictions. The average of \vec{d} is \vec{d}_{avg} .

We propose to use the average deviation vector of predictions to describe the prediction accuracy and to use the distribution of the deviations to describe the prediction precision. See [Figure 3](#).

This description can perfectly match the traditional measurement techniques. The \vec{d}_{avg} matches system error and the distance between a cross and M point can be seen as random error. In [\[15\]](#), the pixelwise deviation distribution of YOLOv2 algorithm used in self-driving is investigated. It shows that the x- (horizontal axis) and y-pixel (vertical axis) errors are independently in the x- and y-pixel directions normally distributed, so that the μ and σ are evaluable. Thus, the dynamic reliability can be represented:

$$R_{dym,s} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (5)$$

where $\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$ and s is the above-defined distance.

This definition can benefit the safety functions designed for HAD. Suppose an algorithm has a dynamic reliability μ in driving direction eques 1m and in cross direction eques 0m, σ is 1m, the object is predicted 30m directly in front of the vehicle, it can be concluded that the distance is 95.45% (2σ) farther than 29m and 99.73% (3σ) farther than 28m. Using 28m for motion control is 99.73% suitable for this scenario.

This distribution is not proved valid for other algorithms. Regarding the principle of creating bounding box, we assume that YOLO-like algorithms (One-stage) satisfy normal distribution. The distribution of two-stage methods should be investigated. Even if the deviation is not normally distributed, the corresponding probability density function can represent the dynamic reliability and guide the design of safety functions.

4. Improve the Reliability of safety-related Machine Learning Functions

The reliability we proposed aims to reflect the ability of an accurate perception of the driving surroundings. The more accurate the perception is, the safer and more comfortable the vehicles can drive. Thus, the optimization of reliability plays a role in development. Considering the perception process we propose two directions to improve the reliability, namely sensory and algorithmic.

Camera, the main sensor for pedestrian detection, has inherent insufficiencies like low resolution, color sensitivity and dynamic range because of the limited computing power and the performance of light-sensitive elements. Sensor fusion is an effective approach to counteract the inherent insufficiencies of single sensor. A LiDAR, for example, can sense the environment during the camera blindness when driving in and out of the tunnel. Another widely used method to improve functional safety is redundancy. Additional fault tolerance can generally improve a safety integrity level. It is also a practical solution to the blindness of specific camera caused by specific sunlight angles.

Furthermore, the prediction algorithms like LSTM (Long Short-Term Memory) and Kalman Filter can provide a predicted location of the covered object, or when the camera is blinded. With the help of pedestrian behavior understanding, which is based on Bayesian networks, a multi prediction for pedestrian motion can be constructed. The prediction will reference the past location and velocity as well as the intention of the pedestrian. The interaction between vehicles and pedestrians will be built like it in reality.

5. Conclusions and Future Work

In this work, we investigated the evolution of safety requirements with the development of autonomous driving. We expounded on the insufficiency of existing standards applying in HAD and proposed the application of functional reliability. To evaluate and validate the reliability of machine learning based HAD functions, we put forward the definition of a reliability function under predefined constraints to quantitatively describe the reliability of pedestrian detection regarding safety and comfort. At last, we indicated some practical methods, which can improve the defined reliability.

In the future, we will implement and verify some sensor fusion and algorithm methods we mentioned. We are striving to realize a safe, reliable and robust pedestrian detection proposal in HAD. We suggest the automated driving community and automobile industry taking closely part in creating validation datasets and defining key parameters in HAD to accelerate the standardization of HAD vehicles development.

References

- [1] Burton S, Gauerhof L, Heinzemann C. Making the Case for Safety of Machine Learning in Highly Automated Driving. In: Tonetta S, Schoitsch E, Bitsch F, editors. Computer Safety, Reliability, and Security: SAFECOMP 2017 Workshops, ASSURE, DECSoS, SASSUR, TELERISE, and TIPS, Trento, Italy, September 12, 2017, Proceedings. Lecture Notes in Computer Science. Vol 10489. Cham: Springer International Publishing; 2017. p. 5–16.
- [2] Grigorescu S, Trasnea B, Cocias T, Macesanu G. A Survey of Deep Learning Techniques for Autonomous Driving. *J. Field Robotics* 2020;37:362–86.
- [3] ISO 26262:2011: Road vehicles - Functional safety. Switzerland; 2011.
- [4] ISO/PAS 21448:2019: Road vehicles — Safety of the intended functionality. Switzerland; 2019.
- [5] Iannino A, Musa JD. Software Reliability. In: Yovits MC, editor. *Advances in Computers*. New York: Academic Press; 1990. p. 85–170.
- [6] Leveson NG. Engineering a safer world: Systems thinking applied to safety. *Engineering systems*. Cambridge, Massachusetts, London, Englang: The MIT Press; 2017.
- [7] Gauerhof L, Munk P, Burton S. Structuring Validation Targets of a Machine Learning Function Applied to Automated Driving. In: Hoshi M, Seki S, editors. *Developments in Language Theory: 22nd International Conference, DLT 2018, Tokyo, Japan, September 10–14, 2018, Proceedings. Theoretical Computer Science and General Issues*. Vol 11088. Cham: Springer International Publishing; 2018. p. 45–58.
- [8] Bojarski M, Chen C, Daw J, Dégirmenci A, Deri J, Firner B, et al. The NVIDIA PilotNet Experiments; 2020.
- [9] Rasouli A, Kotseruba I, Tsotsos JK. Are They Going to Cross? A Benchmark Dataset and Baseline for Pedestrian Crosswalk Behavior. In: 2017 IEEE International Conference on Computer Vision Workshops: ICCVW 2017: 22–29 October 2017, Venice, Italy: proceedings. Piscataway, NJ: IEEE; 2017. p. 206–13.
- [10] Rasouli A, Kotseruba I, Tsotsos JK. Understanding Pedestrian Behavior in Complex Traffic Scenes. *IEEE Trans. Intell. Veh.* 2018;3:61–70.
- [11] Greibe P. Braking distance, friction and behaviour Findings, analyses and recommendations based on braking trials 2007.
- [12] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. SSD: Single Shot MultiBox Detector. In: Leibe B, Matas J, Sebe N, Welling M, editors. *Computer Vision - ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I*. Lecture Notes in Computer Science. Vol 9905. Cham, s.l.: Springer International Publishing; 2016. p. 21–37.
- [13] Li Z, Zhou F. FSSD: Feature Fusion Single Shot Multibox Detector; 2017.
- [14] Ciaparrone G, Luque Sánchez F, Tabik S, Troiano L, Tagliaferri R, Herrera F. Deep learning in video multi-object tracking: A survey. *Neurocomputing* 2020;381:61–88.
- [15] Miethig B, Huangfu Y, Dong J, Tjong J, Mohrenschmidt Mv, Habibi S. A Novel Method for Approximating Object Location Error in Bounding Box Detection Algorithms Using A Monocular Camera. *IEEE Transactions on Vehicular Technology* 2021;1.

Analysis of Signal Processing Methods for Formation of Radar Range Profiles

Nguyen VAN LOI¹, Le THANH SON, Tran TRUNG KIEN,
Tran VAN TRUONG, Tran VU HOP

Radar Center, Viettel High Technology Industries Corporation, Hanoi, Vietnam

Abstract. A range profile (RP) is a vector of reflected powers from a target by the range direction and is used for the purpose of target recognition. In this paper, a problem of formation of RPs is investigated. The well-known difference operator and window-based methods are analyzed with data from a coastal surveillance radar. The drawbacks of those methods are shown. Then, a new method is presented to improve the performance of formation of RPs.

Keywords. Radar signal processing, radar range profile, difference operator method, window-based method, automatic target recognition

1. Introduction

Radar signal processing contains the consecutive algorithms that focus on target detection and formation of target profiles (range or azimuth). A range (or azimuth) profile of a target is a power vector reflected from the target along the radial (or azimuth) direction (see, Figure 1).

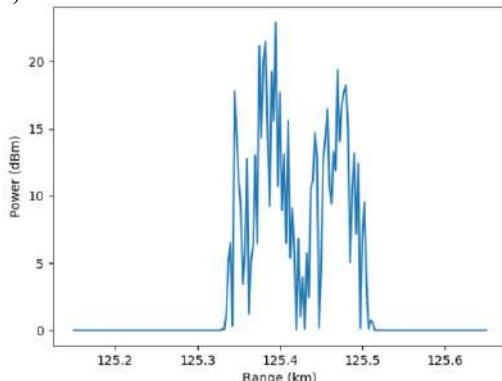


Figure 1. A RP of a cargo ship of length 225m, which is approaching directly in a clear environment region. From RPs important features (such as ship length and width, structure of dominant scattering ...) can be extracted and used for the target recognition [1],[5],[6]. However, for targets moved in a sea clutter environment, the signal-to-clutter ratio (SCR) is low, and therefore, the RPs of a target will be longer than normal (see Figure 2).

¹ Corresponding Author. E-mail: loinv12@viettel.com.vn

In this paper, we deal with the problem of formation of RPs of a target in case of low SCR. Two common approaches for the formation of RPs of a target are considered and analyzed. The tests with data from a coastal surveillance radar show that the window-based method is more accurate than the difference operator method. However, the original window-based method has many disadvantages which are highly sensitive to low SCR. Then, a proposed method on improving the original window-based method is given.

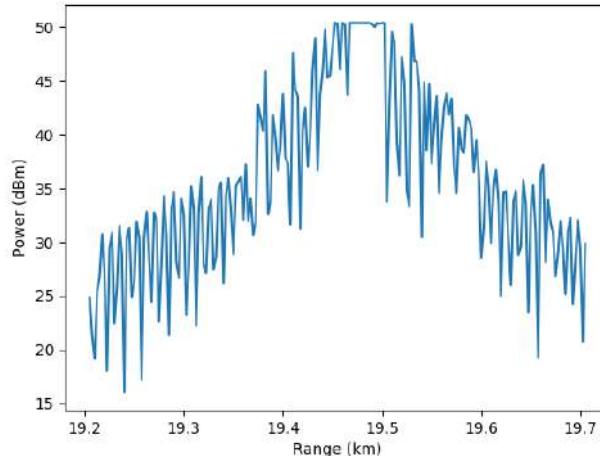


Figure 2. A RP of a cargo ship, which is approaching directly in a sea clutter region. The RP has a length of 500m while the true length of the ship is 181m.

The paper is organized in the following way. In the next section, we give a brief description of the difference operator and window-based methods. The tests and analysis are given in Section 3. In section 4 a modification of the original window-based method is presented. The last section deals with the conclusion and future works.

2. Related works

First of all, let us recall the basic steps of radar signal processing [2],[3],[4]. Figure 3 represents the radar block diagram with Doppler processing.

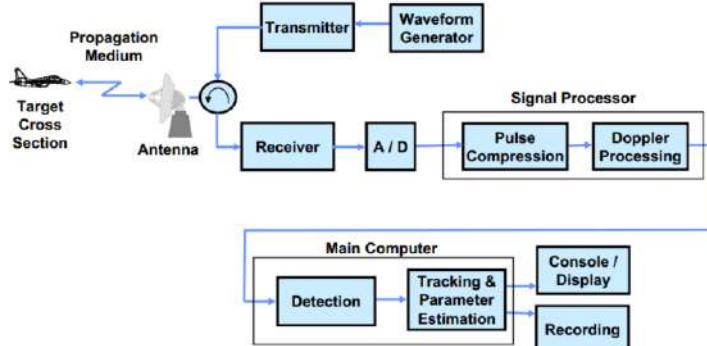


Figure 3. Radar block diagram [4]

The Doppler processing is done by using FFT (Fast Fourier Transform) and its output is considered as the *target's raw range profiles*.

A simple method of formation of RPs is the thresholding method [5]. However, this method is highly sensitive to noise and interference [6]. Then, an adaptive difference operator method is introduced to improve the thresholding method [6]. The idea of the difference operator method (see Figure 4) is the following [6]:

Let $X = [x_1, x_2, \dots, x_N]^T$ be a raw RP of a target, where N denotes the number of reflected sampling points (the number of range cells in the same radial direction that reflect the radar transmitted signal). Set

$$W[k] = \begin{cases} -1, & -M \leq k < 0 \\ 1, & 0 \leq k < M \\ 0, & \text{others} \end{cases} \quad (1)$$

The RP of the target is a vector $Y = [y_1, y_2, \dots, y_{N-2M+1}]^T$ defined as:

$$y_i = \begin{cases} \left(\frac{E_{i2}}{E_{i1}}\right) \times (E_{i2} - E_{i1}), & E_i \geq 0 \\ \left(\frac{E_{i1}}{E_{i2}}\right) \times (E_{i1} - E_{i2}), & E_i < 0 \\ (i = 1, 2, \dots, N - 2M + 1) \end{cases} \quad (2)$$

where,

$$E_i = \sum_{j=0}^{2M-1} x_{i+j} \times W[-M + j] \quad (3)$$

$$E_{i1} = \left| \sum_{j=0}^{M-1} x_{i+j} \right| \quad (4)$$

$$E_{i2} = \left| \sum_{j=M}^{2M-1} x_{i+j} \right| \quad (5)$$

$$\begin{cases} M = m_1, & SNR < S_1 \\ \frac{SNR - m_1}{S_2 - S_1} = \frac{M - m_1}{m_2 - m_1}, & S_1 \leq SNR \leq S_2 \\ M = m_2, & SNR > S_2 \end{cases} \quad (6)$$

$$SNR = 10 \times \log_{10} \left\{ \frac{(\max(x_i)_{i=1}^N)^2}{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \right\}, \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (7)$$

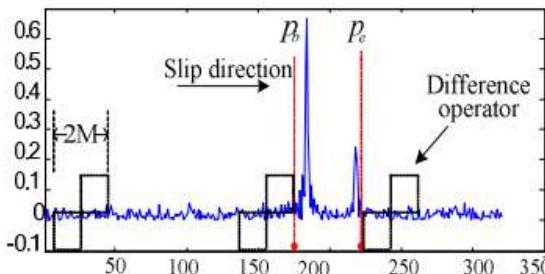


Figure 4. Illustration of difference operator method [6]

Another approach for formation of RPs of a target is the window-based method [1],[7]. This method starts with N_{st} windows, each of them is a range cell containing one of the N_{st} strongest peaks of a raw RP. For each window, at each step, the power sums of three adjacent cells on the left and right side of the window are estimated. Then, the window is extended by the next adjacent cell on the side with the higher power sum. Moreover, the relative power

$$\Delta_i = \frac{p_i}{p_i + \sum p_{win}} \quad (8)$$

is determined, where p_i is the power in the next adjacent cell and $\sum p_{win}$ is the power sum of all cells in the window. The process is repeated until the value Δ_i stays permanently below a certain threshold:

$$\Delta_i < \Delta_{thres} \quad (9)$$

The value of certain threshold $\Delta_{thres} = 0.01$ was found to be effective [1]. So, the classical window-based method is characterized by the number of created windows, the rule for extending the window and the ending condition:

- *Creating the windows:* from N_{st} strongest peaks of a raw RP of a target, N_{st} windows (one range cell for each window) are created.
- *Extending the windows:* for each window, step by step the window is extended by the next adjacent cell on the side with the higher radar-cross-section.
- *Ending condition:* The process is repeated until the inequality

$$\Delta_i = \frac{p_i}{p_i + \sum p_{win}} < \Delta_{thres}$$

stays permanently.

3. Analysis of methods

For analysis of the methods mentioned above we will use the AIS information as ground truths to find the relative errors of estimated lengths of vessels and their standard deviation:

$$Err = \frac{1}{N} \sum_{j=1}^N \left| 1 - \frac{L_{estj}}{L_j} \right| \quad (10.a)$$

$$Std = \text{standard deviation} \left\{ \left| 1 - \frac{L_{estj}}{L_j} \right| \right\} \quad (10.b)$$

where, Err is the relative error; Std is the standard deviation of the relative error; N is the total number of RPs using for evaluation; L_j is the true length of the target corresponding to the j -th RP and L_{estj} is the estimated length of target estimated from the j -th RP:

$$L_{estj} = \frac{\text{length of } j\text{-th RP}}{\cos(\text{aspect angle})} \quad (11)$$

The data for the analysis (see Table 1) is collected from a coastal surveillance radar with the following parameters:

Table 1. Radar parameters

Parameter	Value
Frequency	X-band
Range resolution	3m
Sampling frequency (f)	120MHz

Two tests are investigated.

Test 1: The dataset contains 53 ships (32 cargo ships and 21 tanker ships) with 5600 profiles (3263 profiles of cargo ships and 2337 profiles of tanker ships). All ships moved in the area of radius of 90km slant range and with the aspect angles in the interval $aspAngSet$ (see Figure 5):

$$aspAngSet = [0^\circ; 60^\circ] \cup [120^\circ; 240^\circ] \cup [300^\circ; 360^\circ]$$

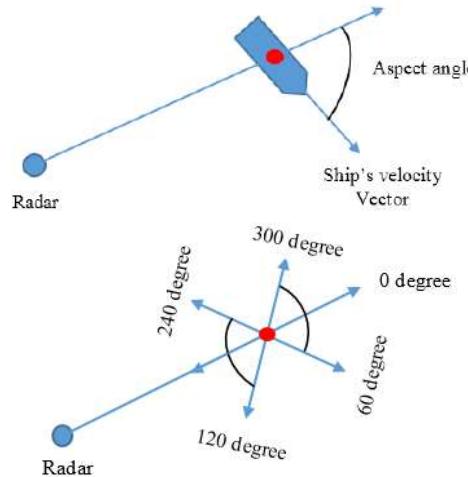


Figure 5. Aspect angle of a ship (above) and considered aspect angles ($aspAngSet$) for Test 1 (below)
The parameters for the difference operator method are obtained by the same way as given in [6], i.e.

Table 2. Parameters for difference operator method

Parameter	Value
S_1	10
S_2	30
m_1	$= 15f/c = 6$ (c is the speed of light)
m_2	$= 7f/c \approx 3$

The parameters for the window-based method are the same with that one given in [1, p.94], i.e.

Table 3. Parameters for window-based method

Parameter	Value
N_{st}	X-band
Δ_{thres}	3m

The test results are given in Table 4.

Table 4. Test result

Difference operator method		Window-based method	
Err	Std	Err	Std
1.5	1.97	0.229	0.21

Test 2: The dataset contains 8 ships (4 cargo ships and 4 tanker ships) with 619 profiles (218 profiles of cargo ships and 401 profiles of tanker ships). All ships moved in the area of 90km slant range and with the aspect angles in the interval $aspAngSet$:

$$aspAngSet = [0^\circ; 10^\circ] \cup [170^\circ; 190^\circ] \cup [350^\circ; 360^\circ]$$

This test is taken for the case when the ships moved directly toward to the radar station or moved directly far away from the radar station. The parameters for difference operator and for window-based methods are the same as given in Tables 2 and 3. The test results are given in Table 5.

Table 5. Test result

Difference operator method		Window-based method	
Err	Std	Err	Std
0.88	0.4	0.175	0.194

The above tests show that the window-based method has a better performance than the difference operator method. However, the original window-based method has some drawbacks that are:

- The use of N_{st} ($N_{st} > 1$) strongest peaks of the RP for “Creating the windows” may lead to the wrong result due to the case of multiple targets or the case of low SCR. In these cases, selected peaks may be clutter or different closed targets (see Figure 6).

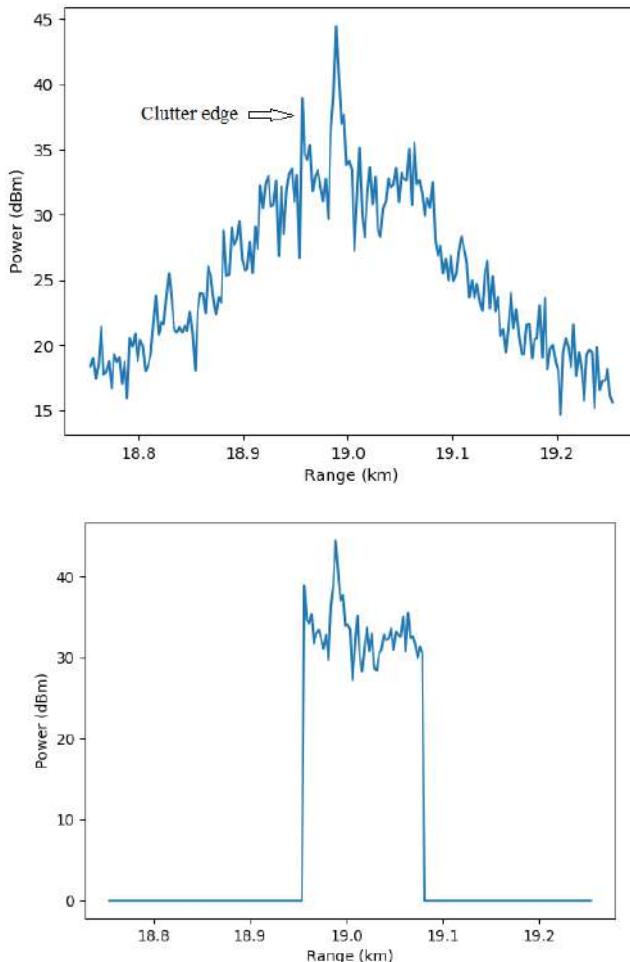


Figure 6. A raw RP of a tanker (length = 127m and aspect angle $\approx 304^\circ$) in a non-homogenous clutter region (above) and the RP formed using the window-based method (below) with $N_{st} = 3$. In this example, one of the selected peaks is the clutter edge and therefore the estimated length of the target is approximately 225m.

Relative error $\approx 77.16\%$.

- The use of power sum of three cells in the left and the right sides of the window to extend the window size. This drawback in case of non-homogeneous clutter may lead to extend the window size into the clutter region (see Figure 7).

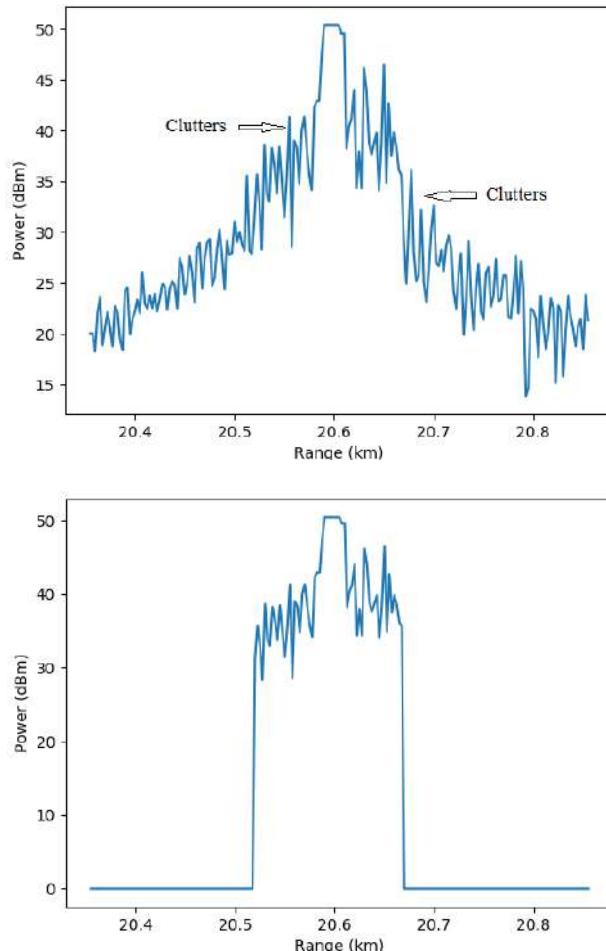


Figure 7. A raw RP of a tanker (length = 127m and aspect angle $\approx 321^\circ$) in a non-homogenous clutter region (above) and the RP formed using the window-based method (below). The estimated length of the target is approximately 183m. Relative error $\approx 44\%$.

- The use of a constant value Δ_{thres} for ending condition for all types of ships. This problem may lead to the extension of the estimated length of target (see Figure 8).

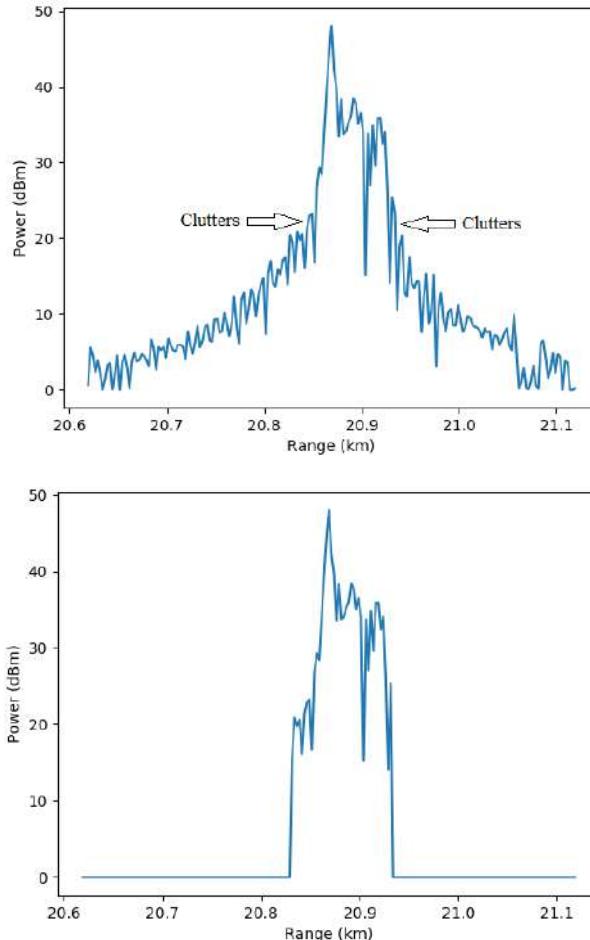


Figure 8. A raw RP of a tanker (length = 63m and aspect angle $\approx 20^\circ$) in a clutter region (above) and the RP formed using the window-based method with $\Delta_{thres} = 0.02$ (below). The estimated length of target is approximately 109m. Relative error $\approx 73\%$.

4. Proposed method and test result

4.1. Proposed method

To overcome the drawbacks analyzed above we propose a modification of the window-based method which contains the following steps:

Step 1: Create a window with only one range cell at target's centroid (i.e. $N_{st} = 1$). Let $X = [x_1, x_2, \dots, x_N]^T$ be a raw RP of a target, where x_i is the power reflected from i -th range bin (or equivalent from a range r_i). Then, the target's centroid range $r_{centroid}$ is defined by (see Figure 9):

$$r_{centroid} = \frac{\sum_{i=1}^N x_i * r_i}{\sum_{i=1}^N x_i} \quad (12)$$

The target's centroid bin ($centroid_{bin}$) is the bin i_* such that $r_{i_*} \approx r_{centroid}$.

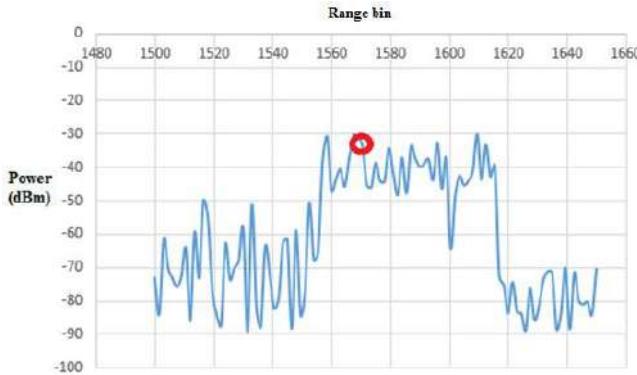


Figure 9. Example of a target's raw profile and its centroid range bin (at the circle position)

Step 2: Extend the window

- Take $2M$ cells near the $centroid_{bin}$ (M adjacent cells on the left and M adjacent cells on the right of the window). The value of M depends on the radar range resolution ΔR and can be chosen by the rule:

$$M \approx \frac{15(m)}{\Delta R (m)} \quad (13)$$

For example, if a radar has range resolution 1,5m we can take $M = 10$. The distance of $2M$ bin is 30m. This distance is suitable for almost marine targets. If we take less range cells, statistics are not guaranteed. If more range cells are taken, the lengths of many types of marine targets may be exceeded.

- Find the means M_{left} and M_{right} of the left M adjacent cells and right M adjacent cells near the $centroid_{bin}$, respectively.
- Calculate the means: λ_{left} and λ_{right} of the bins $[x_1, x_2, \dots, x_M]$ and $[x_{N-M+1}, x_{N-M+2}, \dots, x_N]$, respectively.
- Calculate the values: $M_{left} - \lambda_{left}$ and $M_{right} - \lambda_{right}$. These values show the power level of the target relative to the background noise on the left and right sides of the window (see Figure 10).
- Extend the window one adjacent cell to the left of the window if:

$$M_{left} - \lambda_{left} \geq M_{right} - \lambda_{right} \quad (14)$$

Otherwise, extend the window one adjacent cell to the right of the window.

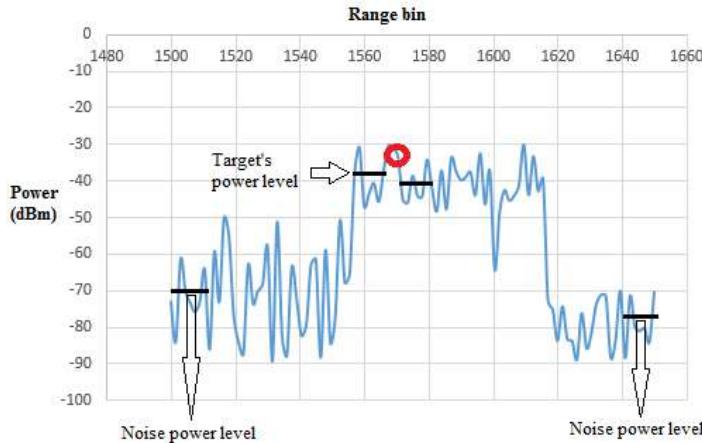


Figure 10. Example of power levels of the target relative to the background noise on the left and right sides of the window

Step 3: End of the process

- The process of extending the window is ended when condition

$$M_{left} \text{ (respectively, } M_{right}) < \Delta_{thres} \times \sum p_{win} \quad (15)$$

always satisfies, where $\sum p_{win}$ is the sum of powers in all cells of the window; Δ_{thres} is a threshold which depends on the length of the window $l(win)$:

$$\Delta_{thres} = \begin{cases} 0.02 & \text{for } l(win) < \frac{60}{\Delta R} \\ 0.01 & \text{for } l(win) < \frac{120}{\Delta R} \\ 0.005 & \text{for } l(win) > \frac{120}{\Delta R} \end{cases} \quad (16)$$

The formula (16) means that for targets of small and medium sizes (target's length $\leq 60m$) the threshold $\Delta_{thres} = 0.02$ is used, for targets of large size ($60m < \text{target's length} \leq 120m$) the threshold $\Delta_{thres} = 0.01$ is used and for targets of extremely large size (target's length $> 120m$) the threshold $\Delta_{thres} = 0.005$ is used.

4.2. Test results

We use the same datasets and parameters used in Test 1 and Test 2 in section 3 for the test and comparison of the proposed method. The test results are given in Tables 6 and 7, respectively.

Table 6. Test result

Window-based method		Proposed method	
Err	Std	Err	Std
0.229	0.21	0.113	0.112

Table 7. Test result

Window-based method		Proposed method	
Err	Std	Err	Std
0.175	0.194	0.069	0.06

As given in (10.a) the value *Err* is the mean value of all relative errors estimated from all test RPs. Figure 11 represents the cumulative probabilities of relative errors of the window-based and proposed methods.

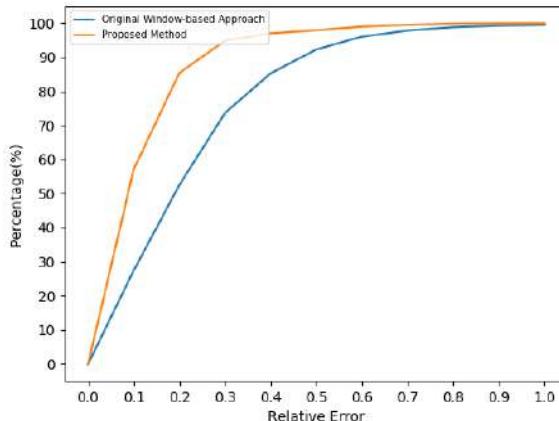


Figure 11. Cumulative probability of relative errors. The curve above corresponds to the proposed method.

5. Conclusion and future works

The paper presented the detailed analysis of two popular signal processing methods for range profile formation that are the difference operator method and the window-based method. The tests shown that although the window-based method has a better performance than the difference operator method, it contains some disadvantages such as the use of $N_{st} > 1$ peaks, the use of reflected powers in three cells in the left and the right sides to extend the window size and the use of a constant value Δ_{thres} for the end of process.

To avoid these drawbacks, we proposed a modification of the original window-based method that uses only one peak (at the target's centroid) for creating a window; uses more statistical information of the reflected powers for extending the window size and uses an adaptive threshold for ending process. The tests shown that the proposed method achieves better performance than the original window-based method.

In future works we will investigate the problem of formation of RPs in case of low SNR (signal-to-noise ratio), multiple closed targets and other interferences.

References

- [1] David Blacknell, Hugh Griffiths (Edi.), *Radar Automatic Target Recognition and Non-Cooperative Target Recognition*, The Institution of Engineering and Technology, 2013.
- [2] Merrill I. Skolnik (Edi.), *Radar Handbook*, McGraw-Hill, 3rd Edition, 2008.
- [3] Mark A. Richards, *Fundamental of Radar Signal Processing*, McGraw-Hill, 2005.

- [4] Robert M. O'Donnell, *Introduction to Radar Systems*, Spring 2007, Massachusetts Institute of Technology: MIT OpenCourseWare, <http://ocw.mit.edu>.
- [5] Q. Ma, *Feature Study for High-Range-Resolution based Automatic Target Recognition: Analysis and Extraction*, Doctor Thesis, OHIO: The Ohio State University, 2001.
- [6] K. Liao, G. Gui, Z. Chen, W. Yang, "High resolution range profile based extraction of radar target length", *Int. J. of the Phys. Sci.*, Vol. 6(23), pp. 5503-5510, Oct. 2011.
- [7] H. Schimpf and H.-H. Fuchs, "The Influence of Multipath on the Classification of Ships", *Proceedings of International Radar Symposium IRS 2009*, Hamburg, Germany, 9–11 September 2009.

Three Term Weighted Type Fractional Fourier Transform Based Generalized Hybrid Carrier and Its Application into PAPR Suppression

Yong LI^{a,b,1}, Zhiqun SONG^{a,b}, Teng SUN^b, and Bin WANG^{a,b}

^a *Science and Technology on Communication Networks Laboratory, China*

^b *The 54th Research Institute of China Electronics Technology Group Corporation, China*

Abstract. To suppress the peak to average power ratio (**PAPR**) of wireless communication based upon multi-carrier system. We, in this paper, proposed the three term weighted type fractional Fourier transform (3-WFRFT) based generalized hybrid carrier (GHC) system. We first provide the definition of 3-WFRFT. Moreover, some useful properties of 3-WFRFT have been presented, in this paper, which will helpful to comprehend the novel 3-WFRFT transform. Furthermore, we take **PAPR** of the proposed algorithm, in comparison with orthogonal frequency division multiplexing (OFDM) system and single carrier modulation (SC) system under typical complementary cumulative density function (**CCDF**) level. It would be demonstrated that, from some numerical simulations, the proposed 3-WFRFT based GHC performs better than OFDM system and will be useful to reduce the **PAPR** level.

Keywords. Three term weighted type fractional Fourier transform (3-WFRFT), peak to average power ratio (**PAPR**), generalized hybrid carrier (GHC))

1. Introduction

Recently, the multi-carrier system seems popularity over wireless communications, such as digital video broadcasting (DVB) and digital audio broadcasting (DAB), wireless local area networks (WLAN) [1–4] and the 5th generation mobile communication technology (5G) and so on. Moreover, given the less complexity and high spectral efficiency, the multiple carrier system has been some standard for most of wireless communications. However, due to the signal envelope and high fluctuation of multi-carrier, the peak to average power ratio (**PAPR**) is high, which will influence the application over wireless communication. Specially, when the non-linear power amplifier (PA) has been used on wireless communication, the high **PAPR** will cause obviously degradation in performance.

¹Corresponding Author: Science and Technology on Communication Networks Laboratory; The 54th Research Institute of China Electronic Technology Group Corporation, Shijiazhuang, Hebei, 050081, China, E-mail: young_li_54@126.com

There are many studies about **PAPR** suppression. In [5], the author proposed the Hadamard recursive carrier interferometry (HRCI) codes and diagonal recursive carrier interferometry (DRCI) codes to suppress the **PAPR** under OFDM systems. Jiang in [6] provided the Iterative Compensation algorithm to **PAPR** suppression. In order to reduce the complexity, Cheng in [7] proposed C-DSLM algorithm. It can be observed, from the studies recently [5–7], the communication scheme is still the DFT based OFDM system. However, the **PAPR** of DFT for wireless signals will be prohibitively high. We need a technology, which can both inherit the superior of OFDM and suppress **PAPR**, to improve the performance of wireless communications.

To this end, we in this paper, proposed the three term weighted type fractional Fourier (3-WFRFT) based generalized hybrid carrier (GHC) system to mitigate **PAPR** of wireless communications. The 3-WFRFT based GHC system, coupling with the components of single carrier modulation (SC) and orthogonal frequency division multiplexing (OFDM) signals. It is expected the **PAPR** level is lower than OFDM signals. It is demonstrated via numerical simulations, that 3-WFRFT based GHC system performs better than OFDM systems under the same channel conditions.

The rest of this paper is listed as follows. We first present The background theory to derive the 3-WFRFT and its properties in Section 2. The GHC based upon 3-WFRFT, will be given in Section 3. Furthermore some simulations based upon PAPR and discussions have been considered in Section 4. Finally, the conclusion of whole paper has been shown in Section 5.

2. The Background Theory

2.1. 3-Weighted Type Fractional Fourier Transform (3-WFRFT)

There are many various forms for Multi-Weighted type Fractional Fourier Transform (M-WFRFT) according to [8–13]. However, the term M can only be $4N, N \in \mathbb{Z}$ or $M > 4, M \in \mathbb{Z}$ and so on. To this end, we will proposed the 3-WFRFT.

In this section, we will provide the concept of 3-WFRFT, which is based upon 4-WFRFT. In additional, M-WFRFT ($M > 4$) proposed in [14], is also based upon 4-WFRFT. However, the relationship between M-WFRFT ($M > 4$) and 4-WFRFT has been derived in [14], but not that of 3-WFRFT and 4-WFRFT. We, in this paper, will provide the concept of 3-WFRFT and some properties, which maybe helpful to solve the problem.

The symbol of N -length $\mathbf{S}^T = \{s_1, s_2, \dots, s_{N-1}, s_N\}$, and its α -order ($\alpha \in [0, 1]$) 3-WFRFT can be defined as:

$$\mathbf{X} = \mathfrak{F}_3^\alpha[\mathbf{S}] = \mathbf{W}_3^\alpha \mathbf{S}^T \quad (1)$$

where \mathbf{W}_3 is called the $N \times N$ 3-WFRFT matrix, and can be expressed as:

$$\mathbf{W}_3^\alpha = \sum_{l=0}^2 B_l(\alpha) \mathbf{W}_4^{\frac{4}{3}l} \quad (2)$$

$\mathbf{W}_4^{\frac{4}{3}l}$ ($l \in \mathbb{Z}, l \in [0, 2]$) is the $N \times N$ 4-WFRFT matrix, consisting of $\mathbf{F}^h, h = 0, 1, 2, 3$. Here, \mathbf{F}^h is the h times of discrete Fourier transform (DFT) matrix. And

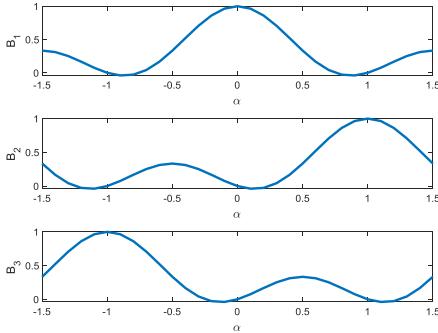
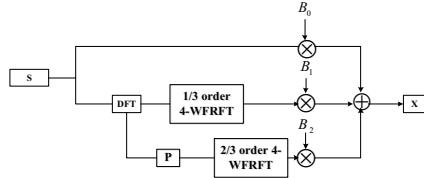
Figure 1. The amplitude of $B_l(\alpha)$.

Figure 2. The Implement of 3-WFRFT

$[\mathbf{F}]_{m,n} = e^{-j2\pi mn/N}$, where $m, n \in \mathbb{Z}$, $m, n \in [0, N-1]$. Specially, \mathbf{F}^h will be degenerated to unit matrix if $l = 0$, moreover \mathbf{F}^h will become the DFT matrix \mathbf{F} when $h = 1$. The 4-WFRFT matrix, i.e., $\mathbf{W}_4^{\frac{4}{3}l}$, can be as:

$$\mathbf{W}_4^{\frac{4}{3}l} = \sum_{h=0}^3 A_h \left(\frac{4}{3}l \right) \mathbf{F}^h, \quad (3)$$

$A_h(\beta)$ ($h \in \mathbb{Z}, h \in [0, 3]$) is the 4-WFRFT weighted coefficient, and can be described as:

$$A_h(\beta) = \frac{1}{4} \frac{1 - e^{-j2\pi(\beta-h)}}{1 - e^{-j2\pi(\beta-h)/4}} \quad (4)$$

And $B_l(\alpha)$ is termed as the 3-WFRFT coefficient, and can be expressed as:

$$B_l(\alpha) = \frac{1}{3} \sum_{n=0}^2 \exp\left[-\frac{2\pi j n(\alpha - l)}{3}\right], l = 0, 1, 2. \alpha \in [-1.5, 1.5] \quad (5)$$

The amplitude of $B_l(\alpha)$ can be described as Figure 1. From Figure 1, it can be observed that, when $\alpha = 0$, only the $B_0 = 1$. The 3-WFRFT matrix will be degenerated to unit matrix.

2.2. Property of 3-WFRFT

Property 1 For any α and β , the additive of 3-WFRFT holds:

$$\mathbf{W}_3^{\alpha+\beta} = \mathbf{W}_3^\alpha \mathbf{W}_3^\beta = \mathbf{W}_3^\beta \mathbf{W}_3^\alpha \quad (6)$$

Property 2 For any α , the periodicity of 3-WFRFT holds:

$$\mathbf{W}_3^\alpha = \mathbf{W}_3^{\alpha+3} \quad (7)$$

Property 3 For any α , the reversibility of 3-WFRFT holds:

$$\mathbf{W}_3^\alpha \mathbf{W}_3^{-\alpha} = \mathbf{W}_3^\alpha (\mathbf{W}_3^\alpha)^T = \mathbf{I} \quad (8)$$

where $(.)^T$ is the transpose of matrix, and \mathbf{I} is the unit matrix.

The proof processing of three properties is straight and omitted here. Specially, 3-WFRFT matrix is the orthogonal one based upon the reversibility property, which is easy to be implemented in reality.

3. The Implement of 3-WFRFT and Its PAPR Analysis

3.1. The Implement of 3-WFRFT

The implement of 3-WFRFT, based 4-WFRFT, DFT and permutation matrix, is demonstrated in Figure.2. The 4-WFRFT operation has been provided in [15], DFT operation is based upon fast Fourier transform (FFT).

The implement of 3-WFRFT is simple. And also, we can find that the 3-WFRFT signal can be consisted by single carrier (SC), weighted type fractional Fourier transform domain signals. In fact, the 4-WFRFT is consisted of single carrier and orthogonal frequency division multiplexing (OFDM) signals. Thence, 3-WFRFT can also be constituted by SC and OFDM signal, since the 3-WFRFT is based upon 4-WFRFT in this paper.

Based upon the previous reasons, it interests us that what about the peak to average power ratio (PAPR) of 3-WFRFT, comparing with OFDM and SC. Thence, we propose the PAPR of 3-WFRFT based generalized hybrid carrier (GHC) in the next section.

3.2. The PAPR of GHC

The peak to average power ratio (**PAPR**) of GHC based with 3-WFRFT signal $X = \{x_0, x_1, \dots, x_{N-2}, x_{N-1}\}$ in a given frame is defined as follows:

$$\mathbf{PAPR}_X = \frac{\max_{n \in [0, N-1]} |x_n|^2}{E|x_n|^2} \quad (9)$$

PAPR, in general, can be considered as random variable and be given by the complementary cumulative density function (CCDF)[3]. It can be defined the probability value that **PAPR** exceeds an preset value $\mathbf{PAPR}_0 > 0$.

$$\mathbf{CCDF} = \Pr(\mathbf{PAPR} > \mathbf{PAPR}_0) \quad (10)$$

Obviously, the **PAPR** of OFDM and SC can be also described as aforesaid. However, the structures of signal for GHC, OFDM and SC system are different. Thence, the **PAPR** level of them will be different. Based upon the above, we will take some numerical simulations to compare the **PAPR** of GHC, OFDM and SC signals in the next section.

4. Simulations and Discussions

To verify the **PAPR** performance superiority of GHC system, in comparison to the SC and OFDM system, we present the numerical simulation under the additive white gaussian noise (AWGN) channels. Here, we set the testing signal are phase-shift keying (PSK) and quadrature amplitude modulation (QAM) signal. Without loss of generalization, we choose quadrature phase-shift keying (QPSK) and 16QAM as the numerical simulation signals.

We can observe that, from Figure 3, given that $CCDF = 10^{-2}$, the **PAPR** of GHC based on 3-WFRFT with $\alpha = 0.2$ is about 5dB, $\alpha = 0.4$ is 2dB, and $\alpha = 0.4$ is 1dB superior over that of OFDM signal level with QPSK modulation. The **PAPR** of SC with QPSK modulation is nearly zero. With the α ($\alpha > 0.7$) increasing, the **PAPR** of GHC with 3-WFRFT signal level is almost the same.

It can be demonstrated that, from the Figure 4, given that $CCDF = 10^{-2}$, the **PAPR** of GHC based on 3-WFRFT with $\alpha = 0.2$ is about 4dB, $\alpha = 0.4$ is 2dB, and $\alpha = 0.5$ is 0.5dB superior over that of OFDM signal level with 16QAM modulation. The **PAPR** of SC with 16QAM modulation is nearly zero. The reason why **PAPR** of GHC is superior over OFDM maybe that, the GHC with 3-WFRFT will be consisted of SC and OFDM signals. When α is small than 0.5, the SC signal components is higher than OFDM signals. Thence, the **PAPR** property is near to that of SC system. And also, the 3-WFRFT will be near to OFDM signal with the transform order α increasing.

The 3-WFRFT based GHC system proposed in this paper, can be applied into **PAPR** suppression, especially in the wireless communication systems with non-linear high non-linear power amplifier (HPA), where the proposed 3-WFRFT will improve the **PAPR** suppression performance.

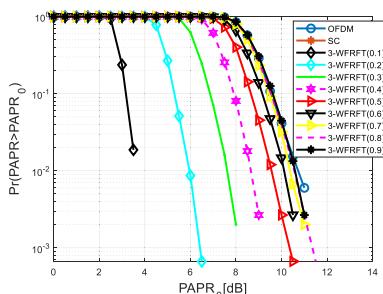


Figure 3. The **PAPR** of GHC ,SCM and OFDM systems with QPSK signals

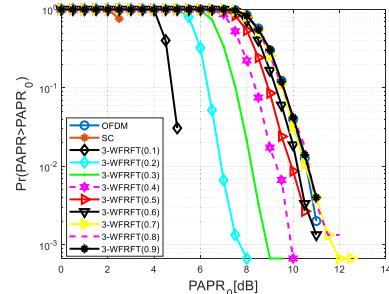


Figure 4. The **PAPR** of GHC ,SCM and OFDM systems with 16QAM signals

5. Conclusion

In this paper, we proposed GHC system based upon 3-WFRFT to reduce the **PAPR** of communication system. The properties of 3-WFRFT has been explored in this paper. It can be demonstrated, from the numerical simulations, the **PAPR** of GHC system with 3-WFRFT is superior over OFDM system. In the future, we will explore 3-WFRFT with more applications over wireless communication.

Acknowledgements

The work is supported by the fund of the National Key Research and Development Program of China (under grant 254). And this work is also supported by Science and Technology on Communication Networks Laboratory.

References

- [1] T. Jiang and Y. Wu: An overview: Peak-to-Average power ratio reduction techniques for OFDM signals, *IEEE Trans. Broadcast.* 2008 **54**(2): 257–268.
- [2] S. B. Weinstein: The history of orthogonal frequency-division multiplexing [history of communications], *IEEE Communications Magazine.* 2009 **47**(11): 26–35.
- [3] D.D. Prasad and M.V Nageswara Rao: A Hybrid Companding Transform Technique for PAPR Reduction of OFDM Signals, *International Conference on Electromagnetic Interference and Compatibility*, 2015 **1**, 263-268.
- [4] H. Joo: New PTS Schemes for PAPR Reduction of OFDM Signals Without Side Information, *IEEE Transaction on Broadcasting*, 2017 **63** (3), 562-570.
- [5] H. Hu, L. Zhang, X. Chen and Z. Wu: Recursive Carrier Interferometry Aided High Data Rate OFDM Systems With PAPR Suppression, Phase Noise Rejection, and Carrier Frequency Offsets Compensation, *IEEE Transactions on Vehicular Technology*, 2019 **68** (4), 3655-3671.
- [6] W. Jiang, X. Kuai, X. Yuan, W. Liu and Z. Song: Sparsity-Learning-Based Iterative Compensation for Filtered-OFDM With Clipping, *IEEE Communications Letters*, 2020 **24** (11), 2483-2487.
- [7] X. Cheng, D. Liu, W. Shi, Y. Zhao and D. Kong: A Novel Conversion Vector-based Low-Complexity SLM Scheme for PAPR Reduction in FBMC/OQAM Systems., *IEEE Transaction on Broadcasting*, 2020 **66** (3), 656-666.
- [8] Q. Ran, D. S. Yeung, C. C. Eric Tsang and et.al: General Multifractional Fourier Transform Method Based on the Generalized Permutation Matrix Group, *IEEE Trans. Signal Process.*, 2005, **53** (1), 83–98.
- [9] H. M. Ozaktas, O. Arikan, A. Kutay and G. Bozdagi: Digital computation of the fractional Fourier transform, *IEEE Trans. Signal Process.*, 1996, **44** (9), 2141–2150.
- [10] G. Cariolaro, T. Erseghe, P. Kraniauskas and N. Laurenti: A unified framework for the fractional Fourier transforms, *IEEE Trans. Signal Process.*, 1998, **46** (12), 3206-3219.
- [11] G. Cariolaro, T. Erseghe, P. Kraniauskas and N. Laurenti: Multiplicity of Fractional Fourier Transform and Their Relationship, *IEEE Trans. Signal Process.*, 2000, **48** (1), 227-241.
- [12] T. Erseghe, P. Kraniauskas and G. Cariolaro: Unified fractional Fourier transform and sampling theorem, *IEEE Trans. Signal Process.*, 1999, **47** (12), 3419-3423.
- [13] S. C. Pei, M. H. Yeh and C. C. Tseng: Discrete fractional Fourier transform based on orthogonal projections, *IEEE Trans. Signal Process.*, 1999, **47** (5), 1335-1348.
- [14] Y. Li, Z. Song and X. Sha: The Multi-Weighted type Fractional Fourier Transform Scheme and Its Application over Wireless Communications, *EURASIP Journal on wireless communications and networking*, 2018 **41**: 1–10.
- [15] Y. Li, X. Sha and K. Wang: Low Complexity Equalization of HCM Systems with DPFFT Demodulation over Doubly-Selective Channels, *IEEE Signal Process. Lett.*, 2014 **21** (7), 862-865.

A Novel Dual Mode Decision Directed Multimodulus Algorithm (DM-DD-MMA) for Blind Adaptive Equalization

Sagi TADMOR, Sapir CARMI and Monika PINCHAS ¹

*Department of Electrical and Electronic Engineering, Ariel University, Ariel 40700,
Israel*

Abstract. In this paper, we propose for the 16 quadrature amplitude modulation (QAM) input case, a dual-mode (DM), decision directed (DD) multimodulus algorithm (MMA) algorithm for blind adaptive equalization which we name as DM-DD-MMA. In this new proposed algorithm, the MMA method is switched to the DD algorithm, based on a previously obtained expression for the step-size parameter valid at the convergence state of the blind adaptive equalizer, that depends on the channel power, input signal statistics and on the properties of the chosen algorithm. Simulation results show that improved equalization performance is obtained for the 16 QAM input case compared with the DM-CMA (where CMA is the constant modulus algorithm), DM-MCMA (where MCMA is the modified CMA) and MCMA-MDDMA (where MDDMA is the modified decision directed modulus algorithm).

Keywords. Blind equalization; CMA; Adaptive equalization; Dual mode algorithm; Decision directed algorithm

1. Introduction

Data transmission over band-limited communication channels requires the application of equalizers to remove intersymbol interference (ISI) caused by the channel properties [1]. We deal in this paper with the blind adaptive equalization issue where the CMA [2] and the MMA [3], [4] are involved in updating the equalizer's coefficients. According to [1] and [5]-[11], the CMA algorithm is one of the most widely used blind equalization algorithm and according to [12], it has become the workhorse for blind equalization. The CMA [2] is a computationally simple algorithm with outstanding equalization performance for source signals such as M-Ary Phase Shift Keying (MPSK). But, for input signals belonging to the 16QAM constellation, it reaches at the steady state a non negligible residual ISI which may not be sufficient for the system to obtain adequate performance. Thus, the idea proposed by [9] was to switch this computationally simple algorithm to a

¹Corresponding Author: Department of Electrical and Electronic Engineering, Ariel University, Ariel 40700, Israel; E-mail: monikap@ariel.ac.il.

DD algorithm based on the output decision error calculated from the equalizer's output. In other words, the CMA algorithm [2] was used for the initial acquisition phase for removing the heavy ISI and then switched to a tracking algorithm, namely to the DD algorithm in order to obtain a faster convergence speed with a lower residual ISI compared to the algorithm that does not involve the DD algorithm. Indeed, according to [9], the equalization performance was seen improved with the DM-CMA algorithm compared with the CMA method [2] for the 16QAM source constellation input. According to [5], the fact that the CMA is carrier phase independent, is actually considered as an advantage of the CMA since no carrier synchronization is required before blind equalization. According to [7], the constellation seen at the equalized output sequence suffers from an arbitrary phase rotation. Thus, a phase rotator is required at the convergence state of the equalizer in order to rotate the constellation back in the right position. In order to eliminate the need for a rotator to perform separate constellation-phase recovery in steady-state operation, [3], [4] proposed the MMA algorithm that indeed eliminates the need for a phase rotator. Moreover, according to [1], the MMA compensates also minor frequency errors. The advantage of the MMA is due to a separate error-calculation, that is, for real and imaginary part of the received signal, individually [1].

When the equalizer has converged to a relative low residual ISI level, the convolutional noise seen at the equalizer's output is also relative low and considered as Gaussian [13]. But, this is not the case at the beginning of the equalization process. At the beginning of the equalization process, the ISI is usually high and the convolutional noise sequence is more uniform than Gaussian [13], [14]. Please note that the convolutional noise exists at the equalized output in addition to the input signal. The convolutional noise exists since the values for the equalizer's coefficients are not the optimal ones leading to zero residual ISI. The author in [15], considered the case of input signals where the real and imaginary parts of the input signal are independent and where the error involved in the update mechanism of the equalizer is a polynomial function of order three. Please note that the 16QAM source input belongs to this case and that the error involved in the MMA algorithm can be modeled as a polynomial function of order three. According to [15], if the equalizer converges to a very low residual ISI, the convolutional noise probability density function (pdf) can be considered approximately as Gaussian if the step-size parameter complies on some constraints depending on the input constellation statistics, channel power, chosen equalization method and equalizer's tap-length. Based on this outcome, [16] obtained a novel algorithm for a blind adaptive equalizer that supplies the time it takes the equalizer entering the steady state operation without the knowledge of the initial ISI, depending on the input signal statistics and properties of the chosen equalizer. The main idea used in [16] was looking at the squared error obtained by the difference between the step-size parameter obtained from [15] that does not contain any convolutional noise, to the step-size parameter obtained from [15] that involves the convolutional noise by using in the step-size parameter calculation, the equalized output signal statistics instead of the input signal statistics. Please note that for the noiseless case, the equalized output signal statistics are equal to the input signal statistics when the convolutional noise (namely the ISI) tends to zero.

In this paper, we propose a dual-mode, decision directed multimodulus algorithm (DM-DD-MMA) for blind adaptive equalization, where the MMA algorithm is switched to the DD algorithm. The switching mechanism is based on [16] and [15] where we use the normalized squared error obtained by the difference between the step-size parameter ob-

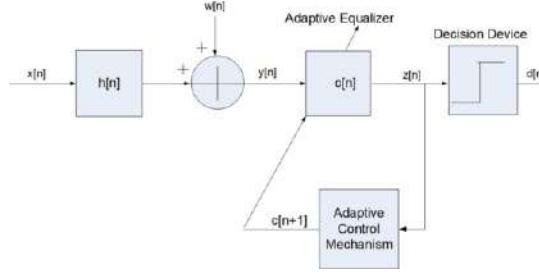


Figure 1. The considered system.

tained from [15] that does not contain any convolutional noise to the step-size parameter obtained from [15] that involves the convolutional noise by using in the step-size parameter calculation, the equalized output signal statistics instead of the input signal statistics. Simulation results will show that for the high SNR case, the new algorithm (DM-DD-MMA) achieves better equalization performance from the convergence speed point of view compared with the DM-CMA, DM-MCMA and MCMA-MDDMA methods while leaving the system with the lowest residual ISI that is approximately obtained also by the DM-MCMA algorithm which has almost the slowest convergence speed compared to all the other methods.

2. System Description

The system under consideration is a linear system, given in Figure 1, where we want to recover its input using an adjustable linear filter (equalizer). In addition, we apply the following assumptions:

- A 16QAM source is used for $x[n]$ where the real and imaginary parts of $x[n]$ are denoted as $x_1[n]$ and $x_2[n]$ respectively.
- The channel $h[n]$ is modeled as a finite impulse response (FIR) filter with channel tap length of \tilde{R} .
- The equalizer $c[n]$ is a FIR filter with equalizer's tap length of N .
- $w[n]$ is modeled as a Gaussian white noise process.
- The real part of the equalized output sequence is denoted as $z_1[n]$.
- The output sequence of the decision device is denoted as $d[n]$.

The equalizer's output is given by:

$$z[n] = y[n] * c[n] = x[n] * s[n] + w[n] * c[n], \quad (1)$$

where “*” stands for the convolutional operation, $y[n] = x[n] * h[n] + w[n]$ and,

$$s[n] = c[n] * h[n] = \delta[n] + \xi[n] \quad (2)$$

where $\delta[n]$ is the Kronecker delta function and $\xi[n]$ stands for the error not having perfect equalization. By perfect equalization we mean that $s[n] = \delta[n]$ which leads for the noiseless case to $z[n] = x[n]$. Now, based on (1) and (2) we may write:

$$z[n] = y[n] * c[n] = x[n] * (\delta[n] + \xi[n]) + w[n] * c[n] = x[n] + p[n] + w[n] * c[n], \quad (3)$$

where $p[n]$ is named as the convolutional noise and is given by $p[n] = x[n] * \xi[n]$. The equalizer's performance is measured via the ISI, given by:

$$ISI = \frac{\sum_t |s_t[n]|^2 - |s|_{max}^2}{|s|_{max}^2}, \quad (4)$$

where $t = N + \tilde{R} - 1$ and $|s|_{max}$ is the maximal absolute value of $s[n]$, given in (2). Based on [9], the equalizer's coefficients are updated for the DM-CMA algorithm by:

$$\begin{aligned} e[n] &= sign\{z[n]\} [|z[n]| - R]; \quad e_D[n] = z[n] - d[n]; \quad R^q = \frac{E[|x[n]|^{2q}]}{E[|x[n]|^q]} \\ \text{if } sign\{Re[e[n]]\} &\neq sign\{Re[e_D[n]]\} \quad \text{or} \quad sign\{Im[e[n]]\} \neq sign\{Im[e_D[n]]\} \quad \text{then} \\ \underline{c}[n+1] &= \underline{c}[n] - \mu_{DM-CMA_1} [z[n] (|z[n]|^2 - R^2)] \underline{y}^*[n] \quad (5) \\ \text{if } sign\{Re[e[n]]\} &= sign\{Re[e_D[n]]\} \quad \text{or} \quad sign\{Im[e[n]]\} = sign\{Im[e_D[n]]\} \quad \text{then} \\ \underline{c}[n+1] &= \underline{c}[n] - \mu_{DM-CMA_2} [z[n] (|z[n]|^2 - |d[n]|^2)] \underline{y}^*[n] \end{aligned}$$

where $|(\cdot)|$, $sign\{(\cdot)\}$ and $(\cdot)^*$ are the absolute operator, signum function and conjugate operator respectively on (\cdot) . μ_{DM-CMA_1} and μ_{DM-CMA_2} are step size parameters, $d[n]$ is the output sequence of the decision device, $Re[\cdot]$ and $Im[\cdot]$ are the real and imaginary parts of $[\cdot]$ respectively and the vector $\underline{c}[n]$ holds the equalizer's coefficients. The input vector $\underline{y}[n]$ is of length N which is the equalizer's tap length. Please note that R^q is a constant modulus depending on the input signal statistics where for $q = 1$ and $q = 2$ we have R and R^2 respectively. Based on (5), when $sign\{Re[e[n]]\} \neq sign\{Re[e_D[n]]\}$ and $sign\{Im[e[n]]\} \neq sign\{Im[e_D[n]]\}$, the CMA algorithm [2] is used to remove the ISI which means that for that case, the probability of correct decisions from the decision device is not very high. Otherwise, the equalizer's coefficients are updated based on the output of the decision device. It should be pointed out that it is assumed in [9], that when $sign\{Re[e[n]]\} = sign\{Re[e_D[n]]\}$ or $sign\{Im[e[n]]\} = sign\{Im[e_D[n]]\}$, the update of the equalizer's coefficients has the right direction and contributes to convergence. According to [1], the equalizer's coefficients are updated for the MCMA-MDDMA algorithm by:

$$\begin{aligned} e_R[n] &= Re[z[n]] \left[|Re[z[n]]|^2 - \frac{E[|x_1[n]|^4]}{E[|x_1[n]|^2]} \right] + Re[z[n]] [|Re[z[n]]|^2 - |sign\{Re[z[n]]\}|^2] \\ e_I[n] &= Im[z[n]] \left[|Im[z[n]]|^2 - \frac{E[|x_2[n]|^4]}{E[|x_2[n]|^2]} \right] + Im[z[n]] [|Im[z[n]]|^2 - |sign\{Im[z[n]]\}|^2] \quad (6) \\ Re[\underline{c}[n+1]] &= Re[\underline{c}[n]] - \mu_{MCMA-MDDMA} [e_R[n] Re[\underline{y}[n]] + e_I[n] Im[\underline{y}[n]]] \\ Im[\underline{c}[n+1]] &= Im[\underline{c}[n]] + \mu_{MCMA-MDDMA} [e_R[n] Im[\underline{y}[n]] - e_I[n] Re[\underline{y}[n]]] \end{aligned}$$

where $\mu_{MCMA-MDDMA}$ is the step size parameter. Please note that according to [1], the MCMA-MDDMA algorithm belongs to a group of algorithms named as joint algorithms. This group additively combines two single algorithms which in our case are the MCMA and MDDMA algorithms. Based on [1], we denote DM-MCMA as the algorithm that updates the equalizer's coefficients according to:

$$\begin{aligned} \text{if } sign\{e_R[n]\} &\neq sign\{Re[e_D[n]]\} \quad \text{or} \quad sign\{e_I[n]\} \neq sign\{Im[e_D[n]]\} \quad \text{then} \\ Re[\underline{c}[n+1]] &= Re[\underline{c}[n]] - \mu_{DM-MCMA_1} [e_R[n] Re[\underline{y}[n]] + e_I[n] Im[\underline{y}[n]]] \\ Im[\underline{c}[n+1]] &= Im[\underline{c}[n]] + \mu_{DM-MCMA_1} [e_R[n] Im[\underline{y}[n]] - e_I[n] Re[\underline{y}[n]]] \quad (7) \\ \text{if } sign\{e_R[n]\} &= sign\{Re[e_D[n]]\} \quad \text{or} \quad sign\{e_I[n]\} = sign\{Im[e_D[n]]\} \quad \text{then} \\ \underline{c}[n+1] &= \underline{c}[n] - \mu_{DM-MCMA_2} [z[n] (|z[n]|^2 - |d[n]|^2)] \underline{y}^*[n] \end{aligned}$$

where $\mu_{DM-MCMA_1}$ and $\mu_{DM-MCMA_2}$ are the step-size parameters. The expression for $e_D[n]$ is given in (5) and those for $e_R[n]$ and $e_I[n]$ are given in (6). Please note that based on (7), when $sign\{e_R[n]\} \neq sign\{Re[e_D[n]]\}$ and $sign\{e_I[n]\} \neq sign\{Im[e_D[n]]\}$, the MCMA-MDDMA algorithm is used to remove the ISI which means that for that case, the probability of correct decisions from the decision device is not very high. Otherwise, the equalizer's coefficients are updated based on the output of the decision device. According to [1], [3], [4], [6], the equalizer's coefficients for the MMA algorithm are updated by:

$$\underline{c}[n+1] = \underline{c}[n] - \mu_{MMA} \left[Re[z[n]] \left[|Re[z[n]]|^2 - \frac{E[x_1[n]^4]}{E[x_1[n]^2]} \right] + j [Im[z[n]]] \left[|Im[z[n]]|^2 - \frac{E[x_2[n]^4]}{E[x_2[n]^2]} \right] \right] y^*[n] \quad (8)$$

where μ_{MMA} is the step size parameter. As was already mentioned earlier in this paper, the MMA algorithm eliminates the need for a phase rotator due to the fact that it uses a separate error-calculation, that is, for real and imaginary part of the received signal, individually, which is not the case in the CMA algorithm [2].

3. The New Proposed Algorithm

In this section we propose a dual-mode, decision directed multimodulus algorithm (DM-DD-MMA) for blind adaptive equalization, where the MMA algorithm [3], [4] is switched to the DD algorithm [9], based on a previously obtained expression for the step-size parameter [15] valid at the convergence state of a blind adaptive equalizer where the error involved in the update mechanism of the equalizer is a polynomial function of order three. According to [15], the Gaussian model holds for the convolutional noise pdf at the steady state operation if the step-size parameter complies to:

$$\mu \ll \frac{2|a_1+3a_3n_2|}{3\left(\sigma_x^2 N \sum_{k=0}^{R-1} |h_k[n]|^2\right) |a_1^2 + 12a_3a_1n_2 + 15a_3^2n_4|} \quad (9)$$

where $n_a = E[x_1^a[n]]$ ($a = 2, 4$ and $E[(\cdot)]$ is the expectation operator) and the values for a_1 and a_3 are given by $a_1 = -\frac{E[x_1[n]^4]}{E[x_1[n]^2]}$ and $a_3 = 1$ respectively for the MMA algorithm [3], [4]. Now, for the noiseless case we have that the equalized output ($z[n]$) tends to the input signal ($x[n]$) when $p[n] \rightarrow 0$. Thus, if we consider the following expression:

$$\underline{nor}_{err} = \frac{\underline{err}}{\max(\underline{err})}; \quad \underline{err}[n] = (\mu_p[n] - \mu_x)^2$$

$$\mu_p[n] = \frac{2|a_1+3a_3\tilde{n}_2|}{3\left(\sigma_x^2 N \sum_{k=0}^{R-1} |h_k[n]|^2\right) |a_1^2 + 12a_3a_1\tilde{n}_2 + 15a_3^2\tilde{n}_4|}; \quad \mu_x = \frac{2|a_1+3a_3n_2|}{3\left(\sigma_x^2 N \sum_{k=0}^{R-1} |h_k[n]|^2\right) |a_1^2 + 12a_3a_1n_2 + 15a_3^2n_4|} \quad (10)$$

where $\tilde{n}_a = \frac{1}{L} \sum_{b=0}^{b=L-1} z_1^a[n-b]$, \underline{err} is a vector (an error vector) containing the elements defined by $\underline{err}[n]$, $\max(\underline{err})$ contains the maximal value of \underline{err} and \underline{nor}_{err} is the normalized error vector. Please note that for the noiseless case, when $p[n] \rightarrow 0$ we have that $\mu_p[n] \rightarrow \mu_x$ which leads to $\underline{err}[n] \rightarrow 0$. Thus, based on (10) we have some indication on how far away we are from the convergence state of the equalizer. Now, based on (10), the equalizer's taps of our new algorithm, denoted as the DM-DD-MMA method are updated according to:

$$\begin{aligned}
& \text{if } nor_{err}[n] < 0.01 \text{ then} \\
& \quad \underline{c}[n+1] = \underline{c}[n] - \mu_2 [z[n] (|z[n]|^2 - |d[n]|^2)] \underline{y}^*[n] \\
& \text{if } nor_{err}[n] \geq 0.01 \text{ then} \\
& \quad \underline{c}[n+1] = \underline{c}[n] - \mu_1 \left[Re[z[n]] \left[|Re[z[n]]|^2 - \frac{E[|x_1[n]|^4]}{E[|x_1[n]|^2]} \right] + j [Im[z[n]]] \left[|Im[z[n]]|^2 - \frac{E[|x_2[n]|^4]}{E[|x_2[n]|^2]} \right] \right] \underline{y}^*[n]
\end{aligned} \tag{11}$$

and μ_1 and μ_2 are the step size parameters. Please note that according to (11), the MMA algorithm [3], [4] is used to remove the ISI when $nor_{err}[n] \geq 0.01$ which means that for that case, the probability of correct decisions from the decision device is not very high. Otherwise, the equalizer's coefficients are updated based on the output of the decision device.

4. Simulation Results

In this section we compare our new proposed equalization method (11) for the 16QAM input case with three different equalization methods (with the MCMA-MDDMA (6), DM-MCMA (7) and DM-CMA (5) methods) for three different channel cases and different values for the SNR (30 dB down to 10 dB). In this work we considered three channels:

Channel1 [17]: $h_n = \{0 \text{ for } n < 0; -0.4 \text{ for } n = 0; 0.8(0.4^{n-1}) \text{ for } n > 0\}$.
Channel2 [18]: $h_n = (0.4851, -0.72765, -0.4851)$.

Channel3 based on [19] where we down decimated the channel parameters by 32 and normalized them so that $h^T h = 1$: $h_n = (0.6069, -0.2023, -0.6069, -0.2529, -0.1517, 0.0506, 0.1011, 0.1517, 0.2023, 0.1517, 0.1517, 0.1011, 0.0506)$

In this work, we use the tap-centering initialization strategy where the equalizers' taps were initialized by setting the center tap equal to one and all others to zero. For Channel1, Channel2 and Channel3 we used an equalizer with 13, 15 and 57 taps, respectively. The step-size parameters μ_1 , μ_2 , μ_{DM-CMA_1} , μ_{DM-CMA_2} , $\mu_{DM-MCMA_1}$, $\mu_{DM-MCMA_2}$, $\mu_{MCMA-MDDMA}$, were chosen for fast convergence with low steady state residual ISI (please refer to Figures 2-9). Figures 2-4 describe the ISI as a function of the iteration number obtained by our new proposed method (DM-DD-MMA) compared to the DM-CMA, DM-MCMA and MCMA-MDDMA methods for the 16QAM constellation input sent via Channel1 and with SNR values of 30dB, 20dB and 10dB. Please note that the iteration number describes the number of updating the equalizer's coefficients during the equalization process. According to Figures 2-3, ($SNR = 30dB$ and $SNR = 20dB$ cases), our new proposed method (DM-DD-MMA) achieves the fastest convergence speed while leaving the system with the lowest residual ISI that is approximately obtained also by the DM-MCMA method which has almost the slowest convergence compared to all the other methods. According to Figure 4, (the $SNR = 10dB$ case), our new proposed method (DM-DD-MMA) has similar equalization performance with the MCMA-MDDMA method and both algorithms lead to a faster convergence speed and to a lower residual ISI compared to the DM-CMA and DM-MCMA methods. Figures 5-7 describe the ISI as a function of the iteration number obtained by our new derived method (DM-DD-MMA) compared to the DM-CMA, DM-MCMA and MCMA-MDDMA meth-

ods for the 16QAM constellation input sent via Channel2 and with SNR values of 30dB, 20dB and 10dB. As was the case with Channel1, for SNR values down to 20dB, our new proposed method (DM-DD-MMA) achieves the fastest convergence speed while leaving the system with the lowest residual ISI that is approximately obtained also by the DM-MCMA method which has almost the slowest convergence speed compared to all the other methods (please refer to Figures 5-6). For the $SNR = 10dB$ case (Figure 7), our new derived method (DM-DD-MMA) has similar equalization performance with the MCMA-MDDMA method and both algorithms lead to a faster convergence speed and to a lower residual ISI compared to the DM-CMA and DM-MCMA methods. Figures 8-9 describe the ISI as a function of the iteration number obtained by our new derived method (DM-DD-MMA) compared to the DM-CMA, DM-MCMA and MCMA-MDDMA methods for the 16QAM constellation input sent via Channel3 and with SNR values of 30dB and 20dB. As was already seen with Channel1 and Channel2 for the $SNR = 30dB$ case, our new derived method (DM-DD-MMA) achieves also for Channel3 the fastest equalization convergence speed while leaving the system with the lowest residual ISI that is approximately achieved also by the DM-MCMA algorithm which has almost the slowest convergence speed compared to all the other methods (please refer to Figure 8). For the $SNR = 20dB$ case (Figure 9), our new derived method (DM-DD-MMA) has similar equalization performance with the MCMA-MDDMA algorithm and both algorithms lead to a faster convergence speed and to a lower residual ISI compared to the DM-CMA and DM-MCMA methods. Figures 2-9 clearly show the effectiveness of the switching mechanism involved in our new proposed algorithm (DM-DD-MMA) that is responsible to switch to the update mechanism of the equalizer's taps based on the decisions made by the decision device, when relative reliable decisions can be done by the decision device. As already was noted earlier in this paper, $\mu_p[n]$ tends to μ_x (for the noiseless case) when the convolutional noise (namely the residual ISI) tends to zero. Thus, for the noiseless case, where $\mu_p[n]$ is very close to μ_x , reliable decisions can be obtained by the decision device. In this paper, we switched to the decision directed algorithm when $nor_{err}[n] < 0.01$ (where $\mu_p[n]$ is not equal to μ_x but is also not too far away from it) in order to get improved equalization performance improvement from the residual ISI as well as from the convergence rate point of view. Since μ_x was obtained for the noiseless case, it is not a surprise that our proposed algorithm does not show the same equalization performance improvement for the very low SNR case as it does for the high SNR situation. It is possible that better equalization performance can be obtained for the very low SNR case if we set $nor_{err}[n] < t_r$ where $t_r \neq 0.01$ or if a new expression for μ_x is obtained that is applicable also for the noisy case.

5. Conclusions

We proposed in this work a dual-mode, decision directed multimodulus algorithm (DM-DD-MMA) for blind adaptive equalization, where the MMA algorithm is switched to the DD algorithm, based on a previously obtained expression for the step-size parameter valid at the convergence state of the equalizer for which the convolutional noise pdf can be considered approximately as Gaussian. The new proposed algorithm (DM-DD-MMA) was compared with the DM-CMA, DM-MCMA and MCMA-MDDMA methods with three different channel cases and with SNR values down to 10dB. Simulation results

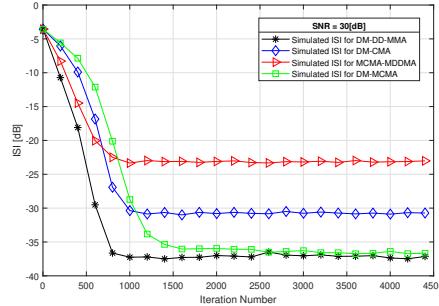


Figure 2. Simulation results (averaged results with 100 trials) with Channel1 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 30dB$, $\mu_1 = 0.0002$, $\mu_2 = 0.0001$, $\mu_{DM-CMA_1} = 0.000025$, $\mu_{DM-CMA_2} = 0.0002$, $\mu_{DM-MCMA_1} = 0.000005$, $\mu_{DM-MCMA_2} = 0.0001$, $\mu_{MCMA-MDDMA} = 0.00005$

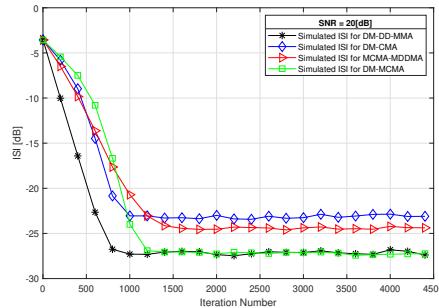


Figure 3. Simulation results (averaged results with 100 trials) with Channel1 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 20dB$, $\mu_1 = 0.0002$, $\mu_2 = 0.0001$, $\mu_{DM-CMA_1} = 0.000025$, $\mu_{DM-CMA_2} = 0.0002$, $\mu_{DM-MCMA_1} = 0.000002$, $\mu_{DM-MCMA_2} = 0.0001$, $\mu_{MCMA-MDDMA} = 0.00003$

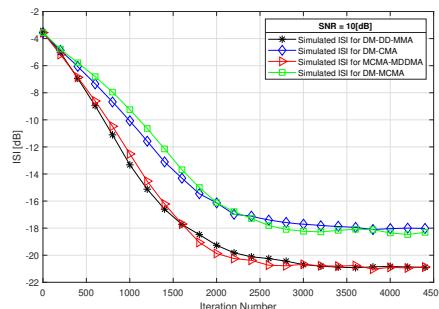


Figure 4. Simulation results (averaged results with 100 trials) with Channel1 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 10dB$, $\mu_1 = 0.00004$, $\mu_2 = 0.00002$, $\mu_{DM-CMA_1} = 0.00005$, $\mu_{DM-CMA_2} = 0.00005$, $\mu_{DM-MCMA_1} = 0.00002$, $\mu_{DM-MCMA_2} = 0.00005$, $\mu_{MCMA-MDDMA} = 0.00002$

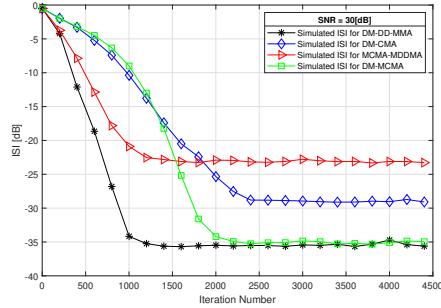


Figure 5. Simulation results (averaged results with 100 trials) with Channel2 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 30dB$, $\mu_1 = 0.0002$, $\mu_2 = 0.0001$, $\mu_{DM-CMA_1} = 0.000027$, $\mu_{DM-CMA_2} = 0.0002$, $\mu_{DM-MCMA_1} = 0.000005$, $\mu_{DM-MCMA_2} = 0.0001$, $\mu_{MCMA-MDDMA} = 0.00004$

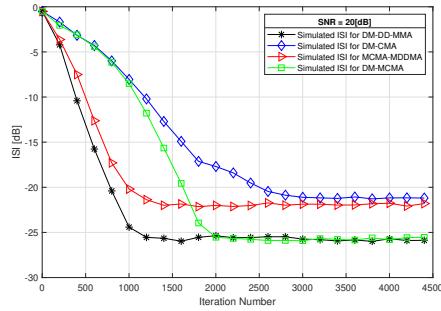


Figure 6. Simulation results (averaged results with 100 trials) with Channel2 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 20dB$, $\mu_1 = 0.0002$, $\mu_2 = 0.0001$, $\mu_{DM-CMA_1} = 0.000027$, $\mu_{DM-CMA_2} = 0.0002$, $\mu_{DM-MCMA_1} = 0.0000027$, $\mu_{DM-MCMA_2} = 0.0001$, $\mu_{MCMA-MDDMA} = 0.00004$

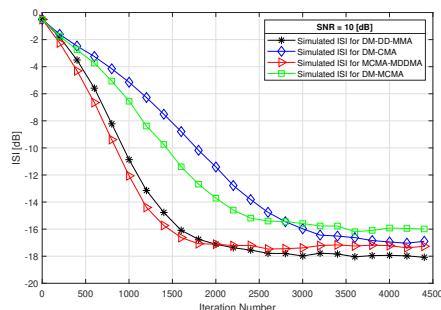


Figure 7. Simulation results (averaged results with 100 trials) with Channel2 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 10dB$, $\mu_1 = 0.00005$, $\mu_2 = 0.00001$, $\mu_{DM-CMA_1} = 0.00005$, $\mu_{DM-CMA_2} = 0.00004$, $\mu_{DM-MCMA_1} = 0.00005$, $\mu_{DM-MCMA_2} = 0.00005$, $\mu_{MCMA-MDDMA} = 0.00003$

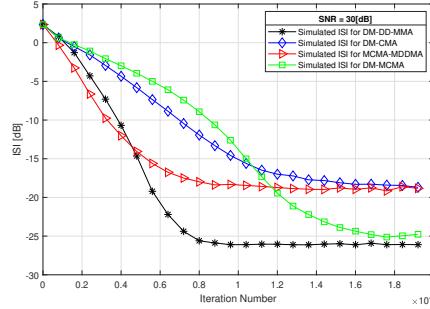


Figure 8. Simulation results (averaged results with 100 trials) with Channel3 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 30dB$, $\mu_1 = 0.00005$, $\mu_2 = 0.00005$, $\mu_{DM-CMA_1} = 0.00003$, $\mu_{DM-CMA_2} = 0.00003$, $\mu_{DM-MCMA_1} = 0.000007$, $\mu_{DM-MCMA_2} = 0.00003$, $\mu_{MCMA-MDDMA} = 0.00002$

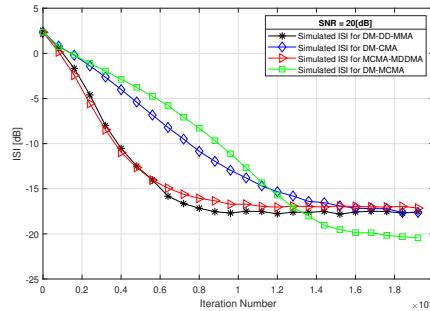


Figure 9. Simulation results (averaged results with 100 trials) with Channel3 for the DM-DD-MMA, DM-CMA, DM-MCMA and MCMA-MDDMA methods. $SNR = 20dB$, $\mu_1 = 0.00005$, $\mu_2 = 0.00005$, $\mu_{DM-CMA_1} = 0.00003$, $\mu_{DM-CMA_2} = 0.00003$, $\mu_{DM-MCMA_1} = 0.000007$, $\mu_{DM-MCMA_2} = 0.00003$, $\mu_{MCMA-MDDMA} = 0.00002$

has confirmed that for the three channel cases with $SNR = 30dB$ as well as for the first two channel cases (Channel1 and Channel2) with $SNR = 20dB$, a faster convergence speed is obtained with our new derived algorithm (DM-DD-MMA) compared with the DM-CMA, DM-MCMA and MCMA-MDDMA methods while leaving the system with the lowest residual ISI that is approximately obtained also by the DM-MCMA algorithm which has almost the slowest convergence speed compared to all the other methods. For the $SNR = 10dB$ case, simulation results have shown that for the first two channel cases (Channel1 and Channel2), our new proposed method (DM-DD-MMA) has similar equalization performance with the MCMA-MDDMA algorithm and both algorithms lead to a faster convergence speed and to a lower residual ISI compared to the DM-CMA and DM-MCMA algorithms. As already mentioned earlier, the new proposed algorithm (DM-DD-MMA) is based on a previously derived expression for the step-size parameter valid at the convergence state of the equalizer. This expression for the step-size was obtained for the noiseless case. Thus, this may be the reason that our new proposed method (DM-DD-MMA) achieved the best equalization performance for SNR values above 20dB compared to the DM-CMA, DM-MCMA and MCMA-MDDMA methods. Therefore, our fu-

ture direction will be deriving a new expression for the step-size parameter valid at the convergence state of the equalizer that is suitable also for the very noisy case.

References

- [1] Meyer, S. Real-time bit error rate analysis of blind equalization algorithms. *SIGNAL PROCESSING ALGORITHMS, ARCHITECTURES, ARRANGEMENTS, AND APPLICATIONS, SPA, IEEE, THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC.* **2013**, 268–273.
- [2] Godard, D.N. Self recovering equalization and carrier tracking in two-dimenional data communication system. *IEEE Trans. Comm.* **1980**, 28(11), 1867–1875.
- [3] Oh, K.N; Chin, Y. O. Modified constant modulus algorithm: Blind equalization and carrier phase recovery algorithm. in *Proc. IEEE Int. Conf. Commun.* **1995**, 1, 498—502.
- [4] Yang, J.; Werner, J.-J.; Dumont, G. A. The multimodulus blind equalization and its generalized algorithms. *IEEE J. Sel. Areas Commun.* **2002**, 20(6), 997—1015.
- [5] Yuan, J.; Lin, T; Equalization and Carrier Phase Recovery of CMA and MMA in Blind Adaptive Receivers. *IEEE Transactions on Signal Processing*, **2010**, 58(6), 3206-3217, June 2010, doi: 10.1109/TSP.2010.2044255. http://www.ee.fju.edu.tw/pages..../032_faculty/yuan/papers/Journal1/2010_2.pdf.
- [6] Yuan, J.T.; Tsai, K.D; Analysis of the Multimodulus Blind Equalization Algorithm in QAM Communication Systems. *IEEE TRANSACTIONS ON COMMUNICATIONS* **2005**, 53(9), 1427–1431.
- [7] Hamzehyan, R.; Dianat, R.; Shirazi, N.C. New Variable Step-Size Blind Equalization Based on Modified Constant Modulus Algorithm. *International Journal of Machine Learning and Computing* **2012**, 2(1), 30–34. <http://ijmlc.org/papers/85-A30092.pdf>, (accessed on 01 June 2021).
- [8] Samarasinghe, P.D.; Kennedy, R.A. Minimum Kurtosis CMA Deconvolution for Blind Image Restoration. *4th International Conference on Information and Automation for Sustainability, 2008. ICIAFS 2008.* **2008**, 271 - 276.
- [9] Gao, H.; Rao, W.; He, A.; Tu, Y.; Li, Z.; New Dual-mode CMA Equalizer. *Physics Procedia* **2012**, 25 1973 – 1979.
- [10] Rao, W.; New Concurrent Modulus Algorithm and Soft Decision Directed Scheme for Blind Equalization. *2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011), Procedia Environmental Sciences* **2011**, 10, 1264 — 1269.
- [11] Sun, J.; Li, X.; Chen, K.; Cui, W.; Chu, M.; A Novel CMA+DD LMS Blind Equalization Algorithm for Underwater Acoustic Communication, in *The Computer Journal*, **2020**, 63(1), 974–981, doi: 10.1093/comjnl/bxaa013.
- [12] Alsaid, S.; Abdel-Raheem, E.; Mayyas, K; New Constant Modulus Blind Equalizer with Optimum Tap-length for QAM Signals. *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) 2019*, 1-4, doi: 10.1109/ISSPIT47144.2019.9001754.
- [13] Haykin, S. Adaptive Filter Theory. In *Blind deconvolution*; Haykin, S., Ed.; Prentice-Hall: Englewood Cliffs, NJ, USA, 1991; Chapter 20.
- [14] Godfrey, R.; Rocca, F. Zero memory non-linear deconvolution. *Geophys. Prospect.* **1981**, 29, 189–228.
- [15] Pinchas, M.; Convolutional noise PDF at the convergence state of a blind adaptive equalizer. *22nd International Conference on Circuits, Systems, Communications and Computers (CSCC 2018), MATEC Web of Conferences*, **210**, 2018, https://www.researchgate.net/publication/328097740_Convolutional_Noise_PDF_at_the_Convergence_State_of_a_Blind_Adaptive_Equalizer, (accessed on 01 June 2021).
- [16] Pinchas, M.; The Convergence Time of a Blind adaptive Equalizer. *INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING*, **2019**, 13, 678–683, <https://www.nauj.org/main/NAUN/circuitssystemssignal1/2019/b842005-alj.pdf>.
- [17] Shalvi, O.; Weinstein, E. New criteria for blind deconvolution of nonminimum phase systems (channels), *IEEE Trans. Information Theory* **1990**, 36, 312–321.
- [18] Pinchas, M. A Closed Approximated Formed Expression for the Achievable Residual Intersymbol Interference Obtained by Blind Equalizers. *Signal Process. (Eurasip)*, **2010**, 90(6), 1940–1962.
- [19] Matlab DMTTEQ Toolbox 3.0 Release. Available online: <http://users.ece.utexas.edu/~bevans/projects/adsl/dmtteq/dmtteq.html>.

Efficiency Measurement of Compressed Air Compressors Using High Availability SoC With 1oo2 Redundancy Architecture

MOHAMED ABDELAWWAD^{a,1}, EIKE HAHN^a and JOSEF BOERCSOEK^a
JULIAN FAIRBROTHER^b, TAREK AL SHAHADAT^b and PETER OTTO^b

^a*ICAS, Institute for Computer Architecture and System Programming, University of Kassel, Kassel, Germany*

^b*Postberg + Co. GmbH, Kassel, Germany*

Abstract. Compressed air systems are essential components in various industrial and everyday applications. The efficiency of these systems is very important due to their role in the energy consumption of industrial plants. To increase efficiency, a new concept for compressed air compressors based on Industry 4.0 is presented. Due to the aggressive environmental conditions in which the compressed air compressors operate, a new design of a SoC with high availability based on 1oo2 redundancy architecture is developed.

Keywords. Smart compressed air compressor, compressor efficiency, isentropic efficiency, 1oo2 redundancy architecture, high availability SoC, volume flow measurement

1. Introduction

One of the most important components in many areas of daily life and industry are the Compressed Air Systems (CASs) [1]. Therefore, the market size was estimated at USD 31.26 billion in 2018 and is expected to reach USD 32.73 billion by 2025 [2]. However, the efficiency of the CASs is a major dilemma as 5 to 10% of electricity consumption in industrial applications is used by them [3, 4]. Therefore, measuring and improving efficiency of CASs plays an important role in reducing energy consumption and result in potential economic savings [5].

The current strategic goal of industrial development is to implement the concept of Industry 4.0 [6]. In this concept, intelligent sensor technology is developed to achieve the so-called Cyber Physical Systems (CPS), where a connection between the physical and virtual space is established and improved. To increase reliability and availability, additional functionalities such as fault prediction and self-adaptation are integrated into the new CPSs [7]. For higher level, semantic communication the CPS supports the OPC UA protocol via an ethernet interface.

¹Corresponding Author: Research Engineer, Institute for Computer Architecture and System Programming, University of Kassel, Wilhelmshöher Allee 71, 34121 Kassel, Germany; E-mail: m.abdelawwad@uni-kassel.de.

Wear occurring in industrial compressed air compressors is a long-ignored subject which can lead to potentially efficiency degradation. For example, the efficiency of compressors according to [8] shows a degradation of up to 45% over 30 years, which is a typical life cycle for compressors.

The current Life Cycle Cost (LCC) calculation is based on fixed compressor efficiency values given by the manufacturers which cannot be guaranteed over the whole lifecycle of the compressor. Therefore, the continuous efficiency assessment can be used to get a proper LCC calculation. Changes in efficiency can also be used to initialize maintenance measures to avoid increases in energy consumption and emissions.

The paper at hand presents a novel solution for continuous efficiency assessment of compressed air compressors. The possibility to monitor efficiency degradation over time of on-site in operating compressed air compressors could give vital insights into energy consumption reduction potentials and emission reduction in the industrial sector.

Another potential issue with CASs is that, in most cases, compressed air compressors are installed and operated in an environment that is aggressive to electronic components. This can lead to availability problems. Therefore, a new design of compressed air compressor efficiency assessment based on a loo2 system architecture is proposed.

2. Efficiency calculation of industrial compressor systems

In the last years, there have been several significant international developments in the field of regulations and energy efficiency standards for air compressors. The European Union released the *Ecodesign preparatory study on compressors* (ENER Lot 31) [9], the US Department of Energy (DOE) issued the *Energy Conservation Standards for Air Compressors* [10] and the Peoples Republic of China has released GB 19153-2019 *Minimum allowable values of energy efficiency and energy efficiency grades for displacement air compressors* [11].

These studies setup frameworks for compressor efficiency, i.e. differentiation between different compressor types and define metrics to access the efficiency and costs of a compressed air compressor. So called minimum energy performance standards (MEPS) are adopted under which manufacturers of compressed air compressors are obliged to reduce emission and the energy consumption throughout the product life cycle.

Manufacturers are required to comply to the new standards and have to test their machines according to the regulations. Compliance with DOE rule is required after January 10, 2025. The Ecodesign regulation is still in process.

For existing in-service compressors an energy audit is an effective instrument to detect optimization potentials. A comprehensive evaluation of the compressed air system should expose over-sized, old and inefficient equipment, equipment without proper control, leakages and inappropriate applications. ISO 11011:2013 *Compressed air – Energy efficiency – Assessment* recommends a holistic approach and considers the whole compressed air system as well as heat recovery. Correct measuring and the right interpretation of the collected data are identified as the basis for efficient usage and predictive maintenance.

To determine the efficiency and proper operation of a machine the most relevant international performance standards and test methods for air compressors are ISO 1217:2009, *Displacement Compressor Acceptance Tests* [12], ISO 5389:2005, *Turbo*

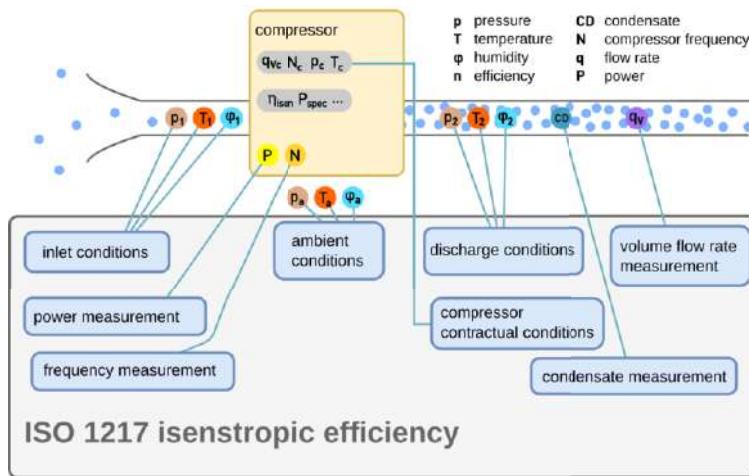


Figure 1. Isentropic efficiency according to ISO 1217 [12].

Compressors – Performance test code [13] and ASME PTC 10-1997, Performance Test Code on Compressors and Exhausters [14].

ISO 1217 specifies operating and testing conditions in arbitrary operating points (full load, part load). Ambient, inlet and discharge conditions, energy consumption, condensate and the volume flow rate have to be measured. Specific test procedures, depending on the type of compressor, should be carried out at certain conditions to assert the compliance to the declared output.

The main metric, adopted by all international regulations, is the *isentropic efficiency* (figure 1), i.e. the ratio of the theoretical isentropic power to the actual power required for a compression process. The isentropic efficiency makes it possible to easily identify compressor losses compared to an ideal standard and cannot be misunderstood as no conversion of units is required.

$$\eta_{isen} = \frac{P_{isen}}{P_{real}} \quad (1)$$

Where:

η_{isen} : isentropic efficiency

P_{isen} : isentropic power

P_{real} : electric input power of package

$$P_{isen} = q_V \cdot \frac{p_1 \cdot \kappa}{\kappa - 1} \cdot \left[\left(\frac{p_2}{p_1} \right)^{\frac{\kappa-1}{\kappa}} - 1 \right] \quad (2)$$

Where:

q_V : volume flow rate

p_1 : inlet pressure

Table 1. Parameters according to Ecodesign study [9]

compressor type	coefficients		
	A	B	C
fixed speed rotary compressor	-0.928	13.911	27.110
variable speed rotary compressor	-1.549	21.573	0.905

p_2 : discharge pressure

The isentropic process is an idealized thermodynamic process and assumes no heat exchange with the environment. For the calculation of isentropic efficiency and the relationship with the former metric of the standard, *specific energy*, see ISO 1217:2009/Amendment 1:2016 [12]. The target efficiency for air compressor packages can be calculated using the following regression curve:

$$\eta_{target} = A \cdot \ln^2(q_{Vmax}) + B \cdot \ln(q_{Vmax}) + C + (100 - (A \cdot \ln^2(q_{Vmax}) + B \cdot \ln^2(q_{Vmax}) + C)) \cdot d/100 \quad (3)$$

Where:

η_{target} : isentropic efficiency that the product shall achieve

q_{Vmax} : full-load volume flow [l/s]

d : proportional loss factor

The parameters for A, B and C according to the Ecodesign study [9] are shown in table 1. The d-value (proportional loss factor) expresses the difference in the average efficiency and the theoretical optimum package efficiency and is specified in the respective regulations.

3. Availability and safety architectures

Air compressors are often used in industrial environments. This environment may contain or be surrounded by aggressive gases, high temperatures, moisture, or oil, which constitute very harsh conditions especially for electronics and sensitive measurements. This increases the failure rate of the electronic components as in e.g. microcontrollers, which leads to a reduction of the reliability and availability of the measurement.

One of the most common and successful techniques for increasing the reliability and/or availability of a system is redundancy. In redundancy architecture, multiple subsystems are available to perform the same functionality. Therefore, the system is referred as M out of N (MoN), where N represents the number of subsystems performing or willing to perform the same function and M represents the actual required output subsystems such as 1oo2, 1oo3, 1oo4, 2oo4 etc. If one subsystem fails, a redundant subsystem takes control and performs the system task [15].

There are many methods to provide redundancy, for example, static redundancy uses a voting element to analyze and compare the output of multiple subsystems to select the majority decision from those subsystems. In such a methodology, the configuration of the system does not change after a fault is detected in a subsystem [16].

In dynamic redundancy, the system configuration is dynamically changed based on the evaluation of the system's diagnostic element. In hot dynamic redundancy, all subsystems are running and connected to the load. Within a very short time, it is possible to switch from one subsystem to another in order to maintain the safe operation of the system. In dynamic cold or standby redundancy, only one subsystem is running and performs the system tasks while the other subsystems are in standby state. When a malfunction is detected in the running subsystem, one of the standby subsystems is switched on to handle the system functions [17].

In SoC architecture design with redundancy, there are two main strategies to provide the arbitration, i.e., redundancy with hardware comparator and redundancy with software comparator. The redundancy with hardware comparator is the common implemented technique in the available safety microcontroller such as HiCore1 [18] and Hercules microcontroller family [19]. Such lockstep approach supports only shutdown in case of system failure, which increases the reliability of the system but negatively affects its availability. To increase the flexibility of the system, redundancy with software comparator is developed [20, 21]. In this paper, a new concept of SoC based on 1oo2 system architecture with software comparator is presented to increase the availability of the system.

4. Design of the system

For the concept of the system design, shown in figure 2, several requirements must be fulfilled. The compressed air system contains one to N compressors from different types. For each compressor in the system, an efficiency calculation has to be performed. Efficiency is calculated on the intelligent and high available flow efficiency sensor downstream of the compressor. Relevant data for the efficiency calculation is collected up to 20 m away from the flow sensor. A high-level PLC provides an overview of all collected and calculated data and allows the provision of the data for possible cloud applications.

From the requirements listed before, first points of the specification for the flow sensor are derived. Several sensors must be connected by different interfaces. The integration of the flow efficiency sensor into an over laying network structure will be done by OPC UA as an ethernet based protocol.

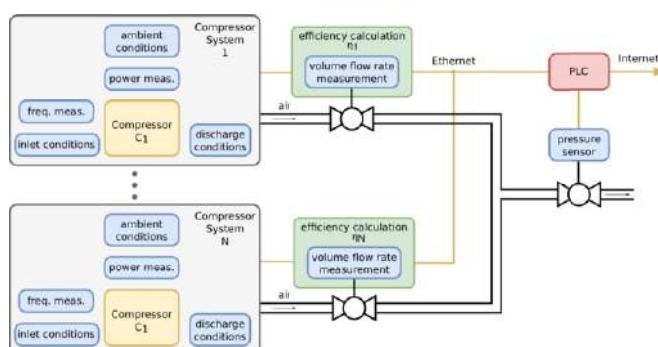


Figure 2. Schematic of the smart air compressed system.

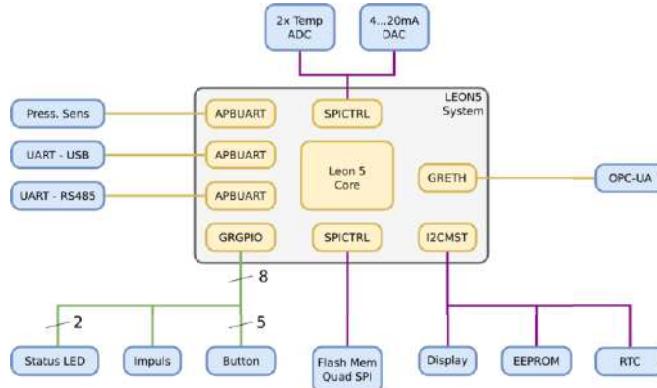


Figure 3. Schematic of the one channel sensor SoC.

4.1. Efficiency SoC design

Based on the theory provided in section 3, the hardware concept for the flow sensor will follow a 1oo2 redundancy approach. For both channels a high level of diagnostic shall be implemented to detect malfunctions in one channel to keep the system running by the second channel. Internally a LEON5 IP core from Cobham Gaisler will be used. It is available under GPL license and as commercial licensed IP for later usage. The benefit seen with this IP is the availability of hundreds of peripherals, also available under GPL license. Since it already brings a small FPU, Buildroot Linux is available for LEON5. According to the specification, the LEON5 core will be extended with different peripherals like UART, SPI, GPIO, ethernet and I2C interfaces. The one channel approach can be seen in figure 3.

For the specification of the sensor SoC, among the needed data for efficiency calculation according to the mentioned calculation methods, several requirements are derived from the international standard OIML 137 [22, 23].

The two-channel system will follow the concept in figure 4. System A and System B will be extended with diagnosis units to get a high degree of diagnostics. An arbiter will be used to realize the dynamic hot standby functionality. For the user, the device will behave like a single channel system, with additional health and predictive maintenance information.

The SoC approach will be realized on a FPGA based hardware platform with the possibility to realize it later as a silicon-based product. For the commercial usage an FPGA based product is not feasible. The costs for the FPGA are not competitive as well as that the susceptibility to errors is too high.

4.2. Efficiency measurement design

Volume flow measurement downstream of a compressed air compressor is a demanding and complex problem. For a reliable measurement the system must be inert against shock, temperature, aggressive medium conditions (oil) and condensate. To meet these conditions the differential pressure (DP) measurement principle as described in ISO 5167 was selected.

Table 2. Efficiency measurement design parameters

Measurand	Unit	Range
Pressure	bar	1...20
Differential pressure	mbar	0...100
Medium temperature	°C	-30...300
Inlet temperature	°C	-20...65
Inlet pressure	mbar	300...1100
Inlet humidity	% RH.	0...100
Compressor power	kWh	0...350

For minimal pressure loss and the ability of retrofitting a special averaging pitot tube (APT) was designed to satisfy the above-mentioned conditions (figure 5). DP volume flow measurement according to ISO 5167-1 [24] requires the additional measurement of medium temperature and system pressure. Since the APT was designed to accommodate a temperature as well as a pressure measurement no further installation points, sensors and transmitters are required.

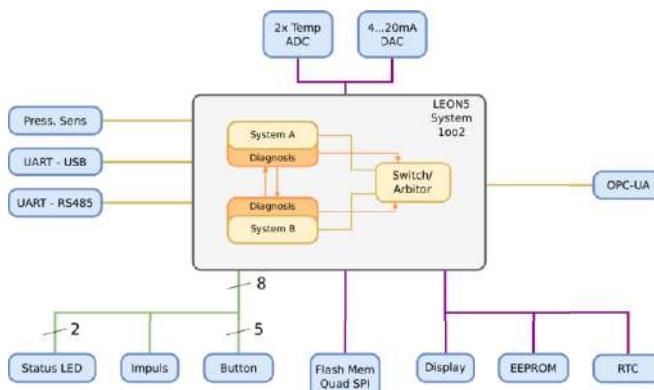
To achieve higher measurement accuracy the APT will be installed in a measuring armature with shut-off function. This ensures controlled installation conditions and easy access for maintenance and replacement.

For DP measurement a custom sensor from KROHNE, a well-known manufacturer of industrial process instrumentation, was selected. The sensor is specially designed for pressure profiles of compressed air compressors with a large DP measuring range as well as a higher sample rate to factor in pulsation effects.

The volume flow measurement device is designed in compliance with OIML 137 [21, 22]. For quality assurance the device will also be certified according to OIML 137.

Process connections from the DP sensor to the DP device are in compliance with EN 61518 in order to facilitate on-site in operating testing of the DP sensor. This also enables the installation of different DP devices according to ISO 5167 (if required by standards or regulations) since the device is in principle transparent to the SoC.

Further input parameters of the efficiency calculation to be measured according to ISO 1217 are inlet conditions (temperature, pressure, humidity) and the electric power

**Figure 4.** Schematic of the two-channel sensor SoC.

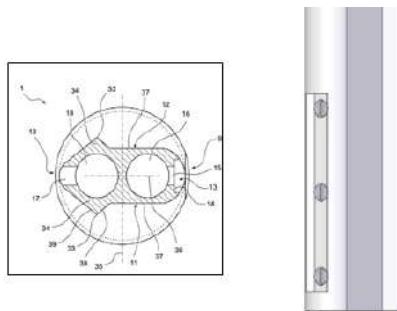


Figure 5. Left: Profile of the APT with differential pressure pipes; Right: View of the APT [German Patent 10 2021 109 414.3, 14.04.2021]

usage of the compressor. An overview of the primary measurands is presented in table 2. For data collection the SoC will provide a Modbus interface and analog input options.

Efficiency assessment, control and maintenance information as well as assessment according to international regulations will be calculated in real-time on the SoC and will be available via display and ethernet interface (OPC UA).

5. Conclusion

In this work, a new concept for an intelligent efficiency measurement system for industrial compressed air systems was presented. The specification of the concept was defined based on the international standards ISO 1217:2009/Amd 1:2016 and OIML 137. A new SoC design with high availability based on the 1oo2 architecture was introduced. Industry 4.0 connectivity is supported in this design by the OPC UA protocol.

For future developments comprehensive compressor control strategies and predictive maintenance assessment will be provided on the PLC for one or several compressed air compressors based on efficiency measurements from the proposed SoC. This concept will be further refined with machine learning capabilities of the PLC.

Since 80-93% of the electric energy used by an air compressor is converted into heat, a properly designed heat recovery has great savings potential [8]. For a more holistic efficiency calculation measurements with a heat meter are planned.

Acknowledgment

We acknowledge the Distr@1 funding program of the Hessian State Chancellery in the area of the Minister for Digital Strategy and Development, project 20_0097_2A *Zuverlässige SoC-Sensorik zur Effizienzsteigerung bei Hochleistungskompressorsystemen auf KI-Basis - Scientia*.

We acknowledge our colleagues and co-workers at the Department of Computer Architecture and System Programming, University of Kassel, Postberg+Co und inTec automation for their support, comments, and discussion.

References

- [1] Sullair, Uses of Compressed Air — Sullair. [Online]. Available: <https://america.sullair.com/en/blog/uses-compressed-air> (accessed: Aug. 10 2021).
- [2] Air Compressor Market Size, Share & Trends — Industry Report, 2025. [Online]. Available: <https://www.grandviewresearch.com/industry-analysis/air-compressor-market> (accessed: Aug. 10 2021).
- [3] “Compressed air in industry: how can digitalization help energy optimization?,” METRON, 24 Jun., 2021. <https://www.metron.energy/blog/compressed-air-industry-energy-optimization/> (accessed: Aug. 10 2021).
- [4] M. P. Roessler, “Energy efficiency in industrial compressed air generation-Survey about developments, trends and alternative generation technologies,” 2016.
- [5] Manuel Unger and Peter Radgen, Eds., Energy efficiency in compressed air systems-a review of energy efficiency potentials, technological development, energy policy actions and future importance, 2017.
- [6] Federal Ministry for Economics Affairs and Energy, Industrie 4.0. [Online]. Available: <https://www.bmwi.de/Redaktion/EN/Dossier/industrie-40.html> (accessed: Aug. 10 2021).
- [7] D. Cogliati, M. Falchetto, D. Pau, M. Roveri, and G. Viscardi, “Intelligent Cyber-Physical Systems for Industry 4.0,” in 2018 First International Conference on Artificial Intelligence for Industries (AI4I), Laguna Hills, CA, USA, Sep. 2018 - Sep. 2018, pp. 19–22.
- [8] Ying Zheng and Michael Bellstedt, “Compressor Degradation Assessment and Wear Mitigation Strategy,” Meat & Lifestock Australia Limited.
- [9] M. van Elburg, M. van der Voort, R. van den Boorn, R. Kemna, W. Li, and Van Holsteijn en Kemna, BV, “Study on Amended Working Plan under the Ecodesign Directive,” Final Report of Task, vol. 3, 2011.
- [10] Energy Conservation Program: Energy Conservation Standards for Air Compressors, Department of Energy.
- [11] Minimum allowable values of energy efficiency and energy efficiency grades for displacement air compressors, GB 19153-2019, SAMR AND SAC, Dec. 2019.
- [12] Displacement compressors — Acceptance tests, ISO 1217:2009/Amd 1:2016, International Organization for Standardization (ISO), Apr. 2016.
- [13] Turbocompressors — Performance test code, ISO 5389:2005, International Organization for Standardization (ISO), Dec. 2005.
- [14] American Society of Mechanical Engineers, Performance test code on compressors and exhausters: An American national standard. New York: The American Society of Mechanical Engineers, 1998.
- [15] J. Börscök, “Functional Safety: Basic Principles of Safety-related Systems,” Hüthig, Heidelberg, 2007.
- [16] R. W. Butler and S. C. Johnson, “Techniques for modeling the reliability of fault-tolerant systems with the Markov state-space approach,” 1995.
- [17] L. Xing, G. Levitin, and C. Wang, Dynamic System Reliability: Modeling and Analysis of Dynamic and Dependent Behaviors: Wiley, 2019. [Online]. Available: <https://books.google.com.eg/books?id=ee-BDwAAQBAJ>
- [18] A. Hayek, B. Machmur, M. Schreiber, J. Borcsok, S. Golz, and M. Epp, “HICore1: “Safety on a chip” turnkey solution for industrial control,” in 2014 IEEE 25th International Conference on Application-Specific Systems, Architectures and Processors, Zurich, Switzerland, Jun. 2014 - Jun. 2014, pp. 74–75.
- [19] Texas Instruments, Safety Manual for TMS570LS31x and TMS570LS21x Hercules™ and ARM® Safety Critical Microcontrollers: User’s Guide.
- [20] M. Abdelawwad, M. Drabesch, M. Schwarz, M. I. Hafiz, and J. Börscök, “Communication SoC based on 1002D architecture for industrial human-robot-collaboration,” in 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 2021, pp. 954–959.
- [21] J. Börscök, W. Müller, E. Hahn, M. Schwarz, and M. Abdelawwad, “Safe-System-on-Chip for Functional Safety,” in 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 2021, pp. 619–624.
- [22] Gas meters - Part 1: Metrological and technical requirements, OIML R 137-1, International Organization of Legal Metrology, 2012.
- [23] Gas meters - Part 2: Metrological controls and performance tests, OIML R 137-2, International Organization of Legal Metrology, 2012.
- [24] Measurement of fluid flow by means of pressure differential devices inserted in circular cross-section conduits running full: Part 1: General principles and requirements, ISO 5167-1:2003, Mar. 2003.

A New Identity-Based Encryption Scheme with Accountable Authority Based on SM9

Ke WANG ^{a,b}, Yuan ZHAO ^c, Song LUO ^c and Zhi GUAN ^{a,d,1}

^a*Key Laboratory of High Confidence Software Technologies (Peking University), MoE, Beijing, China*

^b*Department of Computer Science and Technology, EECS, Peking University, Beijing, China*

^c*School of Computer Science and Engineering, Chongqing University of Technology, Chongqing, China*

^d*National Engineering Research Center for Software Engineering, Peking University, Beijing, China*

Abstract. Accountable authority identity-based encryption (A-IBE) is an extension of identity-based encryption (IBE) in which private key's source can be traced, i.e., whether the key comes from a private key generator or a user. SM9 is an official cryptography standard of China which defines a practical IBE scheme. In this paper, we construct a practical A-IBE scheme from the SM9-IBE scheme. Our A-IBE scheme has public traceability and is proven secure if the based SM9-IBE scheme is secure. Compared with other A-IBE schemes, our A-IBE scheme has better efficiency in encryption and decryption.

Keywords. Accountable authority, identity-based encryption, SM9, bilinear group

1. Introduction

Identity-based encryption (IBE) [1] was proposed to simplify the public-key infrastructure. Unlike traditional public key encryption, any meaningful strings related to a user's identity, such as IP address, e-mail address, phone number or ID number, can be used to form a public key. Messages can be encrypted to any identity, but the ciphertext can only be decrypted by the owner of the target identity.

Though IBE has many appealing advantages, it has an inherent problem called key escrow. A trusted authority named as private key generator(PKG) exists in an IBE scheme which selects the system parameters, generates master keys and all users' private keys. If a private key is used for illegal purposes, it is hard to distinguish this key comes from the PKG or the user.

The key escrow problem is hard to solve until Goyal [2] introduced the concept of accountable authority IBE (A-IBE). In A-IBE schemes, to obtain his own private key,

¹Corresponding Author: guan@pku.edu.cn

the user should run a interactive key generation protocol with the PKG. Only the user knows the generated private key. If the PKG is dishonest and generates another key for malicious usage, we can use an additional tracing algorithm to catch it and sue it in the court of law.

SM9 [3] is an official cryptography standard of China in which a set of pairing-based cryptographic schemes from pairings are defined, including identity-based encryption, digital signature, authenticated key exchange protocol and one recommended 256-bit BN curve. SM9 can be implemented in different platforms and it has better computation efficiency and shorter ciphertext than three other schemes included in ISO/IEC 18033-5 [4]. Nowadays, the International Organization for Standardization adopts the signature scheme and the encryption scheme of SM9 as ISO/IEC 14888-3:2018 [5] and ISO/IEC 18033-5 [6], respectively.

Our Contribution. In our paper, we build a new A-IBE scheme based on the SM9-IBE scheme. As a result, our A-IBE scheme conforms to the Chinese standard and the international standard since the based SM9-IBE scheme is standard. To the best of our knowledge, it is the first A-IBE scheme which is compatible with the Chinese cryptographic standard. Our A-IBE scheme has public traceability, i.e., everyone can trace the source of a decryption key with the help of a public tracing key. Furthermore, it is secure and efficient. Our A-IBE scheme is proven secure if the based SM9-IBE scheme is secure. Analysis shows that our A-IBE scheme has better efficiency in encryption and decryption than other A-IBE schemes, which is very important for current online applications.

Related Works. Goyal [2] first introduced the notion of A-IBE and he proposed a white-box A-IBE scheme and a weak black-box A-IBE scheme, respectively. Goyal et al. [7] presented the first full black-box A-IBE scheme. A weak black-box A-IBE scheme was proposed by Libert and Vergnaud [8] in which the private keys and ciphertexts have constant sizes. Lai et al. [9] proposed the first A-IBE scheme with public traceability in which the tracing key is public and anyone can trace a decryption key. Two generic constructions of A-IBE were proposed by Sahai and Seyalioglu [10], Kiayias and Tang [11] respectively. The former is full black-box secure while the latter is weak black-box secure. To achieve full and efficient black-box A-IBE, Zhao et al. [12] presented a new generic construction and gave an efficient instantiation from Park-Lee IBE scheme [13]. Recently, Zhao et al. [14] and Zhao et al. [15] extend A-IBE to accountable authority identity-based broadcast encryption and accountable authority identity-based revocable encryption, respectively.

Organization. The rest of this paper is organized as follows. Some necessary background knowledge is provided in Section 2. We first propose a modified SM9-IBE scheme in Section 3. Based on this modified SM9-IBE scheme, we then propose a new A-IBE scheme and prove its security in Section 4. Next we give a brief analysis for our A-IBE scheme in Section 5. In Section 6 the paper is concluded with future work.

2. Preliminaries

Definition 2.1. Let $\mathbb{G}_1, \mathbb{G}_2$ be two additive cyclic group and \mathbb{G}_T be a multiplicative groups and they all have prime order N . Let P_1 be a generator of \mathbb{G}_1 while P_2 be a generator of \mathbb{G}_2 . Let $e : \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ be a bilinear map which satisfies the following properties:

- (i) $\forall x, y \in \mathbb{Z}_N, e([x]P_1, [y]P_2) = e(P_1, P_2)^{xy}$ (Bilinearity).
- (ii) $e(P_1, P_2) \neq 1$ (Non-degeneracy).

We assume that the bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$ and the group operation in \mathbb{G}_1 and \mathbb{G}_2 can be computed efficiently. We also assume that $(N, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e, P_1, P_2)$ can be obtained by an efficient algorithm \mathcal{G} which takes a system security parameter λ as input.

If there exists an isomorphism between \mathbb{G}_1 and \mathbb{G}_2 (or \mathbb{G}_1 equals to \mathbb{G}_2), we say that the bilinear groups are symmetric. Symmetric groups are easier to describe the construction of a scheme but have less computation efficiency than asymmetric groups.

The following four algorithms are included in an IBE scheme: the **Setup** algorithm for system setup, the **KeyGen** algorithm to generate private keys, the **Encrypt** algorithm to encrypt messages, and the **Decrypt** algorithm to decrypt ciphertext. The accountable authority IBE scheme has two extra algorithms: **Trace** algorithm is used to trace a decryption box \mathbb{D}_{ID} whether it is generated from the PKG or the user with the help of the public tracing token list, and **Judge** algorithm is used to judge a tracing key for some identity or not.

The CPA security or semantic security game defined for IBE schemes consists of five stages: **Setup**, **Query Phase 1, 2**, **Challenge** and **Guess**. The adversary submits a challenging identity ID^* and two equal-length messages M_1, M_2 at the **Challenge** stage and can query any identities in **Query Phase 1, 2** except the challenging identity. Finally at the **Guess** stage the adversary will be returned a ciphertext encrypted to ID^* and M_μ where $\mu \in \{0, 1\}$ is random. The adversary should give a guess μ' of μ and his advantage is $|\Pr[\mu' = \mu] - \frac{1}{2}|$. An IBE scheme is semantically secure if no p.p.t adversary has non-negligible advantage in winning the above game. A-IBE has two extra security definitions called **DishonestPKG** security which is used to trace whether a decryption box(white-box or black-box) is generated by the PKG, while **DishonestUser** security is used to prove that nobody except the key owner can produce a valid decryption box.

3. A Modified SM9-IBE Scheme

We present a modified SM9-IBE scheme in this section. Compared with the original SM9-IBE scheme, we simplify the encryption and the decryption algorithm. This modified scheme has only semantic (CPA) security, while the initial SM9-IBE scheme is CCA2 secure. We use this modified scheme because in the security proof we would modify the original SM9-IBE ciphertext to a valid A-IBE ciphertext. However, this is infeasible since the original SM9-IBE scheme has an integrity check for its ciphertext.

Let $H_1()$ be a cryptographic hash function and hid be the identifier of the encryption private key generating function, we propose the modified SM9-IBE scheme as follows.

- **Setup**(1^λ): The algorithm gets public parameter $\text{PP} = (N, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, P_1, P_2, e)$ from $\mathcal{G}(\lambda)$. It randomly chooses an integer $ke \in [1, N - 1]$, computes $P_{\text{pub}-e} = [ke]P_1$ and $\Omega = e(P_{\text{pub}-e}, P_2)$. The public key is $\text{PK} = (\text{PP}, P_{\text{pub}-e}, \Omega)$ and the master secret key MSK is ke .
- **KeyGen**(MSK, ID): The algorithm first computes $t_1 = H_1(\text{ID} || hid, N) + ke$. If $t_1 = 0$, it runs the **Setup** algorithm again. Otherwise, it computes $t_2 = ke \cdot t_1^{-1}$ and $K = [t_2]P_2$. The private key for ID is $\text{SK}_{\text{ID}} = (K)$.

- **Encrypt**($\text{PK}, M \in \mathbb{G}_T, \text{ID}$): The algorithm randomly chooses an integer $s \in [1, N - 1]$ and computes $Q_{\text{ID}} = [H_1(\text{ID} || hid, N)]P_1 + P_{\text{pub}-e}$. It then computes $C_1 = [s]Q_{\text{ID}}$ and $C = M \cdot \Omega^s$. The ciphertext is $\text{CT}_{\text{ID}} = (C, C_1)$.
- **Decrypt**($\text{CT}_{\text{ID}} = (C, C_1), \text{SK}_{\text{ID}} = (K)$): The algorithm computes $Y = e(C_1, K)$ and the message M is recovered as $C \cdot Y^{-1}$.

Theorem 3.1. *The modified SM9-IBE scheme is semantically secure.*

The proof of this theorem is similar to the proof of the original SM9-IBE scheme except using the decryption oracle [4], so the concrete proof is omitted due to space limitation.

4. Our A-IBE Scheme

4.1. Description

We propose our A-IBE scheme which has public traceability as follows:

- **Setup**(1^λ): The algorithm first gets public parameter $\text{PP} = (N, \mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, P_1, P_2, e)$ from $\mathcal{G}(\lambda)$. It chooses two random integers $ke, \alpha \in [1, N - 1]$, computes $P_{\text{pub}-a} = [\alpha]P_1$, $P_{\text{pub}-e} = [ke]P_1$, $P_{\text{pub}-t} = [ke \cdot \alpha]P_1$ and $\Omega = e(P_{\text{pub}-e}, P_2)$. The public key PK is $(\text{PP}, P_{\text{pub}-a}, P_{\text{pub}-e}, P_{\text{pub}-t}, \Omega)$ and the master secret key MSK is (ke, α) .
- **KeyGen**(MSK, ID): The PKG interacts with a user U for identify ID in the following protocol to generate key.
 - U first randomly chooses $r \in [1, N - 1]$. Next, U computes $R = [r]P_2$ and sends it to the PKG. It runs an interactive zero-knowledge proof of knowledge (ZK-POK) with the PKG for the discrete log of R with respect to P_2 .
 - PKG checks that the validity of the ZK-POK and it will abort if the check fails. Otherwise, the PKG randomly picks two integers $t, r' \in [1, N - 1]$ and sends (R', r', K', Q', S', V') to U, where

$$R' = [\alpha]R = [r \cdot \alpha]P_2, K' = \left[\frac{ke}{H_1(\text{ID} || hid, N) + ke} \right] P_2 - [t \cdot \alpha]P_2,$$

$$Q' = [t \cdot \alpha]R + [r' \cdot \alpha]P_2, S' = [t]P_2, V' = [t \cdot \alpha]P_1.$$

- U checks the validity of the following three equations:

$$e(P_{\text{pub}-t}, R) = e(P_{\text{pub}-e}, R') \quad (1)$$

$$\begin{aligned} & e([H_1(\text{ID} || hid, N)]P_1 + P_{\text{pub}-e}, K') \\ & = e(P_{\text{pub}-e}, P_2) \cdot e([H_1(\text{ID} || hid, N)]P_{\text{pub}-a} + P_{\text{pub}-t}, S') \end{aligned} \quad (2)$$

$$e(P_1, Q') = e(V', R) \cdot e(P_{pub-a}, P_2)^{r'} \quad (3)$$

U will abort if the check fails. Otherwise, the decryption key SK_{ID} is (r_{ID}, K) , where $r_{\text{ID}} = r'/r$ and

$$K = K' + [r^{-1}]Q' = \left[\frac{ke}{H_1(\text{ID}||hid, N) + ke} + \frac{r'}{r} \cdot \alpha \right] P_2.$$

- (iv) U sends $R_{\text{ID}} = e([H_1(\text{ID}||hid, N)]P_{pub-a} + P_{pub-t}, P_2)^{r_{\text{ID}}}$ and interacts with the PKG for a zero-knowledge proof $\text{ZK}\{(r_{\text{ID}}, K) : R_{\text{ID}} = e(P_{pub-t}, P_2)^{r_{\text{ID}}} \wedge e([H_1(\text{ID}||hid, N)]P_1 + P_{pub-e}, K) = e(P_{pub-e}, P_2) \cdot R_{\text{ID}}\}$ with the PKG.
- (v) PKG verifies the ZK proof. If the check fails, it will abort. Finally the tracing key $t_{\text{ID}} = (\text{ID}, R_{\text{ID}})$ is added to a public tracing key list \mathcal{TK} .
- **Encrypt**(PK, M , ID): The algorithm randomly chooses an integer $s \in [1, N - 1]$ and computes $Q_{\text{ID}} = [H_1(\text{ID}||hid, N)]P_1 + P_{pub-e}$. It then computes $Z_1 = [s]Q_{\text{ID}}$ and $C = M \cdot \Omega^s$ and $Z_2 = e([H_1(\text{ID}||hid, N)]P_{pub-a} + P_{pub-t}, P_2)^s$. The ciphertext CT_{ID} is (C, Z_1, Z_2) .
- **Decrypt**($\text{CT}_{\text{ID}} = (C, Z_1, Z_2)$, $\text{SK}_{\text{ID}} = (r_{\text{ID}}, K)$): The algorithm computes $Y = e(C_1, K)$ and the message M is recovered as $C \cdot Y^{-1} \cdot C_2^{r_{\text{ID}}}$.
- **Trace**(PK, ID, \mathcal{TK} , \mathbb{D}_{ID}): Let \mathbb{D}_{ID} be an ε -useful decryption box. The algorithm will abort if the tuple $(\text{ID}, R_{\text{ID}})$ isn't in the list of public tracing keys \mathcal{TK} . It then runs the following steps.

- (i) Let $ctr = 0$ be a counter. The following steps are repeated for $8\lambda/\varepsilon$ times:
 - Choose two random integers $s, s' \in [1, N - 1]$ and $s \neq s'$.
 - Compute $Q_{\text{ID}} = [H_1(\text{ID}||hid, N)]P_1 + P_{pub-e}$.
 - A random message M is randomly chosen in \mathbb{G}_T . Compute $Z_1 = [s]Q_{\text{ID}}$, $Z_2 = e([H_1(\text{ID}||hid, N)]P_{pub-a} + P_{pub-t}, P_2)^{s'}$, and $C = M \cdot \Omega^s \cdot R_{\text{ID}}^{(s-s')}$.
 - The ciphertext $\text{CT} = (C, Z_1, Z_2)$ is put into \mathbb{D}_{ID} . Let \mathbb{D}_{ID} 's output be M' . If $M = M'$, $ctr = ctr + 1$.
- (ii) The output is PKG if $ctr = 0$. If $ctr \neq 0$, the output is User.
- **Judge**: Assume that U has identity ID. He interacts with a judge in the following protocol.
 - (i) Let $\text{SK}_{\text{ID}} = (r_{\text{ID}}, K)$ be the user's decryption key. U first sets $R_{\text{ID}} = e(P_{pub-t}, P_2)^{r_{\text{ID}}}$. U then sends R_{ID} and interacts with the judge for a ZK proof $\text{ZK}\{(r_{\text{ID}}, K) : R_{\text{ID}} = e(P_{pub-t}, P_2)^{r_{\text{ID}}} \wedge e([H_1(\text{ID}||hid, N)]P_1 + P_{pub-e}, K) = e(P_{pub-e}, P_2) \cdot R_{\text{ID}}\}$.
 - (ii) The judge verifies the ZK proof. If the check fails, he will abort. Finally $t_{\text{ID}} = (\text{ID}, R_{\text{ID}})$ is accepted as U's public tracing key by the judge.

Correctness.

$$\begin{aligned}
e(C_1, K)^{-1} &= e([s \cdot (H_1(\text{ID}||hid, N) + ke)]P_1, [\frac{ke}{H_1(\text{ID}||hid, N) + ke} + r_{\text{ID}} \cdot \alpha]P_2)^{-1} \\
&= e(P_1, P_2)^{-s \cdot ke - s \cdot r_{\text{ID}} \cdot (H_1(\text{ID}||hid, N) + ke) \cdot \alpha} \\
C_2^{r_{\text{ID}}} &= e(P_1, P_2)^{s \cdot r_{\text{ID}} \cdot (H_1(\text{ID}||hid, N) + ke) \cdot \alpha} \\
C \cdot e(C_1, K)^{-1} \cdot C_2^{r_{\text{ID}}} &= M \cdot e(P_1, P_2)^{s \cdot ke} \cdot e(P_1, P_2)^{-s \cdot ke} = M.
\end{aligned}$$

4.2. Security

Theorem 4.1. *If the modified SM9-IBE scheme is semantically secure, then our A-IBE scheme is also semantically secure.*

Proof. We will show that we can construct algorithm \mathcal{B} which breaks the modified SM9-IBE scheme based on an adversary \mathcal{A} which can break our A-IBE scheme. \mathcal{B} is constructed as follows.

Setup. At first the challenger \mathcal{C} outputs the public key of the modified SM9-IBE scheme $\text{PK}' = \{\text{PP}, P_{\text{pub}-e}, \Omega = e(P_{\text{pub}-e}, P_2)\}$ and sends it to \mathcal{B} . \mathcal{B} randomly chooses an integer $\alpha \in [N - 1]$, computes $P_{\text{pub}-t} = (P_{\text{pub}-e})^\alpha, P_{\text{pub}-a} = [\alpha]P_1$. The new public key PK is $\{e, P_1, P_2, P_{\text{pub}-a}, P_{\text{pub}-e}, P_{\text{pub}-t}, \Omega\}$ and $\{\alpha\}$ is kept secret. Then PK is published to \mathcal{A} . Note that the master key of the original SM9-IBE scheme is unknown to \mathcal{B} .

Query Phase 1. An identity ID is submitted by \mathcal{A} . \mathcal{C} returns \mathcal{B} the private key for ID $\text{SK}'_{\text{ID}} = (K_0)$. Then \mathcal{B} interacts with \mathcal{A} in the following protocol.

- (i) \mathcal{A} first randomly chooses $r \in [1, N - 1]$ and computes $R = [r]P_2$. It then sends R and interacts with \mathcal{B} for a ZK-POK of the discrete log of R in respect of P_2 .
- (ii) \mathcal{B} verifies the ZK-POK. It will aborts if the check fails. \mathcal{B} randomly chooses two integers $t, r' \in [1, N - 1]$ and sends (R', r', K', Q', S', V') to \mathcal{U} , where $R' = [\alpha]R, K' = K_0 - [t \cdot \alpha]P_2, Q' = [t \cdot \alpha]R + [r' \cdot \alpha]P_2, S' = [t]P_2$ and $V' = [t \cdot \alpha]P_1$.
- (iii) \mathcal{A} checks whether Equations 1, 2 and 3 hold or not. \mathcal{A} will abort if the check fails. \mathcal{A} sets $\text{SK}_{\text{ID}} = (r_{\text{ID}}, K)$, where $r_{\text{ID}} = r'/r$ and $K = K' + [r^{-1}]Q'$.
- (iv) \mathcal{A} sends $R_{\text{ID}} = e([H_1(\text{ID}||\text{hid}, N)]P_{\text{pub}-a} + P_{\text{pub}-t}, P_2)^{r_{\text{ID}}}$ to \mathcal{B} and interacts with \mathcal{B} for a ZK proof $\text{ZK}\{(r_{\text{ID}}, K) : R_{\text{ID}} = e(P_{\text{pub}-t}, P_2)^{r_{\text{ID}}} \wedge e([H_1(\text{ID}||\text{hid}, N)]P_1 + P_{\text{pub}-e}, K) = e(P_{\text{pub}-e}, P_2) \cdot R_{\text{ID}}\}$.
- (v) \mathcal{B} verifies the ZK proof. If the check fails, he will abort. Finally $t_{\text{ID}} = (\text{ID}, R_{\text{ID}})$ for the identify ID is added to the public tracing key list \mathcal{TH} .

Challenge. Let ID^* be a challenging identity and M_0, M_1 be two equal-length messages submitted by \mathcal{A} . \mathcal{B} forwards ID^* and M_0, M_1 to \mathcal{C} and is returned $\text{CT}' = (C, Z_1)$ for some message Z_μ ($\mu \in \{0, 1\}$). It computes $Z_2 = e(Z_1, [\alpha]P_2)$.

Query Phase 2. \mathcal{A} can query like in **Query Phase 1** except ID^* .

Guess. \mathcal{B} outputs the guess μ' based on \mathcal{A} 's guess μ' of μ .

From above we can find that the advantage of \mathcal{A} equals to \mathcal{B} , so if the advantage for \mathcal{A} to break our A-IBE scheme is non-negligible, the advantage for \mathcal{B} to break the modified SM9-IBE scheme is non-negligible. Hence, according to Theorem 3.1, this theorem is established. \square

The following modified DDH-2 assumption is used to prove the **DishonestPKG** security of our A-IBE scheme.

Definition 4.1. Let a, b are two random integers in $[1, N - 1]$ and Z be randomly chosen in \mathbb{G}_T . Let $\vec{D} = \{[a]P_2, [1/a]P_2, e(P_1, P_2)^b\}$. For an algorithm \mathcal{A} , we define its advantage to break the modified DDH-2 assumption to be

$$|\Pr[\mathcal{A}(\vec{D}, e(P_1, P_2)^{b/a}) = 1] - \Pr[\mathcal{A}(\vec{D}, Z) = 1]|.$$

For any p.p.t algorithm, if its advantage to break the modified DDH-2 assumption is non-negligible, we say that the modified DDH-2 assumption holds.

Theorem 4.2. Under the modified DDH-2 assumption, our A-IBE scheme is **DishonestPKG** secure.

Proof. Assume that the adversary \mathcal{A} can break the **DishonestPKG** security of our A-IBE scheme. This means \mathcal{A} can produce a valid decryption box \mathbb{D}_{ID} with $\text{Trace}(\text{PK}, \text{ID}, \mathcal{K}, \mathbb{D}_{\text{ID}}) = \text{User}$. We can construct a new algorithm \mathcal{B} based on \mathcal{A} to break the modified DDH-2 assumption.

\mathcal{B} receives an instance of the modified DDH-2 assumption $([a]P_2, [1/a]P_2, e(P_1, P_2)^b, T)$ and tries to decide whether $T = e(P_1, P_2)^{b/a}$ or not. \mathcal{A} interacts with \mathcal{B} as follows.

Setup. The public key $\text{PK} = (\text{PP}, P_{\text{pub}-a}, P_{\text{pub}-e}, P_{\text{pub}-t}, \Omega)$ and a challenge identity ID^* is sent from \mathcal{A} to \mathcal{B} .

Key Generation. The key for ID^* is generated between \mathcal{B} and \mathcal{A} from the following steps. First it computes $R = [a]P_2$ and sends it to \mathcal{A} . Then \mathcal{B} plays a ZK-PoK proof with \mathcal{A} for the discrete log of R in respect of P_2 . Though \mathcal{B} doesn't know a , he can simulate the right ZK-PoK proof. Then, \mathcal{B} receives (R', r', K', Q', S', V') from \mathcal{A} . \mathcal{B} checks whether Equations 1, 2 and 3 hold or not. If no, \mathcal{B} aborts. The decryption key $\text{SK}_{\text{ID}} = (r_{\text{ID}^*}, K)$ for ID^* is set as $(r'/a, \left[\frac{ke}{H_1(\text{ID}^*||hid, N) + ke} + \frac{r'}{a} \cdot \alpha \right] P_2)$ (but unknown to \mathcal{B}).

Next \mathcal{B} sends $R_{\text{ID}^*} = e([H_1(\text{ID}^*||hid, N)]P_{\text{pub}-a} + P_{\text{pub}-t}, [1/a]P_2)^{r'}$ to \mathcal{A} and gives a ZK proof for (r_{ID^*}, K) . Equally, \mathcal{B} can simulate the right ZK proof without knowing (r_{ID^*}, K) . At last, the tracing key $t_{\text{ID}^*} = (\text{ID}^*, R_{\text{ID}^*})$ is added to the public tracing key list \mathcal{K} .

Output. At this stage, \mathcal{A} outputs a decryption box \mathbb{D}_{ID^*} for ID^* . \mathcal{B} runs the following steps in the tracing stage.

- (i) Randomly choose n messages $M_1, \dots, M_n \in \mathbb{G}_T$. For $i \in \{1, 2, \dots, n\}$, the i -th ciphertext is computed as $\text{CT}^{(i)} = (C^{(i)}, Z_1^{(i)}, Z_2^{(i)})$ as

$$C^{(i)} = M \cdot \Omega^{s_i} \cdot R_{\text{ID}^*}^{(s_i - s'_i)} \cdot T^{-r' \cdot \delta_i}, C_1^{(i)} = [s_i]Q,$$

$$C_2^{(i)} = e([H_1(\text{ID}^*||hid, N)]P_{\text{pub}-a} + P_{\text{pub}-t}, P_2)^{s'} \cdot e(P_1, P_2)^{b \cdot \delta_i}$$

where $Q = [H_1(\text{ID}^*||hid, N)]P_1 + P_{\text{pub}-e}$ and δ_i, s_i, s'_i are chosen from $[1, N - 1]$ randomly.

- (ii) Set a counter $ctr = 0$. Feed \mathbb{D}_{ID^*} with n ciphertexts $\text{CT}^{(1)}, \dots, \text{CT}^{(n)}$ and get decrypted messages M'_1, \dots, M'_n .
- (iii) Set $ctr = ctr + 1$ for every $i = 1$ to n where $M_i = M'_i$.

Analysis. If $T = e(P_1, P_2)^{b/a}$, every $CT^{(i)}$ is an ill-formed ciphertext but can be decrypted by \mathbb{D}_{ID^*} . Hence, if the decryption box \mathbb{D}_{ID^*} is valid, we will get $ctr > 0$ if $T = e(P_1, P_2)^{b/a}$ or $ctr = 0$ if T is random. As a result, the modified DDH-2 assumption is broken. \square

Theorem 4.3. *If the SM9-IBE scheme has semantic security, our A-IBE scheme has **DishonestUser** security.*

Proof. The proof follows the proof of Theorem 4.1. Assume that our A-IBE scheme isn't **DishonestUser** secure. Then there is an adversary \mathcal{A} can output a decryption box for an identity ID . Hence we can use \mathcal{A} to produce a decryption box for ID^* . Therefore, this decryption box can be used to decrypt the returned ciphertext and get a message M_μ . As a result, the modified SM9-IBE scheme is broken. \square

5. Analysis

We now analyze our A-IBE scheme from theoretical perspective and experimental perspective. We give comparisons with other recent A-IBE schemes in the following tables. In Table 1 we compare the public key size($|\mathbf{PK}|$), the private key size($|\mathbf{SK}|$) and the ciphertext size($|\mathbf{CT}|$). In Table 2 we compare the time of encryption(\mathbf{Time}_{enc}) and the time of decryption(\mathbf{Time}_{dec}), respectively. Experiment was run on a computer with Intel Core i5-8250U 1.60GHz, 8G RAM and Ubuntu 16.04. Analysis shows our A-IBE scheme is more desirable in practice. First our scheme is compact in key size and ciphertext size. Second our scheme has shorter running time in encryption and decryption. Note that operations in asymmetric bilinear groups are more efficient than symmetric groups. Therefore, though the storage requirement of our scheme is the same as Lai et al.'s A-IBE scheme, our scheme is more efficient in encryption and decryption.

Table 1. Storage Comparison

Scheme	Group	$ \mathbf{PK} $	$ \mathbf{SK} $	$ \mathbf{CT} $
Sahai et al.[2011]	Symmetric	$(2m+2) \mathbb{G} + \mathbb{G}_T $	$2n \mathbb{G} + n \mathbb{Z}_m $	$2n \mathbb{G} + n \mathbb{Z}_m + \mathbb{G}_T $
Lai et al.[2013]	Symmetric	$3 \mathbb{G} + \mathbb{G}_T $	$3 \mathbb{G} + \mathbb{Z}_N $	$2 \mathbb{G} + \mathbb{Z}_m + \mathbb{G}_T $
Zhao et al.[2019]	Symmetric	$3 \mathbb{G} + \mathbb{G}_T $	$ \mathbb{G} + \mathbb{Z}_N $	$2 \mathbb{G} + \mathbb{G}_T $
Our scheme	Asymmetric	$3 \mathbb{G}_1 + \mathbb{G}_T $	$ \mathbb{G}_2 + \mathbb{Z}_N $	$2 \mathbb{G}_1 + \mathbb{G}_T $

Note: we omit the common public parameters in public key. $m = |\mathbf{ID}|$ and n is a constant fraction of m . $|\cdot|$ means the element size of \cdot .

6. Conclusion

We propose a new A-IBE scheme based on a modified SM9-IBE scheme in this paper. Our A-IBE scheme has public traceability and is proven secure the based modified

Table 2. Computation Comparison(ms)

Scheme	Time _{enc}	Time _{dec}
Sahai et al.[2011]	359	2,276
Lai et al.[2013]	47	35
Zhao et al.[2019]	51	113
Our scheme	28	19

SM9-IBE scheme is secure. Experiments show that our A-IBE scheme has better efficiency in encryption and decryption than other A-IBE schemes. However, our scheme has only CPA security and weak black-box **DishonestPKG** security. We leave it an open problem to find a new A-IBE scheme based on the full SM9-IBE scheme with public traceability and other security properties.

Acknowledgements

This work is supported by the National Key Research and Development Program of China NO.2018YFB0803601.

References

- [1] Adi Shamir. Identity-based cryptosystems and signatures schemes. In *Advances in Cryptology - Crypto 1984*, volume 196 of *LNCS*, pages 47–53. Springer-Verlag, 1984.
- [2] Vipul Goyal. Reducing trust in the PKG in identity based cryptosystems. In Alfred Menezes, editor, *Advances in Cryptology - CRYPTO 2007, 27th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2007, Proceedings*, volume 4622 of *Lecture Notes in Computer Science*, pages 430–447. Springer, 2007.
- [3] GMT 0044-2016. Sm9 identity-based cryptographic algorithms, 2016. <http://www.gmbz.org.cn/main/postDetail.html?id=20180322410400>.
- [4] Zhaohui Cheng. The sm9 cryptographic schemes. Cryptology ePrint Archive, Report 2017/117, 2017. <https://ia.cr/2017/117>.
- [5] ISO/IEC. It security techniques - digital signatures with appendix - part 3: Discrete logarithm based mechanisms. Geneva, Switzerland, 2018. ISO/IEC14888-3:2018.
- [6] ISO/IEC. Information technology - security techniques encryption algorithms - part 5: Identity-based ciphers; amendment 1: Sm9 mechanism. Geneva, Switzerland, 2021. ISO/IEC 18033-5.
- [7] Vipul Goyal, Steve Lu, Amit Sahai, and Brent Waters. Black-box accountable authority identity-based encryption. In *Proceedings of the 15th ACM Conference on Computer and Communications Security*, CCS '08, pages 427–436, New York, NY, USA, 2008. ACM.
- [8] Benoît Libert and Damien Vergnaud. Towards black-box accountable authority IBE with short ciphertexts and private keys. In Stanislaw Jarecki and Gene Tsudik, editors, *Public Key Cryptography - PKC 2009, 12th International Conference on Practice and Theory in Public Key Cryptography, Irvine, CA, USA, March 18-20, 2009. Proceedings*, volume 5443 of *Lecture Notes in Computer Science*, pages 235–255. Springer, 2009.
- [9] Junzuo Lai, Robert H. Deng, Yunlei Zhao, and Jian Weng. Accountable authority identity-based encryption with public traceability. In Ed Dawson, editor, *Topics in Cryptology - CT-RSA 2013 - The Cryptographers' Track at the RSA Conference 2013, San Francisco, CA, USA, February 25-March 1, 2013. Proceedings*, volume 7779 of *LNCS*, pages 326–342. Springer, 2013.
- [10] Amit Sahai and Hakan Seyalioglu. Fully secure accountable-authority identity-based encryption. In Dario Catalano, Nelly Fazio, Rosario Gennaro, and Antonio Nicolosi, editors, *Public Key Cryptography - PKC 2011 - 14th International Conference on Practice and Theory in Public Key Cryptography, Taormina, Italy, March 6-9, 2011. Proceedings*, volume 6571 of *Lecture Notes in Computer Science*, pages 296–316. Springer, 2011.

- [11] Aggelos Kiayias and Qiang Tang. Making any identity-based encryption accountable, efficiently. In Günther Pernul, Peter Y. A. Ryan, and Edgar R. Weippl, editors, *Computer Security - ESORICS 2015 - 20th European Symposium on Research in Computer Security, Vienna, Austria, September 21-25, 2015, Proceedings, Part I*, volume 9326 of *Lecture Notes in Computer Science*, pages 326–346. Springer, 2015.
- [12] Zhen Zhao, Jianchang Lai, Willy Susilo, Baocang Wang, Yupu Hu, and Fuchun Guo. Efficient construction for full black-box accountable authority identity-based encryption. *IEEE Access*, 7:25936–25947, 2019.
- [13] Jong Hwan Park and Dong Hoon Lee. An efficient IBE scheme with tight security reduction in the random oracle model. *Des. Codes Cryptogr.*, 79(1):63–85, 2016.
- [14] Zhen Zhao, Fuchun Guo, Jianchang Lai, Willy Susilo, Baocang Wang, and Yupu Hu. Accountable authority identity-based broadcast encryption with constant-size private keys and ciphertexts. *Theor. Comput. Sci.*, 809:73–87, 2020.
- [15] Zhen Zhao, Ge Wu, Fuchun Guo, Willy Susilo, Yi Mu, Baocang Wang, and Yupu Hu. Black-box accountable authority identity-based revocation system. *Comput. J.*, 63(4):525–535, 2020.

Incremental 2D Grid Map Generation from RGB-D Images

Tingfeng YE^{a,b}, Juzhong ZHANG^{b,1}, Yingcai WAN^c, Ze CUI^a and Hongbo YANG^b

^a*School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China*

^b*Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou, China*

^c*Dept. Faculty of Robot Science and engineering, Northeastern University, Shenyang, China.*

Abstract. In this paper, we extend RGB-D SLAM to address the problem that sparse map-building RGB-D SLAM cannot directly generate maps for indoor navigation and propose a SLAM system for fast generation of indoor planar maps. The system uses RGBD images to generate positional information while converting the corresponding RGBD images into 2D planar lasers for 2D grid navigation map reconstruction of indoor scenes under the condition of limited computational resources, solving the problem that the sparse point cloud maps generated by RGB-D SLAM cannot be directly used for navigation. Meanwhile, the pose information provided by RGB-D SLAM and scan matching respectively is fused to obtain a more accurate and robust pose, which improves the accuracy of map building. Furthermore, we demonstrate the function of the proposed system on the ICL indoor dataset and evaluate the performance of different RGB-D SLAM. The method proposed in this paper can be generalized to RGB-D SLAM algorithms, and the accuracy of map building will be further improved with the development of RGB-D SLAM algorithms.

Keywords. RGB-D SLAM, layout SLAM, pose fusion, navigation, 2D grid maps

1. Introduction

SLAM (Simultaneous localization and mapping) is used to estimate the camera pose and reconstruct unknown environments, which is currently widely used in various fields, relying on sensors to achieve functions such as autonomous positioning, mapping, and path planning of the machine. For example, automatic navigation of robots, unmanned driving of cars, AR/VR technology positioning and three-dimensional reconstruction of objects.

At present, most of the lidar SLAM used to generate indoor navigation maps use lidar as the main sensor. For example, Gmapping[1], Hector SLAM[2], and Cartographer[3] can use single-line lidar to generate more accurate indoor navigation grid maps, but can only explore the environment in two dimensions. At the same time, single-line lidar equipment is large, limited by the fact that it must use a mechanical

¹ Corresponding author: Juzhong Zhang, Suzhou Institute of Biomedical Engineering and Technology Chinese Academy of Science, Suzhou, China; E-mail: jzzhang@sibet.ac.cn

motor to complete the rotation to obtain the exploration of the unknown environment. And the data obtained is with distortion [1, 2], which requires its matching SLAM algorithm to perform pre-processing to remove distortion from the data, resulting in reduced real-time and accuracy of the algorithm operation. There are also LOAM[4] which using multi-line lidar, and V-LOAM[5], LeGo-LOAM[6] based on [4] combined with other sensors and improved, which obtains spatial 3D information, also brings the problem of high cost [7] and excessive computational effort. Loam_livox [7] which uses solid-state lidar as the sensor that characteristics of laser radar have brought major challenges to the navigation and mapping of lidar. The above SLAM algorithms that use lidar as the main sensor are all limited by the defects of the laser sensor itself.

In this environment, the RGB-D sensor provides an opportunity to significantly develop the robot's indoor navigation and interaction capabilities [8]. At present, for SLAM algorithms, RGB-D cameras have been introduced and three-dimensional maps can be created in real-time, and a variety of different RGB-D SLAM algorithms have been proposed. Most of these RGB-D SLAM are used for indoor localization and object dense reconstruction [9]. Mono SLAM [10] and ORB-SLAM2 [11] based on the feature point method can directly obtain the camera's pose in space and the sparse point cloud map, but the obtained map cannot be directly used for navigation. DVO-SLAM [12] based on the direct method estimates the motion of the camera according to the gradient information of the pixels, and can construct a dense map. S-SLAM [13] and Planar-SLAM [14] based on characteristic lines and surfaces can handle low-texture, structured indoor scenes. Again, the maps created by SLAM mentioned above do not have the capability to be used for navigation.

In this paper, we exploit the ability of RGB-D SLAM to output depth maps as well as camera poses in real time, combined with algorithms for navigation grid map building of 2D LiDAR SLAM, to design a robust RGB-D SLAM processing system specifically for building maps of indoor environments. We first obtain the depth map and the camera pose corresponding to the image information through RGB-D SLAM, then convert the depth map into laser information, and pass the camera pose as the predicted pose to SLAM to obtain the indoor two-dimensional grid map. The grid map is different from [1][2][3]. We pass the camera pose obtained by RGB-D SLAM as the predicted pose into SLAM, which is a more robust prediction. At the same time, we use the loop closure detection provided by itself to improve the real-time performance of the algorithm. In terms of equipment hardware, the system we propose can run in real-time on depth-sensing equipment, reducing the size and cost of the equipment. We have filled the functional gap of sparse RGB-D SLAM in the field of navigation by significantly reducing the computational effort. At last, we evaluated the mapping performance of the two-dimensional grid map obtained by the system using different RGB-D SLAM in the ICL dataset and showed the stability of our system in different RGB-D systems. With the improvement of RGB-D SLAM, there are better map building results.

2. Related Work

There are many related SLAM algorithms that can be used for indoor navigation, which can be roughly divided into two categories, visual SLAM and lidar SLAM. We will mainly summarize the RGB-D SLAM systems and the two-dimensional laser algorithm used in laser SLAM methods, respectively.

2.1. Two-dimensional laser SLAM

Two-dimensional laser SLAM is relatively mature in theory and practice and has been widely used in scientific research and industrial fields. Gmapping is based on the Fast-SLAM [15] scheme and is equipped with a two-dimensional algorithm based on particle filtering of the lidar sensor. The core idea of particle filtering is to randomly sample and estimate the map through selective resampling of particles [16]. Its main contribution is to improve the proposal distribution and selective resampling. However, due to the lack of closed-loop detection, it can only be used in simple, Maintain reliable accuracy in low-feature indoor environments. [17] proposed the corrective gradient refinement algorithm, which is a new method to improve the positioning based on the particle filter, which extends the traditional particle filter algorithm and is more universal. [18] proposed an improved algorithm based on information fusion, combining the odometer and inertial measurement unit, using Kalman filter for information fusion, assisting the robot in mapping, and improving the robot's positioning and mapping performance in degraded environments. Cartographer [3] based on graph optimization is different from [2] whose back-end adopts a closed-loop detection link based on branch and bound method. Cartographer can eliminate errors in robot motion and map large-area scenes. However, its optimization is more computationally intensive, and its real-time performance cannot be guaranteed.

2.2. RGB-D SLAM

Visual SLAM can be divided into sparse and dense maps according to the characteristics of mapping. Semi-dense and dense map can be directly used for indoor navigation maps, while sparse map cannot provide enough information for robots to navigate. Kinect Fusion [19] achieves real-time operation on the GPU. The system uses voxel grids to build maps, and does not restrict the cumulative error generated by the motion, so the application range is small, and when the environment is mainly composed of parallel planes, ICP Will fail. [20] proposed the Kintinuous system and added loop detection to eliminate accumulated errors, which improved the space expansion on the basis of [19]. Different from Kintinou, ElasticFusion [21] uses a direct optimization method for map points in order to improve accuracy, and densely reconstructs and repositions the three-dimensional environment through the surfel model, making full use of color and depth information, but the scope of the map is However, it is only suitable for reconstruction of room-sized scenes. [11] improved on the basis of ORB-SLAM [22], adding two modes of binocular SLAM and RGB-D SLAM. [11] adopts monocular and binocular beam adjustment optimization (BA), which improves the accuracy but can only create relatively simple 3D point clouds. Planer-SLAM [14] can handle low-structure, textured indoor scenes, and generate indoor three-dimensional plane environments from sparse point clouds. Dense RGBD-SLAM reconstruction can achieve good pose estimation and high-quality scene representation, but this requires high computational cost and complex equipment, such as high-performance GPU, which limits the platform used in different devices.

3. System overview

The schematic diagram of our proposed system is shown in Figure 1. Our work is inspired by RGB-D SLAM that generates sparse point cloud maps, such as ORB-SLAM 2 [11]. The above mentioned RGB-D SLAM can operate in an indoor environment and is also robust to strenuous exercise. However, the point cloud map generated by this system is sparse and cannot be directly used for indoor navigation. We can use the depth information in RGB-D SLAM, convert it into lidar data (See next subsection), and add its output pose information to build an indoor two-dimensional grid map. Our system first uses the pose provided by RGB-D SLAM as the initial pose, and then predicts the next new pose based on scan matching. These two poses obtain a new estimated pose through pose fusion. Insert the converted lidar data from the corresponding depth map into the final pose. After inserting a certain number of frames of lidar information, a submap is generated. We refer to the processing algorithm in [3], and on this basis, combined with the loop detection that comes with RGB-D SLAM, which reduces the cumulative error of previous submaps and improves the accuracy of mapping

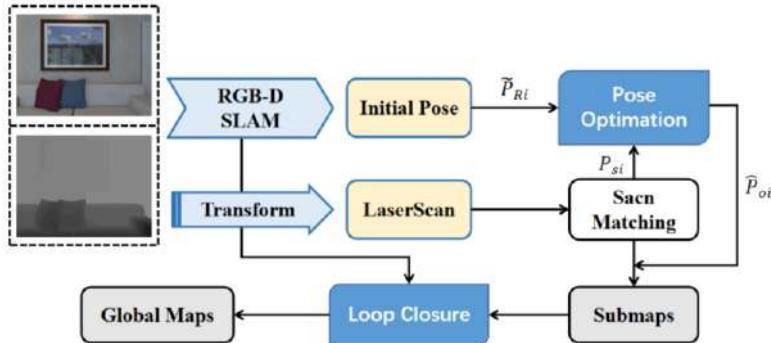


Figure 1. Main system pipeline

3.1. Depth map converted to laser scan

We first convert each pixel in the depth image (commonly the pixel in the middle of the image) into laser data. Refer to the pinhole camera model, the conversion matrix from a point $M(x, y, z)$ in the world coordinate system to a pixel point $m(u, v)$ in the camera coordinate system can be derived from a camera internal parameter matrix and an external parameter matrix, as shown in Equation 1.

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} [R \ T] = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (1)$$

where u, v are any coordinate points under the pixel coordinate system, u_0, v_0 are the central coordinates of the image. x, y, z represent the 3D coordinate points under the world coordinate system. z_c represents the z-axis value of the camera coordinates, the distance from the target to the camera. R, T are denoted as the unit matrices of 3×3 rotation and 3×1 translation in the external reference matrix respectively. Since the ICL

simulation set and the camera parameters it provides are used, distortions can be disregarded. We then project the point cloud data obtained above according to Equation 3 based on the pose information obtained from RGB-D SLAM, thus converting a point cloud derived from non-horizontal depth map information into a horizontal frame of point cloud data. We finally obtain the coordinates of the laser point cloud as follow:

$$\begin{cases} x = q_{pose} (z_C \cdot (u - u_0) \cdot dx / f) \cdot q_{pose}^{-1} \\ y = q_{pose} (z_C \cdot (v - v_0) \cdot dy / f) \cdot q_{pose}^{-1} \\ z = q_{pose} \cdot z_C \cdot q_{pose}^{-1} \end{cases} \quad (2)$$

where q_{pose} represents the pose information obtained by RGB-D SLAM.

3.2. Pose Optimization

Our system undergoes a pose optimization to process the initial pose from RGB-D SLAM and obtain the final pose. This optimization step starts with the calculation of a new pose by scan matching between the current frame of the LIDAR and the submaps, then the two poses are fused by Kalman filtering to obtain a more robust pose:

$$\hat{P}_{oi} = \tilde{P}_{Ri} + K(P_{Si} - C_i \tilde{P}_{Ri}) \quad (3)$$

where \hat{P}_{oi} represents the fused poses, \tilde{P}_{Ri} represents the predicted poses obtained from RGB-D SLAM, P_{Si} represents the observed poses obtained from the scan matching, Kalman gain K depends on the noise of the sensor, and the higher the camera noise, the higher the value of K .

Referring to the scan matcher in [3], scan matching works on the principle of finding the optimal probability value of the scanned points in a grid-based submap. In processing laser data and submap matching, a violent search matching algorithm is used to match the current laser frame traversing the historical grid map to obtain the position with the highest correlation, i.e., the position with the highest confidence. We subjected the grid map to bicubic interpolation and then converted the laser frame to map correlation matching to the least squares problem as follows:

$$\operatorname{argmin}_f \sum_{k=1}^K \left(1 - M_s(T_f h_k) \right)^2 \quad (4)$$

where h_k is the laser frame transformed into the submap by T_f and M_s is the probability-valued bicubic interpolation smoothing filter in the submap. The mathematical optimization of this smoothing function is usually higher resolution than the grid and has better accuracy. However, because we use local optimization, a good initial pose is required that using the pose given by RGB-D SLAM will be a good choice.

3.3. Loop Closure Detection

The map created by SLAM is divided into two main parts. The first part is a local landmark map called a submap, which is made up of a certain number of laser data. The second part is a global map built from a combination of accumulated submaps. The map

building process as described above accumulates errors, which are small for a few tens of consecutive laser data, but not negligible when building larger scenes or rotating scans of an interior. Laser SLAM requires a lot of computation to do loop closure in large map.

The approach taken in this paper is that the global map is corrected when loop closure is identified during the RGB-D SLAM, optimizing the pose of all scans and submaps and reducing the accumulated error in the global map. We directly use the loop closure information provided by RGB-D SLAM, making full use of that excellent loop closure algorithm to optimize the global pose. The global map optimization problem can be formulated as a non-linear least squares problem, using Ceres [23] to calculate problem as follow:

$$\underset{\Xi^m, \Xi^s}{\operatorname{argmin}} \frac{1}{2} \sum_{ij} \rho(E^2(\xi_i^m, \xi_j^s; \Sigma_{ij}, \xi_{ij})) \quad (5)$$

where the submap poses are $\Xi^m = \{\xi_i^m\}_{i=1, \dots, m}$ and the poses are $\Xi^s = \{\xi_j^s\}_{j=1, \dots, n}$ for each frame of laser data are optimized in the world coordinate system. These optimized submaps poses and Scan poses give constraints, which are expressed in terms of the poses ξ_{ij} and the covariance matrix Σ_{ij} . The residuals E are calculated as follows:

$$\begin{aligned} E^2(\xi_i^m, \xi_j^s; \Sigma_{ij}, \xi_{ij}) &= e(\xi_i^m, \xi_j^s; \xi_{ij})^T \Sigma_{ij}^{-1} e(\xi_i^m, \xi_j^s; \xi_{ij}), \\ e(\xi_i^m, \xi_j^s; \xi_{ij}) &= \xi_{ij} - \begin{pmatrix} R_{\xi_i^m}^{-1} (t_{\xi_i^m} - t_{\xi_j^s}) \\ \xi_{i;\theta}^m - \xi_{j;\theta}^s \end{pmatrix}. \end{aligned} \quad (6)$$

where the loss function, ρ (e.g. Huber loss), can be used to reduce the impact of outliers, which may occur in similar environments.

4. Experiment Results

In order to evaluate the performance of our system, in this section we will validate our proposed system using different RGB-D SLAM in two datasets. The first test uses the Living Room dataset, and the second test uses the Office Room dataset. There are open public datasets developed by Imperial College London [24]. 2D grid maps were generated on the ICL-NUIM dataset using different RGB-D SLAM. This paper also tests the performance of the system through trajectory comparison experiments. As shown in Figure 2, the RGB-D SLAM generates corresponding pose information based on the depth image and color image provided by ICL-NUIM, and the system uses the corresponding depth map from the dataset, converts it into a laser data and generates a 2D planar grid map. All experiments were conducted using an Intel Core i5-5200 (with @2.20GHz) and without any use of GPU.

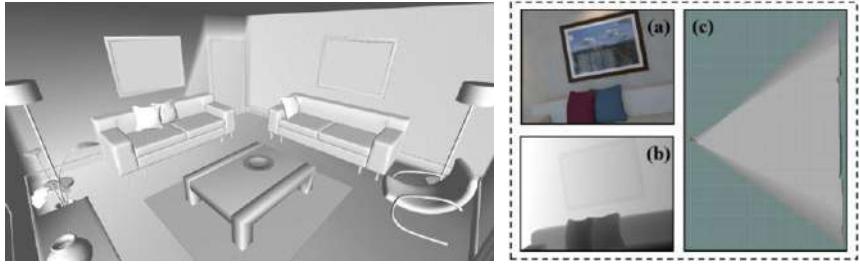


Figure 2. The image on the left shows the interior of the Living Room scene from the ICL dataset (the colors have been removed to highlight the geometry). (a) is a color frame of the Living Room. (b) is the depth frame of the Living Room at the same moment. (c) is the 2D grid map being generated.

4.1. Two-dimensional Grid Map

We use the ICL-NUIM Living Room dataset and the Office Room through different RGB-D SLAM respectively. ORB-SLAM2, Planar-SLAM and Planar-SLAM (with HM) were used for the experiments, and the ground truth data from the ICL-NUIM dataset were also used to generate 2D grid maps and trajectory maps as a reference. Figure 3 shows the experimental results of our system.

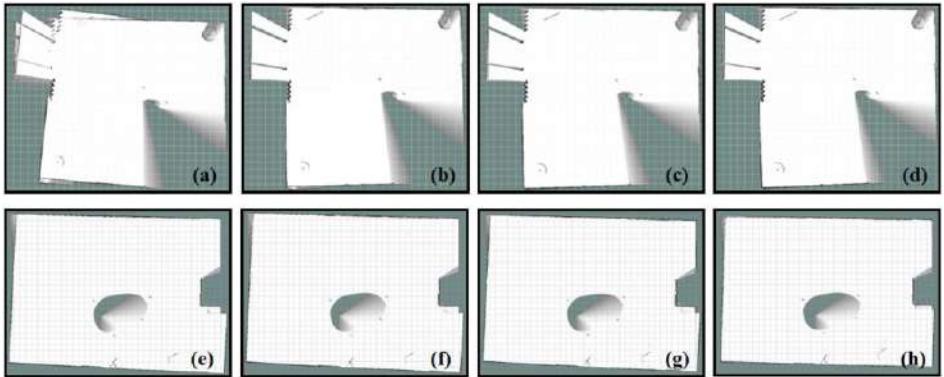


Figure 3. Results obtained using different RGB-D SLAM in the ICL dataset: the first row is Living Room: (a) ORB SLAM2 (b) Planar-SLAM (c) Planar-SLAM with Manhattan (d) ground truth; the second row is Office Room: (e) ORB SLAM2 (f) Planar-SLAM (g) Planar-SLAM with Manhattan (h) ground truth.

As shown in Figure 3, we see that with fewer feature points in the image data, the ORB-SLAM2 outputs less accurate poses resulting in fewer effective maps being built than Planar-SLAM or Planar-SLAM (with HM), which can improve the accuracy of the poses by detecting line and surface features. In the next section, we explain quantitatively by using the experimental data from the trajectory comparison in next section.

4.2. Trajectory curves

We obtained eight different trajectory curves from three different RGB-D SLAM and ground truth on the two datasets, as illustrated in Figure 4. We also calculated the average Euclidean distance of other trajectories compared with the ground truth to judge the similarity, and the results are shown in Table 1.

Table 1. Average Euclidean distances of trajectories obtained with different RGB-D SLAM

RGB-D SLAM	Living Room	Office Room
ORB-SLAM 2	0.407	0.454(m)
Planar SLAM	0.309	0.364(m)
Planar SLAM with MH	0.288	0.245(m)

As can be seen from the table, as the performance of the RGB-D SLAM positioning performance gets better, the closer the trajectories obtained are to the true values and the more similar the resulting maps are to the ground truth.

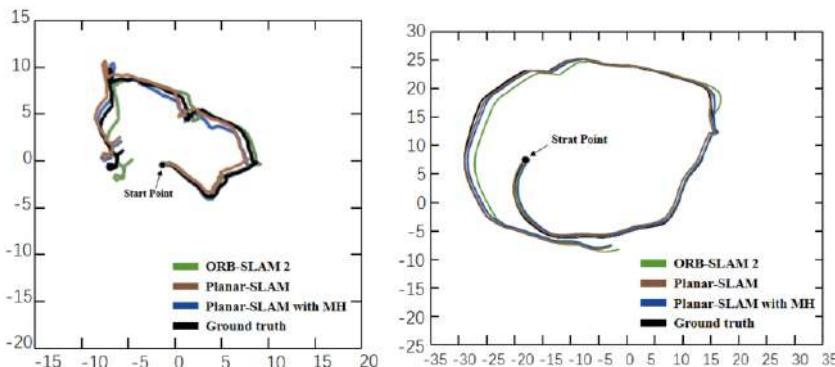


Figure 4. Experimental paths (unit meter) in Living Room (left) and Office Room (right) using different RGB-D SLAM: ORB-SLAM2 (green), Planar-SLAM (brown), Planar-SLAM with Manhattan (blue), ground truth (black).

5. Conclusion

We present a method for rapidly generating indoor 2D grid navigation maps by combining RGB-D SLAM and laser SLAM systems. The robust localization algorithm of RGB-D SLAM is utilized to determine the camera's poses and loops, and a depth map to laser data method is used to enable the depth map information to be used for laser SLAM, achieving advanced results by fusing pose. The lightness of the sensor and the fact that the algorithm does not require the use of large computational performance allows us to use the system for a robot with a depth sensor, using an IPC for indoor navigation of the robot. As a next step, we hope to make full use of the 3D information provided by the depth camera to generate indoor 3D navigation maps, suitable for complex indoor environments.

References

- [1] Grisetti G , Stachniss C , Burgard W . Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters[J]. IEEE Transactions on Robotics, 2007, 23:p.34-46.
- [2] Kohlbrecher S, Von Stryk O, Meyer J, et al. A flexible and scalable SLAM system with full 3D motion estimation[C]//2011 IEEE international symposium on safety, security, and rescue robotics. IEEE, 2011: 155-160.
- [3] Hess W, Kohler D, Rapp H, et al. Real-time loop closure in 2D LIDAR SLAM[C]//2016 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2016: 1271-1278.

- [4] Zhang J, Singh S. LOAM: Lidar Odometry and Mapping in Real-time[C]//Robotics: Science and Systems. 2014, 2(9).
- [5] Zhang J, Singh S. Visual-lidar odometry and mapping: Low-drift, robust, and fast[C]//2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2015: 2174-2181.
- [6] Shan T, Englot B. Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018: 4758-4765.
- [7] Lin J, Zhang F. Loam livox: A fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 3126-3131.
- [8] ElGhor H E C, Roussel D, Ababsa F, et al. 3D Planar RGB-D SLAM System[C]//International Conference on Advanced Concepts for Intelligent Vision Systems. Springer, Cham, 2016: 486-497.
- [9] Zhang, Shishun, Longyu Zheng, and Wenbing Tao. "Survey and Evaluation of RGB-D SLAM." IEEE Access 9 (2021): 21367-21387.
- [10] Davison A J, Reid I D, Molton N D, et al. MonoSLAM: Real-time single camera SLAM[J]. IEEE transactions on pattern analysis and machine intelligence, 2007, 29(6): 1052-1067.
- [11] Mur-Artal R, Tardós J D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras[J]. IEEE transactions on robotics, 2017, 33(5): 1255-1262.
- [12] Kerl C, Sturm J, Cremers D. Robust odometry estimation for RGB-D cameras[C]//2013 IEEE international conference on robotics and automation. IEEE, 2013: 3748-3754.
- [13] Li Y, Brasch N, Wang Y, et al. Structure-slam: Low-drift monocular slam in indoor environments[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 6583-6590.
- [14] Li Y, Yunus R, Brasch N, et al. RGB-D SLAM with Structural Regularities[J]. arXiv preprint arXiv:2010.07997, 2020.
- [15] Stentz A , Fox D , Montemerlo M , et al. FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem with Unknown Data Association[J]. Proc.aaai Natl.conf.on Artificial Intelligence Edmonton Alberta, 2002, 50(2):240-248.
- [16] Konolige K, Grisetti G, Kümmerle R, et al. Efficient sparse pose adjustment for 2D mapping[C]//2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2010: 22-29.
- [17] Wongsuwan K, Sukvichai K. Generalizing corrective gradient refinement in RBPF for occupancy grid LIDAR SLAM[C]//2017 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2017: 495-500.
- [18] Yu N, Zhang B. An improved hector SLAM algorithm based on information fusion for mobile robot[C]//2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS). IEEE, 2018: 279-284.
- [19] Newcombe R A, Izadi S, Hilliges O, et al. Kinectfusion: Real-time dense surface mapping and tracking[C]//2011 10th IEEE international symposium on mixed and augmented reality. IEEE, 2011: 127-136.
- [20] Whelan T, Johannsson H, Kaess M, et al. Robust real-time visual odometry for dense RGB-D mapping[C]//2013 IEEE International Conference on Robotics and Automation. IEEE, 2013: 5724-5731.
- [21] Whelan T, Salas-Moreno R F, Glocker B, et al. ElasticFusion: Real-time dense SLAM and light source estimation[J]. The International Journal of Robotics Research, 2016, 35(14): 1697-1716.
- [22] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE transactions on robotics, 2015, 31(5): 1147-1163.
- [23] S. Agarwal, K. Mierle, and Others, "Ceres solver," <http://ceres-solver.org>.
- [24] Handa A, Whelan T, McDonald J, et al. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM[C]//2014 IEEE international conference on Robotics and automation (ICRA). IEEE, 2014: 1524-1531.

Flat OFC Generation Based on DPMZM Cascaded Dual-Parallel PolM with Frequency Multiplication Circuit

Xin WANG, Sijia LIU, Ximing WANG, Yingxi MIAO, Caili GONG, Yongfeng WEI¹

*Department of Electronic Information Engineering, Inner Mongolia University,
Hohhot, 010021, China*

Abstract. An innovative scheme to generate the high-quality OFC based on DPMZM cascaded dual-parallel PolM with frequency multiplication circuit is proposed and demonstrated. In the scheme, 5 comb lines are generated in the first-stage generator, and the comb line is expanded to 11 times in the second-stage generator. The theoretical model of the overall scheme is established and analyzed. In this scheme, 66-line OFC is generated and the flatness is 0.73 dB, the side mode suppression ratio (SMSR) is 14.19 dB, and the optical signal noise ratio (OSNR) is about 29 dB.

Keywords. OFC generation, DPMZM, PolM, frequency multiplication circuit

1. Introduction

Microwave photonics has many unique superiorities, for instance the large bandwidth, low transmission loss, multiplexable and resistance to electromagnetic interference (EMI), and will be applied widely in radar technology, which can improve the performance of radar.

In microwave photonic radar system, high-quality multi-wavelength light sources are required, and the high-quality optical frequency comb (OFC) with high flatness, a large number of comb lines and the tunable frequency interval are considered as a research hotspot, and can be applied in WDM system [1], photonic filtering [2] and optical millimeter wave generating [3], etc. In the past decade, various schemes have been proposed for generating OFC, for instance the optical nonlinear effect method [4], the mode-locked laser method [5], the photoelectric oscillator method [6], and the external modulator method [7-10]. The method of generating OFC based on the external modulator has superiorities of stability, flexibility and adjustment, which has attracted great attention. In 2013, a novel approach to generating an OFC by applying a polarization modulator (PolM) and a polarizer was proposed which can generate 25-line OFC, and the flatness of OFC is 1 dB [8]. In 2014, a low-cost and simple OFC generation scheme was proposed by applying a dual-parallel Mach-Zehnder modulator (DPMZM) and generate 5-line and 7-line OFC with the flatness of 0.1 dB and 1 dB, respectively [9].

¹ Corresponding author: Yongfeng Wei, Department of Electronic Information Engineering, Inner Mongolia University, Hohhot, 010021, China ; E-mail: wyfly@bupt.edu.cn .

In 2015, a new scheme to generate OFC using one MZM and DPMZM which generate 9-line OFC, and the flatness is 0.8 dB [10].

In this scheme, a new scheme to generate the high-quality OFC is proposed and demonstrated. The result shows that 66-line OFC is generated, the flatness is 0.73 dB and SMSR is 14.19dB.

2. Scheme principle and analysis

The schematic diagram of flat OFC generation based on DPMZM cascaded dual-parallel PolM with frequency multiplication circuit is shown in Figure 1.

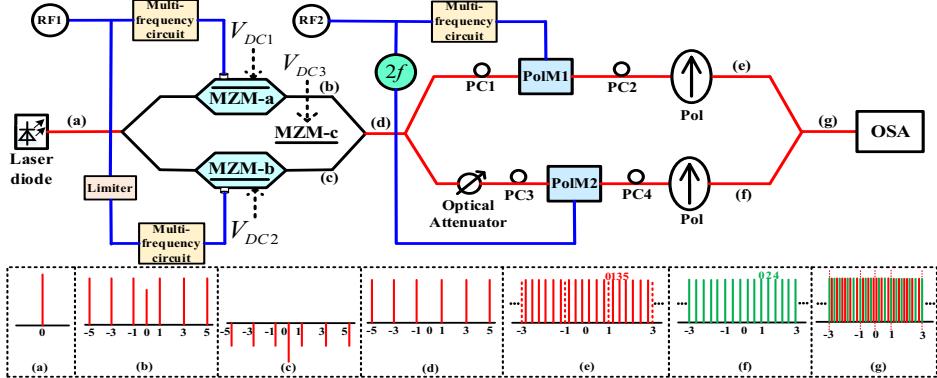


Figure 1. Schematic diagram of the flat OFC generation based on DPMZM cascaded dual-parallel PolM with frequency multiplication circuit.

Assuming the light wave launched from the laser diode can be expressed as $E_{in}(t) = E_c e^{i\omega_c t}$. On the upper branch, the RF1 is injected into the frequency multiplication circuit. The schematic diagram of frequency multiplication circuit is shown in Figure 2. On the lower branch, the RF1 signal is limited by a limiter, and then injected into the frequency multiplication circuit.

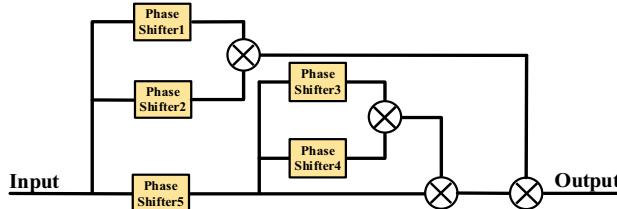


Figure 2. Schematic diagram of frequency multiplication circuit.

The phase shifts of electric frequency shifters are $\theta_{PS1} = \pi/3$, $\theta_{PS2} = -\pi/3$, $\theta_{PS3} = \pi/3$, $\theta_{PS4} = -\pi/3$ and $\theta_{PS5} = \pi/2$ respectively. The RF1 signal is $V_{RF1}(t) = V_{RF1} \cos(\omega_l t)$, after frequency multiplication circuit, the output is:

$$\begin{aligned}
 V_1(t) &= V_{RF1}^5 \sin(\omega_l t + \frac{\pi}{3}) \cdot \sin(\omega_l t - \frac{\pi}{3}) \cdot \sin(\omega_l t - \frac{\pi}{3}) \cdot \cos(\omega_l t + \frac{\pi}{3}) \cdot \cos(\omega_l t) \\
 &= \frac{V_{RF1}^5}{16} [\cos(\omega_l t) + \cos(3\omega_l t) + \cos(5\omega_l t)]
 \end{aligned} \tag{1}$$

The output spectrogram is shown in Figure 3, assuming that $\omega=4\text{GHz}$.

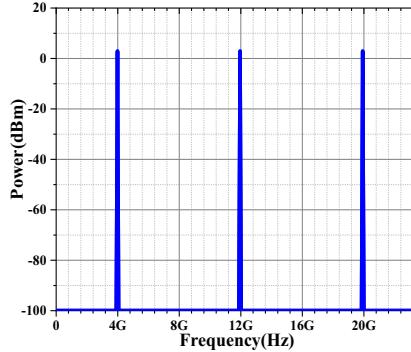


Figure 3. Electrical spectrum of the output signal.

The drive signals output by the frequency multiplication circuit are injected into DPMZM for modulation. MZM-a and MZM-b are biased at minimum transmission point, and the MZM-c is applied to introduce phase difference of π on the upper and lower branch, therefore $V_{DC1}=V_{DC2}=V_{DC3}=4\text{V}$. When the modulation index is small enough, the influence of the high-order sidebands can be ignored, and ± 1 order sidebands are preserved. The output expression of the upper branch is:

$$E_{out-a} = j \frac{1}{4} E_{in}(t) J_1(V_{RF1}) (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \quad (2)$$

The output expression of the lower branch is:

$$E_{out-b} = -\frac{1}{4} j E_{in}(t) J_1(V_{RF1^*}) (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \quad (3)$$

Combine the upper and lower branch of the DPMZM, the expression of the output signal can be defined as:

$$E_{out-DPMZM}(t) = j \frac{1}{4} E_{in}(t) [J_1(V_{RF1}) - J_1(V_{RF1^*})] \cdot (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \quad (4)$$

In the actual simulation, the optical carrier cannot be completely suppressed. In order to suppress the optical carrier, it is necessary to adjust the limiter to change the amplitude of the lower-branch optical signal, so that $J_0(V_{RF1}) \approx J_0(V_{RF1^*})$, and realize the suppression of the optical carrier without affecting other harmonic components. The output spectrum of DPMZM is shown in Figure 4.

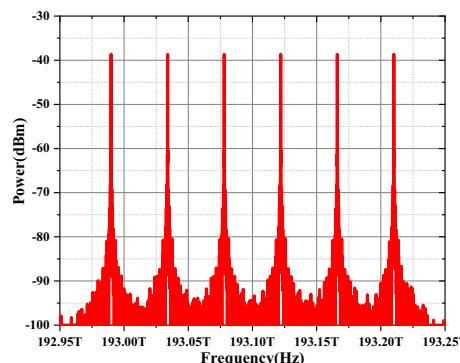


Figure 4. The optical spectrum at the output of DPMZM.

On the upper branch of dual-parallel PolM, by adjusting PC1, the optical signal is injected into PolM at an angle of 45° with the main axis of PolM. Assuming that the optical signal is $E_{in}(t)$, which is injected into PolM, and the expression of RF is $V(t) = V \cos \omega t$. Set the angle between the Pol and the main axis of PolM as $\theta = 45^\circ$, the output expression after Pol is:

$$\begin{aligned} E_{out_pol}(t) &= \frac{E_{in}(t)}{2} (\cos \theta \cdot e^{jm \cos \omega t} + \sin \theta \cdot e^{-jm \cos \omega t + j\varphi}) \\ &= \frac{E_{in}(t)}{2} \left\{ \sum_{n=-\infty}^{\infty} [1 - (-1)^n e^{j\varphi}] J_n(m) e^{jn\omega t} \right\} \end{aligned} \quad (5)$$

Where φ is the phase difference introduced by the upper and lower phase modulators, $m = \pi V / V_\pi$ is the modulation index of PolM, and $J_n(m)$ is the n th-order Bessel function. By adjust PC2, we can make the phase difference $\varphi = \pi$. If the value of m is small, high-order Bessel function should be ignored, only the +1 order and -1 order harmonics are preserved. The output expression of the signal can be expressed as:

$$E_{out_pol}(t) = j \frac{E_{in}(t)}{2} J_1(m) (e^{j\omega t} + e^{-j\omega t}) \quad (6)$$

When RF signal is $V_{RF2}(t) = V_{RF2} \cos(\omega_2 t)$, which is injected into the frequency multiplication circuit. The output signal of upper branch can be expressed as:

$$E_{upper}(t) = j \frac{E_{in}(t)}{2} J_1 \left(\frac{V_{RF2}}{16} \right) (e^{j\omega_2 t} + e^{-j\omega_2 t} + e^{j3\omega_2 t} + e^{-j3\omega_2 t} + e^{j5\omega_2 t} + e^{-j5\omega_2 t}) \quad (7)$$

When RF signal is $V_{RF2}^*(t) = V_{RF2} \cos(2\omega_2 t)$, the output is:

$$\begin{aligned} E_{lower}(t) &= \frac{E_{in}(t)}{2} [J_0(V_{RF2}) (1 + e^{j\varphi}) \\ &\quad + J_1(V_{RF2}) (1 + e^{j\varphi}) e^{j2\omega_2 t} + J_{-1}(V_{RF2}) (1 + e^{j\varphi}) e^{-j2\omega_2 t} \\ &\quad + J_2(V_{RF2}) (1 + e^{j\varphi}) e^{j4\omega_2 t} + J_{-2}(V_{RF2}) (1 + e^{j\varphi}) e^{-j4\omega_2 t}] \end{aligned} \quad (8)$$

If $J_0(V_{RF2}) = J_2(V_{RF2})$ and $\cos \varphi = [J_1^2(V_{RF2}) - J_0^2(V_{RF2})] / [J_1^2(V_{RF2}) + J_0^2(V_{RF2})]$, by adjusting the φ , we can make $J_0(V_{RF2}) = J_{\pm 1}(V_{RF2}) = J_{\pm 2}(V_{RF2})$. When a beam of optical carrier is injected into PolM, the upper and lower output signals are shown in Figure 5.

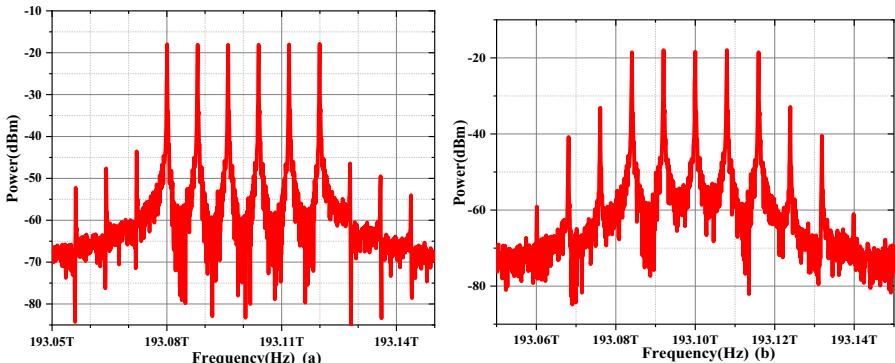


Figure 5. Optical spectrum at the output of upper (a) and lower(b) branch of dual-parallel PolM.

On the upper branch of the dual-parallel PolM, each optical comb generated by DPMZM will be injected into PolM as an optical carrier for modulation. The output expression of the PolM1 is:

$$\begin{aligned}
 E_{PolM1} &= E_{out-DPMZM}(t) * \frac{1}{2} J_1 \left(\frac{V_{RF2}}{16} \right)^5 \cdot [e^{j\omega_2 t} + e^{-j\omega_2 t} + e^{j3\omega_2 t} + e^{-j3\omega_2 t} + e^{j5\omega_2 t} + e^{-j5\omega_2 t}] \\
 &= j \frac{1}{8} E_{in}(t) J_1 \left(\frac{V_{RF2}}{16} \right)^5 [J_1(V_{RF1}) - J_1(V_{RF1}^*)] \cdot \\
 &\quad (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \\
 &\quad (e^{j\omega_2 t} + e^{-j\omega_2 t} + e^{j3\omega_2 t} + e^{-j3\omega_2 t} + e^{j5\omega_2 t} + e^{-j5\omega_2 t})
 \end{aligned} \tag{9}$$

On the lower branch of dual-parallel PolM, the OFC generated by DPMZM will be injected into an optical attenuator, by optical attenuator we can attenuate the input optical signal power. And then, each of optical comb line is injected into PolM as an optical carrier for modulation. The output expression of the PolM2 is:

$$\begin{aligned}
 E_{PolM2} &= E_{out-DPMZM}(t) * \frac{1}{2} J_0 \left(V_{RF2} \right) (1 + e^{j\varphi}) \cdot (1 + e^{j2\omega_2 t} + e^{-j2\omega_2 t} + e^{j4\omega_2 t} + e^{-j4\omega_2 t}) \\
 &= j \frac{1}{8} E_{in}(t) J_0 \left(V_{RF2} \right) (1 + e^{j\varphi}) [J_1(V_{RF1}) - J_1(V_{RF1}^*)] \cdot \\
 &\quad (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \\
 &\quad (1 + e^{j2\omega_2 t} + e^{-j2\omega_2 t} + e^{j4\omega_2 t} + e^{-j4\omega_2 t})
 \end{aligned} \tag{10}$$

According to Eq.(9) and Eq.(10), we can get 36 comb lines and 30 comb lines respectively. By coupling the upper and lower branch, we can obtain 66-line OFC, and the expression can be expressed as:

$$\begin{aligned}
 E_{out-DPPolM} &= j \frac{1}{8} E_{in}(t) J_1 \left(\frac{V_{RF2}}{16} \right)^5 [J_1(V_{RF1}) - J_1(V_{RF1}^*)] \cdot \\
 &\quad (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \\
 &\quad (e^{j\omega_2 t} + e^{-j\omega_2 t} + e^{j3\omega_2 t} + e^{-j3\omega_2 t} + e^{j5\omega_2 t} + e^{-j5\omega_2 t}) \cdot \\
 &\quad \{ j \frac{1}{8} E_{in}(t) J_0 \left(V_{RF2} \right) (1 + e^{j\varphi}) [J_1(V_{RF1}) - J_1(V_{RF1}^*)] \cdot \\
 &\quad (e^{j\omega_1 t} + e^{j3\omega_1 t} + e^{j5\omega_1 t} + e^{-j\omega_1 t} + e^{-j3\omega_1 t} + e^{-j5\omega_1 t}) \\
 &\quad (1 + e^{j2\omega_2 t} + e^{-j2\omega_2 t} + e^{j4\omega_2 t} + e^{-j4\omega_2 t}) \}
 \end{aligned} \tag{11}$$

In order to obtain an OFC with the same comb line interval, the frequencies of RF1 signal and RF2 signal need to satisfy the following relationship: $\omega_1 = 5.5\omega_2$.

3. Simulation results and analysis

To prove that the scheme is feasible, we demonstrate the scheme by the simulation. First of all, a laser diode (LD) generates a continuous light carrier and the frequency of the light carrier is 193.1 THz, the power is 10 dBm, and the line-width is 10 MHz, which is modulated by DPMZM. The half-wave voltage of the MZM-a and MZM-b are set to 4V, the bias voltage of the upper and lower sub-MZMs are $V_{DC1} = V_{DC2} = 4V$, and the bias voltage of MZM-c is $V_{DC3} = 4V$. The frequency of RF1 is 22 GHz, and the amplitude of RF1 is 1.0V. By adjusting the limiter, the amplitude of the input RF signal of the MZM-b can be adjusted. In this scheme, adjusting the limiter to make the amplitude of the RF signal to 0.5V. The frequency of RF2 is 4GHz, and the amplitude is 1.0V. On the upper branch of dual-parallel PolM, the frequency offset of the upper branch of PolM1

is 90 deg, and the lower branch is -90 deg. We set the rotation angle of PC1 to 45 deg, so that making the input signal and the main axis of PolM at the angle of 45 deg. Change the PC2 and setting its rotation angle to 45 deg, and the signal is injected into a Pol with a device angle of 45 deg. On the lower branch of dual-parallel PolM, the frequency of RF drive signal is $2f_{RF2} = 8\text{GHz}$, the frequency offset of the upper branch of PolM2 is 210 deg, and the lower branch is -210 deg. Similarly, the parameters of the PC3, PC4 and Pol are the same as the upper branch of dual-parallel PolM.

When the value of above parameters remain unchanged, as shown in Figure 6(a), the interval of the comb lines is 4GHz, 66-line OFC is generated, the flatness is 0.73dB, and the SMSR is 14.19 dB, the OSNR is about 29 dB. When the frequency of the RF2 is 0.5 GHz, as shown in Figure 6(b), the interval of the comb lines is about 0.5 GHz, 66 line OFC is generated and the flatness is 1.16 dB, and the SMSR is 13.69 dB, the OSNR is about 19.5 dB.

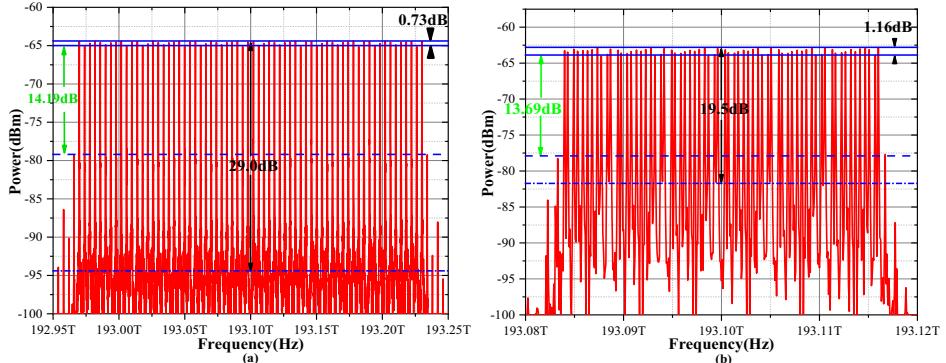
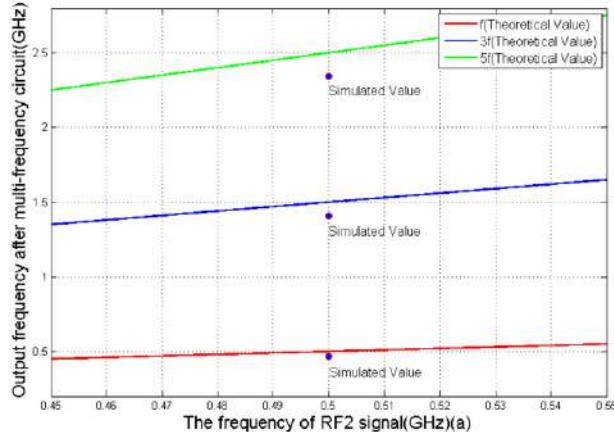


Figure 6. The generated optical comb with 66 comb lines when the interval is 4GHz (a) and 0.5GHz (b).

According to Fig.6, we can generate 66 comb lines. When the frequency of RF2 is 4 GHz, the generated OFC is relatively flat, and the interval of the comb line is uniform. When the frequency of RF2 is 0.5 GHz, the generated OFC is also relatively flat, however, the interval of the comb line is not uniform, which is caused by the frequency multiplication circuit. The Fig.7 shows the comparison between the theoretical value and the simulated value of the output signal after different RF2 signals are injected into the frequency multiplication circuit.



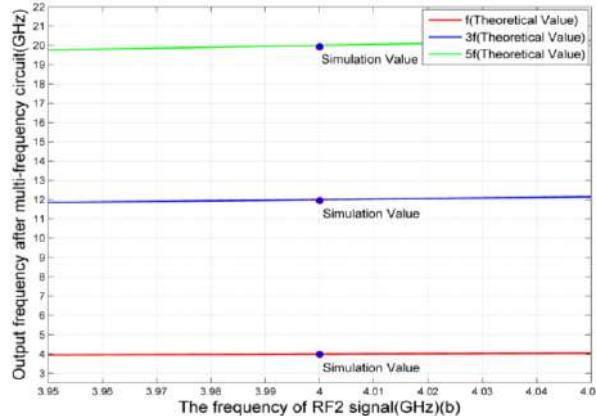


Figure 7. Comparison between the theoretical value and the simulated value of the output signal after 0.5GHz (a) and 4GHz (b) signals are injected into the frequency multiplication circuit.

As shown in Fig.7 (a), when the frequency of RF2 is 4 GHz, the simulated value of the output signal after frequency multiplication circuit is completely consistent with the theoretical value, and the output signal is 4 GHz, 12 GHz and 20 GHz. As shown in Fig.7 (b), when the frequency of RF2 is 0.5 GHz, the simulated value of the output signal after frequency multiplication circuit is not completely consistent with the theoretical value, and the output signal is 0.47 GHz, 1.4 GHz and 2.34 GHz, which causes the interval of the comb line is not so uniform.

4. Conclusion

In this scheme, an innovative scheme to generate the high-quality OFC based on DPMZM cascaded dual-parallel PolM with frequency multiplication circuit is proposed and demonstrated. In this scheme, when the interval of the comb lines is 4GHz, 66 optical comb lines with a flatness of 0.73dB is generated, and the SMSR is 14.19 dB, the OSNR is about 30dB. When the frequency of the RF2 is 0.5 GHz, the interval of the comb lines is about 0.5 GHz, 66 optical comb lines with a flatness of 1.16 dB is generated, and the SMSR is 13.69 dB, the optical signal noise ratio is about 21.14 dB. In our proposal, by adjusting the frequency of the RF, we can adjust the interval of OFC. The results show the considerable approach for generation OFC. Compared with other schemes, this scheme requires low-power RF signal, and the minimum interval of comb lines can reach to about 0.5 GHz.

Acknowledgment

This work was supported by the National Nature Science Foundation of China (No.62161034, No.61561037, No.61861034, No.61763034).

References

- [1] Yan, J., Xia, Z., Zhang, S., et al.: A flexible waveforms generator based on a single dual-parallel Mach-Zehnder Modulator. *J. Opt. Commun.* 334, 31–43 (2015).

- [2] W. Dai, W. Rao, H. Wang and H. Fu, "Microwave Photonic Filter by Using Recirculating Frequency Shifter to Generate Optical Frequency Comb," in *IEEE Photonics Journal*, vol. 13, no. 2, pp. 1-8, April 2021, Art no. 5500308, doi: 10.1109/JPHOT.2021.3069104.
- [3] Atsushi Kanno, G. K. M. Hasanuzzaman, Naokatsu Yamamoto, Stavros Iezekiel, "Optical frequency comb applied optoelectronic oscillator for millimeter-wave signal generation and its application," *Proc. SPIE 10634, Passive and Active Millimeter-Wave Imaging XXI*, 106340E (8 May 2018); <https://doi.org/10.1117/12.2304821>.
- [4] E. A. Anashkina and A. V. Andrianov, "Kerr-Raman Optical Frequency Combs in Silica Microsphere Pumped Near Zero Dispersion Wavelength," in *IEEE Access*, vol. 9, pp. 6729-6734, 2021, doi: 10.1109/ACCESS.2021.3049183.
- [5] X. -Z. Li, M. Seghilani, L. R. Cortés and J. Azaña, "Sub-THz Optical Frequency Comb Generation by Efficient Repetition-Rate Multiplication of a 250-MHz Mode-locked Laser," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 2, pp. 1-10, March-April 2021, Art no. 7601310, doi: 10.1109/JSTQE.2020.3039280.
- [6] Zhengyang Xie, Shangyuan Li, Haozhe Yan, Xuedi Xiao, Xiaoxiao Xue, Xiaoping Zheng, Bingkun Zhou, "Tunable ultraflat optical frequency comb generator based on optoelectronic oscillator using dual-parallel Mach-Zehnder modulator," *Opt. Eng.* 56(6) 066115 (27 June 2017) <https://doi.org/10.1117/1.OE.56.6.066115>.
- [7] K. Qu, S. Zhao, X. Li, Z. Zhu, D. Liang and D. Liang, "Ultra-Flat and Broadband Optical Frequency Comb Generator via a Single Mach-Zehnder Modulator," in *IEEE Photonics Technology Letters*, vol. 29, no. 2, pp. 255-258, 15 Jan. 2017, doi: 10.1109/LPT.2016.2640276.
- [8] Cihai Chen, Chao He, Dan Zhu, Ronghui Guo, Fangzheng Zhang, and Shilong Pan, "Generation of a flat optical frequency comb based on a cascaded polarization modulator and phase modulator," *Opt. Lett.* 38, 3137-3139 (2013).
- [9] Qiang Wang, Li Huo, Yanfei Xing, and Bingkun Zhou, "Ultra-flat optical frequency comb generator using a single-driven dual-parallel Mach-Zehnder modulator," *Opt. Lett.* 39, 3050-3053 (2014).
- [10] Q. Chang, J. Gao and Y. Su, "Generation of optical comb frequency signal with high spectral flatness using two cascaded optical modulators," *2008 Asia Optical Fiber Communication & Optoelectronic Exposition & Conference*, 2008, pp. 1-3. Yan, J., Xia, Z., Zhang, S., et al.: A flexible waveforms generator based on a single dual-parallel Mach-Zehnder Modulator. *J. Opt. Commun.* 334, 31-43 (2015).

Algebra Based Human Skeleton Sequence Reduction and Action Recognition

Shibin XUAN^{a,b,1}, Kuan WANG^a, Lixia LIU^a, Chang LIU^a and Jiaxiang LI^a

^aSchool of Artificial Intelligence, Guangxi University for Nationalities, China

^bGuangxi Key Laboratory of Hybrid Computation and IC Design Analysis, China

Abstract. Skeleton-based human action recognition is a research hotspot in recent years, but most of the research focuses on the spatio-temporal feature extraction by convolutional neural network. In order to improve the correct recognition rate of these models, this paper proposes three strategies: using algebraic method to reduce redundant video frames, adding auxiliary edges into the joint adjacency graph to improve the skeleton graph structure, and adding some virtual classes to disperse the error recognition rate. Experimental results on NTU-RGB-D60, NTU-RGB-D120 and Kinetics Skeleton 400 databases show that the proposed strategy can effectively improve the accuracy of the original algorithm.

Keywords. Action recognition, redundant frames reduction, joint adjacency graph, virtual classes

1. Introduction

Human action recognition has always been one of the research hotspots in computer vision. In recent years, human action recognition has attracted a large number of researchers' interest because of its wide application fields, such as intelligent video surveillance, advanced human-computer interaction, virtual reality, taking care for the old man and the children, computer-assisted teaching and intelligent assisted driving, etc. In the development of human behavior recognition technology, researchers have put forward many methods, among which the typical representatives are: the action recognition method based on tracking technology [1,2], the action recognition method based on human column structure [3-5], action recognition method based on spatio-temporal descriptor [6-8], Recognition method based on image depth information [9,10], the action recognition method based on "word bags" [11,12] and the action recognition method based on human skeleton [13,14,15]. Because the human skeleton data has the advantages of not being affected by human body shape and background, and the decrease of the price of depth sensors used to acquire human skeleton data and the emergence of skeleton data generation algorithms, the research of human action recognition algorithm based on skeleton data become a research hotspot. The skeleton-based action recognition deep learning method has attracted a large number of researchers' attention.

At present, deep learning models of skeleton-based human action recognition mainly include: Recurrent Neural Networks (RNN)[16,17,18,19,20], Convolutional

¹ Corresponding Author, Shibin Xuan E-mail: sbinx@qq.com.

Neural Networks(CNN)[21,22,23], and Graph Convolutional Networks(GCN) [18, 24, 25,26], Semantics-Guided Neural Networks[27].

Although these methods have achieved great success, they mostly learn the relationship between nodes in a frame through convolutional neural network and the change relationship between frames on the time axis through convolutional neural network or recursive neural network. Due to the diverse characteristics of human movement, even if the same movement is made by different people in different time periods, there are also great differences. At present, most methods are to expand and rotate the position of the key node, but the speed difference of action is only solved by multi-size window, which is obviously only a trade-off. In fact, it can be seen from the case analysis that human movement is completed through the activities of limbs and trunk, and each movement of limbs has its starting point and ending point. The number of video frames between the starting point and ending point of limbs movement in the slow video segment are more than those in the fast one, which increases the uncertainty of the convolution calculation on the time axis. For this purpose, in this paper, we propose a method based on algebraic operation to reduce the number of frames between the starting frame and the ending frame, and try to reduce the inconsistency caused by the different speed of action. At the same time, in order to improve the relative relationship between nodes, some connections between nodes that are not physically adjacent are added to the skeleton graph. In addition, the number of target classes is increased to disperse the error rate. Since the proposed measures do not involve any network structure, they can be applied to any existing skeleton-based human action recognition models.

2. Introduction of related basic work

The earliest and most successful application of graph convolutional neural network in skeleton-based human behavior recognition is ST-GCN [15], where the skeleton data is represented as a spatio-temporal graph $G = (V, E)$, in which the vertices in the graph is correspond to the human joints, and the edge represents the human bones and the line between same joints in the consecutive frame. A video segment can be represented as $X \in R^{N \times T \times C}$, where N represents the number of joints in an frame image, T represents the number of frames in a video, and C represents the number of channels. A relations graph between joints in a single frame can be represented by an adjacency matrix $A \in \{0,1\}^{N \times N}$, I denotes the identity matrix. $A + I$ can be divided into three parts $A_j (j = 1,2,3)$ according to the three cases {root, centripetal, earth}, and satisfies $A + I = \sum_{j=1}^3 A_j$. In the case of single frame, spatial convolution calculation can be expressed as:

$$F_{out} = \sum_{j=1}^3 \Lambda_j^{-\frac{1}{2}} A_j \Lambda_j^{-\frac{1}{2}} F_{in} W_j \quad (1)$$

Where $\Lambda_j^{ii} = \sum_k (A_j^{ik}) + \alpha$, α is a regulating constant, in generally set as $\alpha = 0.001$.

Ke Cheng et al. [28] introduced the shift-CNNS idea into ST-GCN and proposed shift-GCN, which greatly reduced the model complexity. Shift-GCN includes Shift space graph convolution and Shift time graph convolution. Shift space graph convolution is divided into local Shift graph convolution and non-local Shift graph convolution. The neighborhood of node v is denoted as $B_v = \{B_v^1, B_v^2, \dots, B_v^n\}$, and the channel at node v is divided into $n + 1$ parts. Among them, the first part retains the

characteristics of node v , and the other n parts are gained from shifting $B_v^1, B_v^2, \dots, B_v^n$ respectively. If $F \in R^{N \times C}$ is supposed to be a human skeleton feature in a video frame, the corresponding shift features $\tilde{F} \in R^{N \times C}$ can be calculated according to the following formula:

$$\tilde{F} = F_{(v,:c)} \parallel F_{(B_v^1,c:2c)} \parallel F_{(B_v^2,2c:3c)} \parallel \dots \parallel F_{(B_v^n,nc:)} \quad (2)$$

Where $c = \left\lfloor \frac{C}{n+1} \right\rfloor$, \parallel represents channel-wisely cascade, and a learnable mask is used to realize non-local Shift graph convolution.

Ziyu Liu[29] et al. proposed loose graph convolution network and consistent spatio-temporal graph convolution operator to solve the unbiased long-domain connection relation under multi-scale operator and barrier-free capture of complex spatio-temporal dependence from spatio-temporal information flow. For this purpose, First of all, the k adjacency matrix $\tilde{A}_{(k)}$ can be defined as follows:

$$[\tilde{A}_{(k)}]_{i,j} = \begin{cases} 1 & \text{if } d(v_i, v_j) = k, \\ 1 & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

Then the output of the spatial convolution model can be expressed as:

$$X_t^{(l+1)} = \sigma(\sum_{k=0}^K \tilde{D}_{(k)}^{-\frac{1}{2}} \tilde{A}_{(k)} \tilde{D}_{(k)}^{-\frac{1}{2}} X_t^{(l)} W_{(k)}^{(l)}) \quad (4)$$

Let the size of the sliding window along the time axis be τ and the window expansion rate be d ,

$$\tilde{A}_{(\tau)} = \begin{bmatrix} \tilde{A} & \cdots & \tilde{A} \\ \vdots & \ddots & \vdots \\ \tilde{A} & \cdots & \tilde{A} \end{bmatrix} \in R^{\tau N \times \tau N}. \quad (5)$$

Similarly, we can define joint features $X_{(\tau,d)} \in R^{\tau N \times \tau N \times C}$.

The final MS-G3D model compute can be expressed as:

$$[X_{\tau}^{(l+1)}]_t = \sigma(\sum_{k=0}^K \tilde{D}_{(\tau,k)}^{-\frac{1}{2}} \tilde{A}_{(\tau,k)} \tilde{D}_{(\tau,k)}^{-\frac{1}{2}} [X_{\tau}^{(l)}]_t W_{(k)}^{(l)}) \quad (6)$$

3. Skeleton data reduction based on algebraic computation and category optimization

From the last section 2, their work is more about network structure improvement, but network structure improvement has certain limits. We look forward to improving the correct recognition rate without changing the network structure. In this part, we will introduce in detail how to use algebraic method to reduce the video frames on the time axis, analyze the changing characteristics of the relationship between joints in human activities and the theoretical basis of category optimization.

3.1. Framework frame data reduction based on algebraic calculation

Although the human skeleton data has greatly simplified the video data, because the human body motion curve described by the skeleton is still a high-dimensional curve, the motion difference makes the motion curve of the same action also have a great difference. In addition, the start and end frames of the obtained skeleton data are also inconsistent, and the number of frames expressing a complete action is also

inconsistent. All these make the activity recognition based on skeleton data have great uncertainty.

Human action is formed by the movement of human limbs and head, and the movement of limbs is the key factor of action recognition. No matter how complex the movements of the human body are, the movements of the limbs are made up of a combination of fragments of movement in a particular direction, and the characteristics of movement are mainly determined by the starting and ending points of this movement in a particular direction. The starting and ending video frames play a key role in motion recognition, and the video frames determining the activity category are referred to as key frames.

As the skeleton data of video sequence is a high-dimensional data stream, in general, it is difficult to directly calculate the corresponding key frame for such high-dimensional data. However, if we can establish a mapping from such high-dimensional data to one-dimensional data, and this mapping can retain the trend characteristics of the original human movement, we can easily obtain the corresponding video key frames from this one-dimensional data. After many experiments we found that there is a characteristic function that has this function. Here, we present a mapping function that can be used for fast and efficient key frame extraction.

Assuming that the number of joints in the skeleton is n , x_{ij} represents the j th joint node of i th frame, and the internal structure matrix of k th frame is expressed as follows:

$$R_k = (r_{st}^k)_{n \times n} \quad (7)$$

Where $r_{st}^k = \text{dist}(x_{ks}, x_{kt})$, $s, t = 1, \dots, n$, the function $\text{dist}(\cdot, \cdot)$ represents the distance between two points. Then the mapping function of $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ is defined as:

$$y_k = \text{smooth}(\text{std}(\text{eig}(R_k))). \quad (8)$$

Where the function $\text{eig}(\cdot)$ represents the eigenvalue vector of the matrix, $\text{std}(\cdot)$ denotes the standard variance of the vector, and $\text{smooth}(\cdot)$ refers to as the data smoothing process.

Let $S_{\max} = \text{maxid}([y_k])$, $S_{\min} = \text{minid}([y_k])$, $S_{\text{ext}} = \text{uion}(S_{\max}, S_{\min})$, where $\text{maxid}(\cdot)$ represents the indies of the maximum value, $\text{minid}(\cdot)$ denotes the indies of the minimum value, and $\text{uion}(\cdot)$ is the set union operation, so that set S_{ext} includes the indies set of the key frames in the corresponding video.

For example, we select a video in MSRAction 3D database and solve key frames according to the above calculation formula. The mapping function curve of the video segment in Figure 1 is shown in Figure 2, here Minimum frame numbers: 4, 16, 26, 37, Maximum frame numbers: 1, 12, 21, 27.

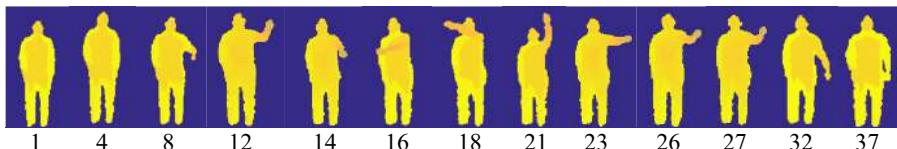


Figure 1. The depth information image of a video with extreme frames in MSRAction 3D

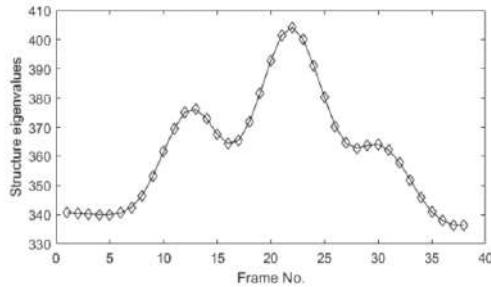


Figure 2 Mapping function of video segment in Fig. 1

It can be seen from Fig.1 that the mapping function defined in this way from skeleton data to motion trend basically meets the segmentation requirements of action units.

In order to further illustrate the rationality of the mapping function established above, we randomly generate a symmetric matrix with a size of 20×20 , multiply the values of the k th row and the k th column by a periodic change function such as $\sin(\pi x)$, and let $x = 0.1, 0.2, 0.3, \dots, 6$, then the mapping function curve is obtained as shown in Fig.3. Because when a joint of human body is moving, the distance between this node and other joints will change, the row and column values corresponding to this joint in the corresponding distance matrix will change accordingly. Since the human

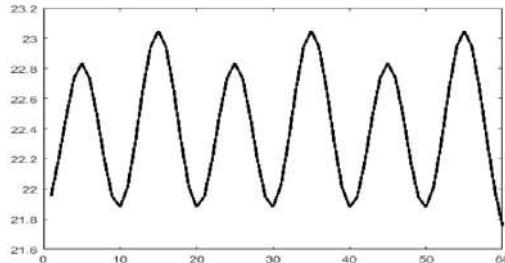


Figure 3 The mapping function curve of the random symmetric matrix is obtained by multiplying the 5th row and 5th column by the periodic function $\sin(\pi x)$.

body movement is in a specific direction at a certain time, it represents the increase or decrease of values in the same row and column in the structural matrix. Fig.3 shows that the structure mapping function defined can better reflect human body changes.

Advantages of the algorithm: fast calculation speed. Generally, the number of joints in a frame is 15, 20 or 25. The corresponding distance matrix does not exceed 25×25 dimensions. For such a low dimension, it takes very little time for the computer to calculate its eigenvalue and the corresponding variance, filtering and extreme value. We summarize the above calculation process into the following algorithm.

Algorithm 1 Video frame reduction

Input: original video sequence X , smoothing filter $\text{smooth}()$, extremum extraction function $\text{maxid}()$, $\text{minid}()$.

Output: reduced video sequence Y .

Step 1: Compute structure matrix R_k of k th frame by formula 7.

Step2: Compute feature value y_k of structure matrix R_k of k th frame by formula 8.

Step3: Compute extremum frame number set S_{ext} of X by functions $maxid()$, $minid()$, $uion()$.

Step4: Output the frames set of X with frame number in S_{ext} .

3.2. Skeleton graph

In most models, the vertices in the skeleton structure graph are taken as each joint and the edges as the physiological bone. In some new models, some new edges are added on the natural skeleton graph to express the importance of the position relations between these joints, so as to make up for the fact that the convolution operation of the natural skeleton graph cannot take into account these relations. For example, Yansong Tang [24] et al., in order to show the importance of the relative position relation between the joints of limbs for activity description, added the connection between the corresponding joints of limbs on the basis of the original natural connection, as shown in Fig.4.

Considering the importance of limbs in motion, in addition to the natural connection between nodes, this paper adds the connection edge between some limbs nodes and the central position of the human body to enhance the role of limbs in the representation of human posture. See Fig.5

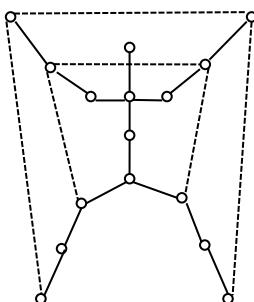


Figure 4. Solid lines represent bones and dotted lines represent new added edges.

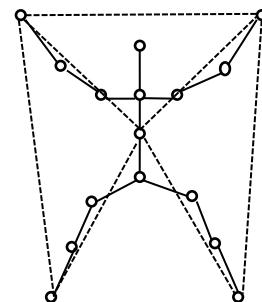


Figure 5. Solid lines represent bones and dotted lines represent new added edges.

3.3. Dispersing error rate

At present, almost all algorithm models adopt one-hot mode when they express the sample target category, that is, the target category is set as a vector y with length N of the total number of categories, where the i th element of vector y representing class i is 1 and other elements are 0. It is found in the experiment that if the length of vector y is doubled, that is, $2N$, and the i th element of vector y representing class i is 1, and all other elements are 0, the model accuracy will be improved. For the test sample x with class i , whether the system model can correctly identify sample x can be judged according to whether its output value on sample x is i . Therefore, in the case of expanding the number of categories, even if the output of the system model on the input sample x exceeds the initial sample category range, it only indicates that the system cannot correctly identify sample x , but the system can still operate normally. The i th element of the category vector y_x of the sample x of class i is 1 and the rest

elements are 0. According to the output structure of the model, when training the model with the training sample x of class i , it is expected that the model output is as close as possible to the class vector y_x of sample x , that is, the i th element of the output vector \hat{y}_x of the model should be as close to 1 as possible, and other elements should be as close to 0 as possible. For any test sample x' , when it is fed into the trained model, it is assumed that the probability of its output correct category is p and the error probability is q . When $p \gg q$, the model outputs almost the correct category, but when p approaches q or is less than or equal to q , the model outputs the wrong category more likely. The model output should be one of the other $N - 1$ categories of the incorrect category, denoted j . The probability of the model output class j is $q/(N - 1)$ under the assumption of equal probability. If the length of the class vector y becomes $2N$ without changing the total number of classes, the model outputs class j with a probability of $q/(2N - 1)$. Thus, under the assumption of equal probability, appropriately extending the length of category vector may reduce the probability of outputting a certain error category. However, experiments show that the length of category vector should not be increased too much, which will reduce the accuracy of model output. The reason may be that adding too many categories reduces the probability of output correct categories while reducing the error categories of the model. Therefore, in this paper, all experiments increase the length of category vector by 1 or 0.5 times.

4. Analysis of experimental results

To demonstrate the effectiveness of the proposed approach, three databases commonly used to evaluate skeleton-based human recognition models, NTU-RGB-D60, NTU-RGB-D120, and Kinetics Skeleton 400, were selected as benchmark databases. All video segments of the database are processed by the method of eliminating redundant frames introduced in section 3.1, and the empty frames at the tail are filled with zero. Four models ST-GCN[15], 2s-AGCN[26], SGN[27], Shift-GCN[28] and MS-G3D[29] were selected as the benchmark models.

4.1. Experiments on NTU-RGB-D60 database

NTU-RGB-D60[30] contains 56,578 skeleton sequences and 60 action categories, all of which were completed by 40 different people and captured by 3 cameras from different perspectives. Each frame contains the three-dimensional coordinates of 25 body nodes. In cross-subject (X-sub), the training set contained 40,091 samples and the test set contained 16,487 test instances. In cross-view (X-view), the test set consisted of all 18,932 samples collected by camera 1, and the remaining 37,646 samples were used for training. All experimental results are listed in Table 1.

Table 1 Comparison of Top1 results on NTU-RGB-D60 database

The algorithm name	Xsub		Xview	
	Original	+our	Original	+our
ST-GCN	81.5	85.3	88.3	91.8
2sAGCN	88.5	89.3	95.1	95.7
SGN	89.0	89.5	94.5	95.1
Shift-GCN	90.7	91.0	96.5	96.5
MS-G3D	91.5	91.7	96.2	96.3

It can be seen from Table 1 that the proposed method improves the recognition results on the basis of the original algorithm, and greatly improves the recognition rate

of the basic GCN algorithm ST-GCN, and the correct recognition rate of Xsub and Xview are increased by 4.66% and 3.96% respectively. However, the improvement of correct recognition rate of MS-G3D algorithm is limited. The possible reason is that there is a limit on the correct recognition rate of GCN-based algorithm on this database, and there is not much range for improvement of correct recognition rate of algorithm.

4.2. Experiments on NTU-RGB-D120 database

NTU RGB+D 120[31] added an additional 57,367 skeleton sequences on the basis of NTU RGB+D 60, with a total sample number of 113,945, including more than 120 categories. In CrossSetup (X-set), the training Set contains 54,468 samples collected by half of the cameras, with the remaining 59,477 samples used for testing. In cross-subject, the training set consisted of 63,026 samples from 53 subjects, and the remaining 50,919 samples were used for testing. All experimental results are listed in Table 2.

Table 2 Comparison of Top1 results on NTU-RGB-D120 database

The algorithm name	Xsub		Xset	
	Original	+our	Original	+our
2sAGCN	82.9	84.5	84.9	86.3
SGN	79.2	81.0	81.5	83.2
Shift-GCN	85.9	86.7	86.2	87.4
MS-G3D	86.9	87.3	87.6	88.1

NTU-RGB-D120 has one times more activity categories than NTU-RGB-D60, and the correct recognition rate of each algorithm is obviously lower than that of NTU-RGB-D60. However, the correct recognition rate of these algorithm models is improved with the measures proposed by us. Table 2 shows that the correct recognition rate of 2sAGCN increases by 1.92% and 1.65% in Xsub and Xview respectively after adding our proposed strategy.

4.3. Experiments on Kinetics Skeleton 400 database

The Kinetics Skeleton 400[32] dataset contained 240,436 training samples and 19,796 test Skeleton sequences, and over 400 classes. The experimental results on this database are shown in Table 3.

Table 3 Comparison of Top1 results on Kinetics Skeleton 400 database

The algorithm name	Top-1(%)		Top-5(%)	
	Original	+our	Original	+our
ST-GCN	30.7	33.5	52.8	56.4
2sAGCN	36.1	37.0	58.7	59.3
MS-G3D	38.0	38.5	60.9	61.2

The Kinetics Skeleton 400 DATASET contains much higher activity categories than NTU-RGB-D. Table 3 shows that the performance of each algorithm is very sensitive to activity categories, and the correct recognition rate reduces significantly. However, the proposed strategy can still effectively improve the performance of each algorithm. Table 3 shows that the correct recognition rate of ST-GCN improved by 9.1% and 6.8% in Top-1 and Top-5 respectively after adding our proposed strategy.

5. Conclusion

In this paper, we propose a new method of video frame reduction, which transforms a high-dimensional video frame sequence into a one-dimensional time series by algebraic method. Experiments show that the one-dimensional sequence can effectively represent the movement trend of limb movement, and the extreme point of the sequence corresponds to the time when the movement trend changes, while the non-extreme point of the sequence has little significance for the identification of the whole movement. In this way, we can select the video frame corresponding to the extreme point of the one-dimensional sequence to form a new sequence, so as to achieve the purpose of reducing the redundant video sequence signal. At the same time, we added lines between some non-physiologically connected joints in the human skeleton diagram to enhance the ability of representing actions by the relative position relationship between key joints. Finally, the false recognition rate is reduced by adding virtual categories. Experimental results show that our strategy can effectively improve the correct recognition rate of the original algorithm model.

Acknowledgment

This research is supported by the National Natural Science Foundation of China (61866003)

References

- [1] WANG H, KLASER A, SCHMID C, et al. Action recognition by dense trajectories[C]//Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, Jun 20-25, 2011. IEEE, 2011: 3169-3176.
- [2] WANG H, SCHMID C. Action recognition with improved trajectories[C]//International Conference on Computer Vision, Sydney, NSW, Australia, Dec 1-8, 2013. IEEE, 2014: 3551-3558.
- [3] D. Marr, H.K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 1978, 200 (1140): 269-294.
- [4] D. Hogg. Model-based vision: a program to see a walking person. *Image and vision computing*, 1983, 1 (1): 5-20.
- [5] K. Rohr. Towards model-based recognition of human movements in image sequences. *CVGIP-Image Understanding*, 1994, 59 (1): 94-115.
- [6] LAPTEV I. On space-time interest points[J]. *International Journal of Computer Vision*, 2005, 64(2-3): 107-123.
- [7] WANG H, ULLAH M M, KLASER A, et al. Evaluation of local spatio-temporal features for action recognition[C]//CAVALLARO A, British Machine Vision Conference, 2009. BMVA, 2009: 124.1 - 124.11.
- [8] WILLEMS G, TUYTELAARS T, GOOL L V. An efficient dense and scale-invariant spatio-temporal interest point detector[C]//FORSYTH D, 2008 European Conference on Computer Vision, Springer, Berlin, Heidelberg. 2008, 5303: 650-663.
- [9] YANG X D, TIAN Y L. Effective 3D action recognition using eigenjoints[J]. *Journal of Visual Communication and Image Representation*, 2014, 25(1): 2-11.
- [10] ZHANG H X, YE Y S, CAI X Z, et al. Efficient algorithm of action recognition based on joint points[J]. *Computer Engineering and Design*, 2020, 41(11) : 3168-3174
- [11] Wangqing Li, Zhenyu Zhang, Zicheng Liu, Li, W., Zhang, Z. & Liu, Z. (2010). Action recognition based on a bag of 3D points. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010 (pp. 9-14). Piscataway, New Jersey, USA: IEEE.
- [12] [Muhammet Fatih Aslan, Akif Durdu, Kadir Sabanci](#). Human action recognition with bag of visual words using different machine learning methods and hyperparameter optimization. *Neural Computing and Applications*, 32, 8585-8597(2020).

- [13] Yong Du, Wei Wang, and Liang Wang. Hierarchical recurrent neural network for skeleton based action recognition. In CVPR, 2015.
- [14] QiuHong Ke, Mohammed Bennamoun, Senjian An, Ferdous Sohel, and Farid Boussaid. A new representation of skeleton sequences for 3d action recognition. In CVPR, 2017.
- [15] Sijie Yan, Yuanjun Xiong, and Dahu Lin. Spatial temporal graph convolutional networks for skeleton-based action recognition. In AAAI, 2018.
- [16] Wentao Zhu, Cuiling Lan, Junliang Xing, Wenjun Zeng, Yanghao Li, Li Shen, and Xiaohui Xie. Co-occurrence feature learning for skeleton based action recognition using regularized deep lstm networks. In AAAI, 2016.
- [17] Pengfei Zhang, Cuiling Lan, Junliang Xing, Wenjun Zeng, Jianru Xue, and Nanning Zheng. View adaptive recurrent neural networks for high performance human action recognition from skeleton data. In ICCV, 2017.
- [18] Chenyang Si, Ya Jing, Wei Wang, Liang Wang, and Tieniu Tan. Skeleton-based action recognition with spatial reasoning and temporal stack learning. In ECCV, 2018.
- [19] Amir Shahroudy, Jun Liu, Tian-Tsong Ng, and Gang Wang. Ntu rgb+ d: A large scale dataset for 3d human activity analysis. In CVPR, 2016
- [20] Jun Liu, Amir Shahroudy, Dong Xu, and Gang Wang. Spatio-temporal lstm with trust gates for 3d human action recognition. In ECCV, 2016.
- [21] Pengfei Zhang, Cuiling Lan, Junliang Xing, Wenjun Zeng, Jianru Xue, and Nanning Zheng. View adaptive neural networks for high performance skeleton-based human action recognition. TPAMI, 2019.
- [22] Junwu Weng, Mengyuan Liu, Xudong Jiang, and Junsong Yuan. Deformable pose traversal convolution for 3d action and gesture recognition. In ECCV, 2018.
- [23] Congqi Cao, Cuiling Lan, Yifan Zhang, Wenjun Zeng, Hanqing Lu, and Yanning Zhang. Skeleton-based action recognition with gated convolutional neural networks. TCSVT, 29(11):3247–3257, 2019.
- [24] Yansong Tang, Yi Tian, Jiwen Lu, Peiyang Li, and Jie Zhou. Deep progressive reinforcement learning for skeleton-based action recognition. In CVPR, 2018.
- [25] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. arXiv, 2016.
- [26] Lei Shi, Yifan Zhang, Jian Cheng, and Hanqing Lu. Twostream adaptive graph convolutional networks for skeletonbased action recognition. In CVPR, 2019 In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 12026–12035, 2019.
- [27] Pengfei Zhang, Cuiling Lan, Wenjun Zeng, Junliang Xing, Jianru Xue, Nanning Zheng. Semantics-Guided Neural Networks for Efficient Skeleton-Based Human Action Recognition. In CVPR2020.
- [28] Ke Cheng, Yifan Zhang, Xiangyu He, Weihan Chen, Jian Cheng, Hanqing Lu. Skeleton-Based Action Recognition with Shift Graph Convolutional Network. In CVPR2020.
- [29] Ziyu Liu, Hongwen Zhang, Zhenghao Chen, Zhiyong Wang, Wanli Ouyang. The Disentangling and Unifying Graph Convolutions for Skeleton-Based Action Recognition, In CVPR2020.
- [30] Amir Shahroudy, Jun Liu, Tian-Tsong Ng, and Gang Wang. Ntu rgb+ d: A large scale dataset for 3d human activity analysis. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1010–1019, 2016.
- [31] Jun Liu, Amir Shahroudy, Mauricio Perez, Gang Wang, Ling-Yu Duan, and Alex C. Kot. NTU RGB+D 120: A large-scale benchmark for 3d human activity understanding. CoRR, abs/1905.04757, 2019.
- [32] Will Kay, Joao Carreira, Karen Simonyan, Brian Zhang, Chloe Hillier, Sudheendra Vijayanarasimhan, Fabio Viola, Tim Green, Trevor Back, Paul Natsev, Mustafa Suleyman, and Andrew Zisserman. The kinetics human action video dataset, 2017.

An Empirical Study on the Fossilization of English Language Learning in the Context of Multimedia Network Teaching

Lihua HUANG¹

Guangdong Vocational College of Hotel Management, Dongguan, China

Abstract. To carry out the pilot teaching of multimedia assisted teaching and self-learning mode of network platform for experimental class, the data analysis results obtained by using tools such as questionnaire survey show that: The multimedia network teaching environment has a great effect on the students' vocabulary knowledge fossilization, but has little effect on the grammar knowledge fossilization; The fossilization of reading ability among the four communicative abilities of English language is alleviated obviously, and the self-assessed oral ability of students improves faster after the addition of oral English classes; It accelerates the ossification of writing and listening ability, but it is mainly caused by the characteristics of the network itself and improper learning methods of students.

Keywords. Multimedia network, teaching environment, learning fossilization, empirical research

1. Introduction

Multimedia network-assisted English teaching has been given the sacred mission of updating teaching methods, reforming teaching models, improving teaching efficiency, and promoting teaching reform since its birth. It provides a large number of convenient and rapid visual, three-dimensional, interactive teaching based on audio-visual and communication, so that learners can become more effective processors of information and constructors of knowledge meaning, which greatly enriches teachers' classroom teaching. The student autonomous learning model under the network platform has greatly exercised and improved the ability of students to use network resources for autonomous learning, and gave full play to the subjective initiative of students in learning. However, after a period of teaching practice, we find that the fossilization of English language learning that is common in the traditional teaching mode still exists in the multimedia and network teaching environment, and has its own characteristics. Multimedia classroom teaching and the autonomous learning model under the network platform can activate the fossilization in some aspects, but in other respects it accelerates the formation of the fossilization in English language learning. It is of great significance to study the characteristics and causes of such fossilization for optimizing multimedia and network teaching and making better use of modern technology for English teaching. In the theory of fossilization, the term language fossilization was first put forward by American linguist Larry Selinker(1972)[1]. He believes that most second language learners will

¹ Corresponding author: 67655200@qq.com

ossify in the acquisition of language ability after the foreign language learning has reached a certain level. The study of fossilization was carried out earlier by foreign scholars. The research focuses on the causes of fossilization, such as Selinker (1972), Lenneberg (1967) [2] and Lamendella (1979)[3], Schumann(1975)[4], Vigil and Oler(1976)[5], Ellis(1994)[6], Krashen(1985)[7], Sims(1989)[8] etc. Domestic scholars have been paying attention to the phenomenon of fossilization since the beginning of this century. Niu Qiang (2000)[9] believes that four factors, including emotional factors, interference of mother tongue and transfer, teachers and teaching materials, are related to the causes of fossilization. Liu Runqing (2003)[10] conducted a large-scale questionnaire survey with students from key universities as the research object. Relevant data statistics and analysis show that language learning fossilization is common among students. He believes that the main causes of fossilization are cognitive perspectives, learners' personal factors and learning atmosphere, etc., and put forward many feasible measures to improve fossilization.

However, with the continuous deepening of higher vocational English teaching reform, multimedia assisted teaching equipment and network platforms have been introduced into English classrooms and extracurricular autonomous learning, and the fossilization of students' English language learning has been given new characteristics. Xu Yushu (2006)[11] explored the phenomenon of English learning fossilization in higher vocational colleges under the multimedia teaching mode. Wang Shunling and Zhang Faye (2006)[12] studied the fossilization of college English learning in the context of multimedia projection, and found that the rigidity of students' vocabulary and grammar was serious due to the factors of multimedia itself, learner's personal factors and teacher's factors. Hu Gang and Sheng Fan (2007)[13] analyzed the causes of transitional language fossilization under the new environment of multimedia assisted foreign language teaching and put forward some solutions. Since multimedia classroom teaching and the autonomous learning model under the network platform have a new perspective on the role of fossilization phenomena, there are few relevant studies. Through experimental class pilot teaching and questionnaire survey, this research aims to study the following questions: First, what is the relationship between multimedia-assisted classroom teaching and autonomous learning on the network platform and rigid phenomena. Second, whether this mode will accelerate the occurrence of some language proficiency fossilization. Through specific operations in the experimental class, students' opinions and suggestions are statistically analyzed, so that the new teaching model is more targeted, more in line with the specific needs of students, and more conducive to comprehensive promotion and spreading.

2. Research Methods

2.1 Research Subjects and Goals

The research subjects are 238 students in 8 administration classes of non-English majors in the 2018 grade, including 138 boys and 100 girls, of which 111 are liberal arts students and 127 are science and engineering students. As experimental classes, these 8 administrative classes have been using multimedia network to assist English class teaching since the first semester of freshman year in September 2018. A one-year English course is offered, with an average of 54 hours of three lessons per week. In addition, students are required to use the network platform to conduct independent learning outside the classroom, in which the listening training part no longer occupies the classroom

teaching time, but becomes a part of the network independent learning content. Questionnaires were sent to students twice at the beginning of the first semester and at the end of the second semester of the first year of college, in order to grasp the expression forms and changes of fossilization in English learning before and after multimedia and network classroom teaching. And the students' opinions and suggestions are statistically analyzed, so that the multimedia network classroom teaching model is more optimized, more activated, to avoid the emergence of new language fossilization.

2.2 Research Tools and Content.

The measurement tools used in this study are the "Questionnaire on the Fossilization of English Learning" and the "Questionnaire on English Learning in the Multimedia Network Teaching Environment". The first questionnaire was issued after the freshmen enrolled in October 2018. The questionnaire mainly refers to Liu Runqing's (2003) "Distribution Table and Causes of the fossilization of Chinese College Students", which involves three major factors: the existence of fossilization in English learning of college students, the distribution of fossilization, and the causes of fossilization. In terms of questions, the questionnaire is based on multiple-choice questions. The second questionnaire was issued in June 2019, that is, after one year of multimedia online classroom teaching in the experimental class. The second questionnaire is formulated in accordance with the actual situation of our school. The main contents include: Does the multimedia network teaching environment have an activating effect on the fossilization of English learning, and in what ways? Does such an environment accelerate the occurrence of certain fossilization phenomena? Suggestions and opinions on this teaching mode and so on. The questionnaire is based on open questions.

3. Research Result

3.1 The Overall Situation of Fossilization

According to the statistics of the two questionnaires, 198 students (83.19% of the total number of participants in the survey) and 152 students think that their English learning is rigid (63.87%). In addition, the majority of students hold a positive attitude towards the promotion of multimedia and network teaching environment on English learning: 195 students (81.93%) believe that the multimedia network teaching environment can alleviate the occurrence of fossilization ; 173 students (72.68%) believe that they are confident to make progress in English language learning.

3.2 Fossilization Distribution

The statistics of the first questionnaire show that: The proportion of fossilization English vocabulary knowledge is the highest, with a total of 167 people, reaching 70.17%. The second most rigid area is grammar, with a total of 127 people, accounting for 53.36%. The four major English communicative competences of speaking, listening, reading, and writing all have fossilization. Among them, the fossilization of oral communicative competence is the most serious, with a ratio of 67.65%, a total of 161 people; the fossilization proportion of listening ability is 61.34%, with a total of 146 people, ranking second; and the proportion of reading ability is the third, up to 50.42%, a total of 120 people; although the fossilization proportion of writing ability is ranked the last, it also reaches 46.64%, totaled 111 people, showing a serious degree of rigidity.

Statistics from the second questionnaire show that the proportion of rigid English grammar knowledge is the highest, with a total of 120 people, reaching 50.42%; the proportion of lexical rigidity falling to the second place, with a total of 92 people, accounting for 38.66%. The proportion of other kinds of English knowledge rigidity is still less than 20%, which is almost the same as the results of the previous questionnaire, and the research is of little significance. Among the four major English communicative competences, the rigidity of listening and communication skills rose to the first place, with a total of 172 people, accounting for 72.27%; the rigidity of writing skills ranked second, reaching 56.3%, a total of 134 people; the proportion of reading skills was still ranked thirdly, it is 41.6%, a total of 99 people, and the rigid proportion of speaking ability drops to 34.87%, a total of 83 people.

3.3 Comparative Analysis of Questionnaire Results

The two questionnaires in this study were issued at the beginning of the freshman enrollment and at the end of the freshman year about a year later. During this period, teachers used multimedia network classroom teaching for two consecutive semesters to the students in the experimental class, and set additional oral English classes and increased the time for independent online learning. From the results of the two questionnaire surveys, it can be seen that the rigidity of English language learning is widespread among the junior college students. In general, the fossilization was greatly alleviated after the use of multimedia network teaching mode, which decreased by about 20 percent.

Since most colleges and universities introduce multimedia assisted teaching equipment in English classes at the time of higher vocational colleges are expanding their enrollment, most of the classes using multimedia assisted teaching equipment are combined classes of two or more administrative classes with at least 40 or 50 students. It is predicted that such a large class is not helpful to the cultivation and improvement of students' oral English ability, therefore, a single-class oral class with an average of one hour per week is offered for non-English majors in freshmen and sophomores. In addition, because students generally report that the effectiveness of collective listening lessons in the classroom is not good, the teaching content of the listening part has also been separated from the English class. Listening homework is a part of the exercises that must be completed, and students need to learn the listening content on an online platform after class. One of the expected effects of the above two teaching reform measures is to prevent the rigidity of oral English and listening learning under the multimedia network teaching mode. According to the results of this survey, the statistical data of the rigidity of students' self-tested English learning after one year of English classroom teaching using the multimedia network teaching model has the following changes:

The number of students who believe that their oral English proficiency is fossilized decreased from 161 to 83, and the fossilized proportion of oral English proficiency changed from the highest to the lowest among the four kinds of English communicative competence, showing a huge change before and after. 83.31% of the students believe that the new oral class is the main reason for this change.

The rigidity of listening and communication skills rose from second to first, and the proportion increased by more than ten percentage points. 85.67% of the students felt that using the online self-learning platform to learn listening has accelerated the occurrence of listening ability fossilization.

The ranking of reading ability fossilization remains unchanged, ranking the third, but it has decreased by nearly 10 percentage points, which proves that the fossilization

of reading ability has been greatly alleviated.

The proportion of writing ability fossilization rose from the last place to the second place, with an increase of 10 percentage points, indicating a decline in students' writing ability.

4. Discussion and Inspiration

Through the above data analysis and comprehensive statistics of students' answers to the open questions in the second questionnaire, we can draw the following conclusions from this teaching practice:

Generally speaking, the multimedia and network teaching environment has a considerable effect on activating and alleviating the rigidity of the English language. Students generally believe that compared with the traditional teaching model (blackboard + chalk + teacher), multimedia-assisted classroom teaching and the self-learning model of the network platform are more helpful to improve their English language comprehensive ability and stimulate learning interest and motivation.

In the aspect of oral communication ability, multimedia network teaching environment is more likely to cause rigidity than traditional classroom teaching. The reasons are as follows: First, the communication between multimedia auxiliary devices and students is mostly one-way indoctrination. Second, under the co-class teaching method, there are a large number of people in class. Third, there are few opportunities for oral communication between teachers and students as well as between students. However, with the addition of oral English classes, the situation is greatly improved, and the students' oral English ability has entered an upward curve from a plateau period or even a decline period. 68.91% of the students think that there is only one class per week in oral English class, and the class hours are too short, and they hope to increase the proportion of oral class teaching. This fact shows that opening oral English class separately is an effective measure to alleviate the rigidity of oral English ability under the multimedia network teaching environment. It can create an environment and atmosphere for students to practice oral English regularly under the guidance of teachers for a long time. This cumulative and imperceptible effect is still very significant.

After students use the online platform self-learning model to learn the content of classroom listening, the number of students who believe that their listening ability is rigid has increased by about 10% compared with the previous one. From the data, it seems that the autonomous learning model of the network platform has more catalyzed the rigidity of students' listening ability. However, after detailed statistics and analysis of the answers to open questions, we found that nearly seventy percent of the reasons why students do not like to train their listening skills in the form of online self-learning are improper learning methods. Such as: The listening time is too long each time, and the learning effect is poor due to fatigue; Students often choose to give up learning because the selected listening materials are too difficult to understand; The time period the students choose to practice listening is not ideal, which leads to their failure to form a virtuous cycle of biological clock; The students lack sufficient personal self-control, so that they are attracted by other content on the Internet, and unable to concentrate on listening exercises, etc. The reasons for this are all because in the traditional classroom listening teaching, from the selection of listening materials, to the control of the duration

and frequency, to the explanation and re-listening of key and difficult points, all are carefully designed and arranged by the teacher. Therefore, the timely guidance of teachers, especially the guidance of listening learning methods, is of vital importance for the students who use the network self-learning mode to train their listening.

The multimedia network teaching environment also has a considerable effect on alleviating the rigidity of English reading ability. As is known to all, traditional English classroom teaching mainly relies on books, chalk and teachers, which can convey quite limited information, while multimedia assisted classroom teaching and independent learning on the network platform provide students with a large amount of information such as English texts, pictures, audio and video. In this case, the average amount of information in a class can be several times or even dozens of times that of a traditional class. The commonly used English vocabulary, sentences and chapters contained in this information are repeatedly presented to students. The huge amount of reading and the continuous recurrence of reading content are undoubtedly beneficial to the improvement of students' reading ability.

From the data point of view, the multimedia network teaching environment will aggravate the rigidity of students' writing ability to some extent. Co-class teaching has doubled the amount of teachers' homework corrections, and the speed of teachers' review and feedback of students' compositions has been reduced. Although in some teaching method reform classrooms, students' mutual evaluation and teachers' selective evaluation are used to reduce the workload of teachers and mobilize students' enthusiasm to participate in the evaluation of classmates' compositions. However, the implementation effect is not satisfactory due to the differences in students' attitudes and limitations. Students still prefer teachers to review their compositions. On the other hand, the main reason for the rigidity of students' writing ability does not lie in the adoption of multimedia network teaching model itself, but in the richness of network resources. The amount of information available on online platforms makes it easier and easier to write an English composition. Some students even sign their names on the searched materials, leaving the rest of the content unchanged and handing it over as their own homework. The massive amount of information also makes it increasingly difficult to identify the originality of an article, which can no longer be done by the teachers alone. It is not surprising that the lack of writing training opportunities leads to rigid writing skills.

The multimedia network teaching environment has a greater activation effect on the rigidity of students' vocabulary knowledge, but has little effect on the rigidity of grammatical knowledge. The repetition of common English vocabulary and the doubling of reading volume will naturally enable students to master more vocabulary. As for grammar, students do not pay much attention to the study of grammar because grammar knowledge has been basically taught in high school English classes, while it is only deepened and consolidated and reviewed in college English classes. When students use the network to study independently, they are more likely to do listening, speaking and reading training rather than learning grammar on the computer. In addition, the mastery of grammar knowledge requires more lectures and guidance from teachers than the mastery of other English knowledge, so the multimedia network environment has little effect on the rigidity of grammar knowledge.

The influence of multimedia network teaching environment on English language fossilization, whether positive or negative, is the result of interaction of various factors. It's believed that more relevant research will enable educators to make better use of educational technology to serve teaching.

References

- [1] Selinker, L. Inter-language. International Review of Applied Linguistics, 1972(10).
- [2] Lenneberg, E. Biological Foundations of Language. New York: Wiley and Sons, 1967, 23-56.
- [3] Lamendella, T. General Principles of Neuro-functional Organization of Their Manifestations in Primary & Non-primary Language Acquisition. Language Learning 1977(27).
- [4] Schumann, J. Second language acquisition pidginization hypothesis. Language Learning, 1975 (26)
- [5] Vigil, F. & Oller. Rules of Fossilization: A Tentative Model. Language Teaching, 1976 (26).
- [6] Ellis, Rod. The Study of Second Language Acquisition. Shanghai: Shanghai Foreign Language Education Press, 1999, pp 55.
- [7] Krashen, S D. The Input and Practice: Issue and Implication. London: Longman, 1985, pp 64-71.
- [8] Sims, W. R. Fossilization and Learning Strategies on Second Language Acquisition (ERIC Document Reproduction Service No. ED339 195), 1989.
- [9] Niu Qiang. The Fossilization of Transition and Its Implications in Teaching [J]. Foreign Language and Foreign Language Teaching, 2000(5).
- [10] Liu Runqing, Dai Manchun. Research on the status quo and development strategies of foreign language teaching reform in Chinese universities [J]. Beijing: Foreign Language Teaching and Research Press, 2003: 263-274.
- [11] Xu Yushu. Research on the rigidity of English learning under the multimedia teaching mode of higher vocational colleges [J]. Journal of Chinese Language and Literature, 2006(S1).
- [12] Wang Shunling, Zhang Fak. A study on the fossilization of college English learning in the context of multimedia projection [J]. Audio-visual Teaching of Foreign Languages, 2006(8).
- [13] Hu Gang, Sheng Fan. On multimedia-assisted foreign language teaching and interlanguage rigidity [J]. Economic and Social Development, 2007(11).

A Study on the Efficiency and Countermeasures of Network-Based Autonomous Learning Platform for Vocational College English Learners

Liping JIANG¹

1School of Business and Foreign Studies, Guangdong Mechanical & Electrical Polytechnic, Guangzhou, China

2School of Social Sciences and Humanities, University College Sedaya International, Kuala Lumpur, Malaysia

Abstract. This paper investigates the efficiency of 289 vocational college English learners using the network-based autonomous learning platform in Guangdong province in China. The results show that learners have a positive attitude toward the network-based extracurricular English autonomous learning, and learners with different English foundation have different requirements for extracurricular online autonomous learning. This paper puts forward some countermeasures on how to make full use of the network-based extracurricular English autonomous learning.

Keywords. Vocational college English learners, network autonomous learning platform, using efficiency, countermeasures

1. Introduction

With the continuous development of modern education and information technology, autonomous learning based on computer and network has gradually become a teaching assistant mode. "Regulations for Teaching College English Curriculum" (2007) puts forward that "a lot of advanced information technology should be used to develop and construct various courses based on network and computer".[1] In order to adapt to the new situation of the development of higher vocational education in China, deepen the teaching reform, improve the teaching quality, and meet the needs of the state and society for the cultivation of higher vocational talents in the new era, the document of *The basic requirements for English Teaching in Higher Vocational Education (Trial Implementation)* and *The outline of the national medium and long term education reform and development plan (2010-2020)* are issued. The *Guidance* of the two documents clearly puts forward the requirements of cultivating digital information competency, and clearly proposes that learners' autonomous learning ability should be enhanced in vocational college English [2]. English autonomous learning based on modern information technology has become a hot topic in the construction and research of higher Vocational English course in college. Network-based autonomous learning platform

¹ Corresponding author: 1531447813@qq.com

model is gradually accepted by the majority of college teachers and learners, and has been used in assisting teaching. Many powerful publishing institutions in China have also launched online courses based on teaching materials; Language testing equipment suppliers are actively developing Network-based Autonomous Learning Platform; Some colleges and universities are also investing greatly in the construction of Network-based Autonomous Learning Environment with school-based characteristics. The construction and application of network-based autonomous learning platform will be one of the key research directions of Vocational College English autonomous learning in the future. However, at the same time, due to the current Chinese network-based autonomous learning platform is still in the early stage of development, learning content, management mode and self-learning mode are still in the initial phase of development, therefore, there are many problems to be solved.

2. Research background

As an indispensable part of higher education in China, higher vocational colleges shoulder the mission of cultivating professional and technical talents. In 2017, According to the document of *The list of national colleges and universities issued* by the Ministry of Education in China, the level of a higher education in China can be divided into undergraduate universities and higher vocational colleges (2017, Ministry of Education in China). As the pioneer of China's reform and opening up, Guangdong Province is one of China's most economically developed provinces. The number and teaching quality of higher vocational colleges rank at the forefront of China. According to the national vocational college professional setting management and public information, the service platform data of January 2020, there are 89 higher vocational colleges in Guangdong province. Vocational college English is a compulsory public course for non-English majors in the vocational colleges, which will open for them to study at the first year of college. Vocational college English learners are from both liberal arts and science. However, the academic performance of learners in vocational colleges is lower than those from undergraduate universities and colleges due to the different level of higher education, therefore, the majority of vocational college learners' English foundation is relatively poor. Therefore, it is more necessary to make full use of the advantages of blended learning and fully improve learners' adaptability of blended learning to improve the efficiency of English learning.

Can network-based autonomous learning platform promote vocational college English learners' autonomous learning? What factors restrict the autonomous learning ability of vocational college English learners? What should teachers do for learners' autonomous learning? What is the current status of vocational college English learners' autonomous learning ability? For those problems, Chinese domestic scholars have conducted in-depth research and discussion. Professor Zhu (2011) investigated the current situation of college learners' English autonomous learning, and concluded that learners have certain English autonomous learning ability, but there are also some problems. He proposed that the learners should be trained and guided in the sense of autonomous learning, learning strategies and learning methods.[3] Professor Guo (2008) discussed the guiding role of teachers in college learners' English autonomous learning and proposed that teachers should play a leading role in cultivating and promoting learners' autonomous learning. Without the teacher's introduction, it is difficult for learners' autonomous learning to achieve results.[4] Professor Zhang (2018) introduced

the application of online autonomous learning platform, pointed out the existing problems, and gave suggestions and countermeasures.[5] Professor Fan and Zhu (2018) demonstrated that the construction and application of college English online learning platform has a significant assistance and positive impact on the improvement of vocational college learners' English autonomous learning ability. [6] Scholars Li and Chen (2019) effectively integrate modern network information technology with business English course teaching, and schools and teachers should scientifically design curriculum system and teaching content.[7] However, some people still worry about the current status of learners' online autonomous learning as they point out that learners' extracurricular autonomous learning status and learning effect are not optimistic due to the influence of learning environment. The previous studies investigated the application and efficiency of network-based autonomous learning platform in colleges. However, there is few researchers did research on the efficiency and countermeasures on the application of the platform based on questionnaire survey. Besides, the research objects have been restricted within students in colleges and universities. Taking the vocational college English learners in researcher's vocational college as a research example, this paper makes a study on the current status and efficiency of learners using the network-based autonomous learning platform based on questionnaire survey.

3. Research design

3.1 Research participant

The research participants are the first year vocational college English learners who have been engaging in Extracurricular Autonomous Learning platform for one semester. The learners were required to log into the college's Blue Pigeon Network-based Autonomous Learning Platform for autonomous learning when they entered the college as freshmen. Blue Pigeon Network-based Autonomous Learning Platform is a network-based autonomous learning system. learners can log into the platform in the school's autonomous learning center, or log into the platform with any computers in the dormitory or library. They are required to study on it for 2 hours a week, and the learning content is regulated and optional. learners' autonomous learning after class will account for 50% of their regular scores which will account for 40% of the final total scores. That is to say, learners have to log into the network platform for autonomous learning after class in order to achieve satisfied marks. Their language teachers are mainly responsible for the specific learning content and learning supervision.

3.2 Research questions

This study mainly tends to answer the following three questions:

- What is the current status of learners using autonomous learning platform after class?
- What is the learners' evaluation on the using effect of self-learning platform?
- What are learners' opinions and suggestions on school autonomous learning and relevant regulations?

3.3 Research tool

In terms of the above research problems, the researcher developed a questionnaire in which one question item is made for one question.

- Whether you actually use the autonomous learning platform for learning after class;
- Your evaluation of the effect of using the autonomous learning platform;
- Improvement measures of the autonomous learning platform;

In order to listen to the learners' voice, there is an open question for short answer at the end of the questionnaire which requires learners to put forward subjective opinions and suggestions to the autonomous learning platform and relevant rules and regulations of the school.

3.4 Research procedure

The questionnaire is carried out through network survey, which is carried out through the questionnaire function of the Blue Pigeon Autonomous Learning Platform. The specific survey time designed from April 6th, 2021 to April 16th, 2021. The platform will automatically pop up a questionnaire after learners logged into the platform. The questionnaire description clearly tells the learners that the questionnaire is anonymous, and the questionnaire results are only for scientific research. The main purpose of the questionnaire is to provide a reliable basis for the college to formulate relevant learning policies. Whether or not to participate in the questionnaire is decided by the learners themselves according to their personal decision. In order to ensure that every student is aware of the questionnaire, the researcher asked teachers to publicize the questionnaire in class. A total of 289 learners (with slightly different number of participants for each question) participated in the survey. After the formal completion of the questionnaire, the researcher conducted statistical analysis on the questionnaire. In order to deeply understand the relationship between learners' options and personal English foundation, the researcher obtained the scores of Practical English Test for Colleges, short for PRETCO of the learners from the school administration department. PRETCO is a very authentic and popular test for vocational colleges to evaluate learners' English efficiency level, and the score range of PRETCO is from 0 to 100 with passing mark of 60. Researcher used the batch query function of Excel to query and analyze the PRETCO scores of the learners participating in the questionnaire.

4. Statistical analysis on the using effect of platform

The researcher used a SPSS 22.0 statistical software to make statistics on the survey results of the number of platform users and the use effect of the platform. The results are shown in the cross table data in table 1.

Table 1. Cross table of percentage of users and percentage of use effect (Number: 289)

Percentage of users	Study every week	Number and percentage of users			total
		Very useful	Certain effect	useless	
	Study every week	48	52	10	110 (42.47%)
	Occasionally used	8	116	16	140 (48.44%)
	Seldom use	4	8	18	30 (10.38%)
Total		60 (20.76%)	176 (60.89%)	32 (11.76%)	289

Researcher conducted a Chi square test on the survey results of the number of

platform users and the use effect of the platform to test whether there is a correlation between the two survey data. The results are shown in table 2.

Table 2. Chi square test

	value	freedom	Asymptotic significance (bidirectional)
Pearson Chi square	143.823 ^a	4	.000
Likelihood ratio	122.121	4	.000
Linear correlation	84.424	1	.000
Number of valid cases	289		

a. Two cells (22.1%) had an expected count of less than 5. The minimum expected count was 2.35.

From the chi square test results in table 2, it is clear that the P value of chi square test is 0.000, less than 0.05, the survey data is significant, and it shows that the row variables and column variables are not mutually independent, that is, the percentage of use effect and the percentage of users are not independent, but related.

Table 3. symmetry measures

		Value	Asymptotic standard error	T^b last read	Significance of last read
Name	to	Phi	.74		.000
name			1		
		Cramer V	.52		.000
			1		
Interval	to	Pearson's R	.56		
interval			2	.047	11.164
Order	to	Spearman correlation	.56		.000 ^c
order			0	.047	10.910
Number of valid cases			28		.000 ^c
			9		

a. There is no null hypothesis.

b. The null hypothesis is assumed using asymptotic standard error.

c. Based on nominal approximation.

From the symmetrical measurement value in table 3, Phi value is equal to 0.741, P = 0.000, indicating that there is a positive correlation between the percentage of users and the percentage of use effect, that is, the higher the percentage of users, the higher the percentage of use effect. The value of Cramer V is less than. 0005, which also shows the same result.

5. Analysis and discussion

From the data of the use effect and number of self-learning platform in Table 1: only 20.76% of the learners said "very useful", 60.89% of the learners chose "have a certain role", and 11.76% of the learners still chose "completely useless". The total percentage of learners who choose "very useful" and "have a certain role" is over 80%, which

indicates that learners hold a positive attitude towards the effect of autonomous learning platform. From the statistics of PRETCO scores, learners with poor English foundation are more inclined to improve their basic English ability through autonomous learning platform. Most of the learners think that all aspects of the autonomous learning platform should be improved. In fact, the network autonomous learning platform of college is still in the initial stage of operation. All aspects are not perfect, and all rules and regulations have not been established. This is the focus of the next step.

Then the researcher collects the improvement measures proposed by learners, and provide five alternatives, as shown in Table 4

	Richer learning resources	Strengthen management and guidance	Simplify study content	More learning materials	Materials can be downloaded
Numbers	194	120	104	66	30
Percentage	67.12%	41.52%	35.98%	22.83%	10.38%

Table 4. Suggestions for using autonomous learning platform (Number: 289)

According to table 4, 67.12% of the learners choose to "provide rich learning resources", which shows that the learners' demand for learning resources is the first. The biggest difference between the blue pigeon Online Autonomous Learning Platform and the web-based online course developed by the publishing house is that the online course has existing learning materials, and teachers do not need or can not add various learning materials, that is to say, network course is a closed resource learning platform.

The percentage of choosing "teachers should strengthen the management of the platform and guide learners" accounted for 41.52%, which indicated that learners' autonomous learning in the autonomous learning platform needed teachers' participation. Some studies have pointed out that at present, the autonomous learning ability of college learners in China is still at a low level, and they can not completely leave the guidance of teachers for autonomous learning [8] [9]. Professor Guo (2008) pointed out that teachers are an important factor affecting college learners' English autonomous learning, and teachers have an unshirkable responsibility and obligation in cultivating learners' autonomous learning ability [3:120]. Teachers should train learners' autonomous learning ability in classroom teaching.

35.98% of the learners put forward the idea of "fragmenting the learning content to facilitate learning". This suggestion is related to the accelerating pace of learners' learning and life. learners hope to use the fragmenting time to learn. And 22.83% of the learners asked to "add the information of PRETCO". It shows that some learners hope to get help in exam materials. The construction of the database should follow the "three centers" principle, which is student-centered, interest centered and comprehensive application ability centered.

The open-ended survey allows learners to put forward their subjective opinions and suggestions on the autonomous learning platform of the college. A total of 160 learners participated in the survey. learners' suggestions can be summarized into the following categories: rules and regulations, learning materials, hardware construction, and suggestions for teachers. The details are as follows:

- Rules and regulations:

Strengthening the management of autonomous learning platform. The platform should increase more listening practising time and more English movies; I hope that I can get more scores for my usual performance because I am also learning listening and reading on the platform; Let learners take the initiative to learn, not just to learn English

for the scores passively, which can improve the enthusiasm of learners, so the learning effect will be improved; It's not very good to study things that are too compulsory;

At present, the interaction of autonomous learning platform is not enough. When learners encounter problems in the process of autonomous learning, they have no way to communicate with teachers timely. Therefore, when formulating rules and regulations, teachers should also be required to participate in the process of learners' autonomous learning.

- Studying materials:

Adding more English movies. Some learning difficulties should be explained in detail; The platform should add more abundant learning resources to our autonomous learning platform, so that the resources of the platform will not be too monotonous; We can watch movies and listen to songs on the independent platform, which will be more dynamic, and the learning effect will be better. Updating foreign language broadcasting tools, such as BBC; I hope that the proportion of time spent on watching movies should be lengthened.

In sum, it is necessary to increase and diversify the learning materials from time to time, and it is also necessary to put more PRETCO learning materials on it. The excellent open courses at home and abroad can also be introduced into the autonomous learning platform system, so that learners can fully enjoy all kinds of excellent teaching resources in the process of autonomous learning, and foreign language teaching can also be extended after class.

- Hardware construction:

Establishing Android version for mobile phone, so that learners can use their mobile phones to use platform independently anytime and anywhere without being restricted by learning sites and time; There are problems with autonomous learning platform, and sometimes we even can not log on it; I hope the college can strengthen the maintenance and management of the system.

Some learners, especially some impoverished learners, have not bought personal computers, so they can only study in the autonomous learning classroom of the college. However, at present, the number of autonomous learning classrooms can not fully meet the needs of learners' autonomous learning, therefore, it is necessary to increase the number of autonomous learning rooms.

- Suggestions for teachers:

I hope teachers should improve learners' understanding of the function of the platform, rather than simply let learners to use it completely independently; It's good to have a teacher to guide them online; To strengthen the communication between learners and teachers on how to use the platform.

An English communication channel between learners and teachers should be added to the autonomous learning system, so that learners can communicate with teachers at any time in the process of autonomous learning.

6. Conclusion and countermeasures

This paper makes a survey on the current situation of vocational college English learners' using of network-based autonomous learning out of class. The results show that learners hold a positive attitude toward the network-based extracurricular English autonomous learning, and learners with different English levels hold different views toward network-based Extracurricular Autonomous Learning platform. It is not enough just to rely on

regulations alone to mobilize the learners' enthusiasm for autonomous learning through the network platform after class,

First, teachers should be organized to participate the learning database in the platform to meet the personalized learning needs of learners. Autonomous learning database is the premise of learners' autonomous learning through the network platform. What resources should be collected in the database and whether there are learning resources that learners are interested in are important guarantee for learners to log into the platform for learning actively. Network-based autonomous learning platform is an open platform, therefore, in addition to its own learning resources, teachers should put various college-based learning resources according to learners' actual needs.

Second, teachers should actively participate in learners' autonomous learning ability training, learning strategy training and supervision and management. Many studies have shown that Chinese college learners' English autonomous learning level and ability are still at a low level therefore it is impossible to achieve success in extracurricular autonomous learning without teachers' participation. On the basis of hardware conditions, whether teachers participate in autonomous learning or not is a decisive factor affecting the success or failure of autonomous learning. Without teachers' active participation, autonomous learning is likely to face failure. Therefore, teachers must participate in learners' autonomous learning, supervise, guide and train learners' extracurricular autonomous learning.

Third, formulating relevant policies to support autonomous learning. The administrative departments should formulate policies to arouse the enthusiasm of teachers and learners. Only by mobilizing the enthusiasm of teachers and learners, can extracurricular network autonomous learning be implemented because the college's policy has a guiding and even decisive role. The formulation of the policy is mainly divided into two directions: To formulate relevant autonomous learning policies for learners to ensure their autonomous learning after class. According to the specific situation of the hardware equipment, college has carried out policy provisions on learners' autonomous learning, that is, they must study for 2 hours in the autonomous learning platform every week, which is one of the important sources of their score of regular performance. Compared with that of learners, the policy-making for teachers can not solve the problem with only a inflexible regulation, but should have certain incentive policies or workload regulations. Some teachers even think that guiding learners' autonomous learning is a kind of "extra work". Of course, this idea is not right. However, Participating in learners' extracurricular autonomous learning really increases the workload of teachers. Whether this part of the workload can be recognized or not and how to be recognized is an urgent issue for policy makers.

Fourth, providing special education technology training for teachers. The premise for teachers to participate in autonomous learning based on network platform is that teachers should have a certain computer and network foundation. It must be admitted that many English teachers have weak computer foundation, which still can not meet the needs of guiding learners' learning. Therefore, it is necessary to carry out special training for teachers to use the network platform.

At last, setting up special teaching research and teaching reform projects to encourage teachers to carry out teaching research based on network platform. The educational administration department of the college should conduct special research projects to encourage teachers to carry out teaching and research based on network autonomous learning platform. At present, the network-based teaching mode is very popular around the world, and attracts many people's attention. "Flipped classroom" and

"MOOC" have been widely accepted. The educational administration department should encourage English teachers to carry out "Flipped classroom", "MOOC" and other teaching practices based on the network autonomous learning platform.

Acknowledgment

2021 Guangdong Education Science Planning Project (Special program for Higher Education) "Research on SPOC-based Blended Learning Adaptability of Higher Vocational English Learners" (Grant No: 2021GXJK558).

References

- [1] Department of higher education. *Ministry of education, requirements for College English teaching*, Shanghai Foreign Language Education Press, 2007.
- [2] Wang L., Cui C., *Implementing the "business English undergraduate teaching guide" and promoting business English talent training*, Foreign language circles, 2020.vol.3, p.5-11.
- [3] Zhu Y., *Current situation and Countermeasures of English autonomous learning of non English majors*, Journal of Southeast University, vol.3, p.5-11, 2011.
- [4] Guo J., *The guiding role of teachers in college students' English Autonomous Learning*, Journal of Capital Normal University, 2008.vol.3, p.122-125.
- [5] Zhang J.. *Current situation and existing problems of College English network platform*, Horizon of science and technology, 2018.vol.9, p:180-183.
- [6] Fan W., Zhu Q., *Research on the application and teaching practice of College English Online Autonomous Learning Platform*, Journal of Guangdong Communications vocational and technical college, 2018.vol.12, p76-78.
- [7] Li Y., Chen X., Xue X., Huang Y.. *Analysis of the current situation and reform of network information teaching of business English*, Education and teaching forum, 2019.vol.12, p.136-138.
- [8] Fu X., *A survey of College Students' English autonomous learning ability and problems*, Journal of Hainan University, 2006.vol.1, p.143-146.
- [9] Gao Ch., Guo J., *An investigation and Research on the current situation of College Students' Autonomous English Listening Learning*, Journal of Panzhihua University, 2011.vol.1, p.91-94.

Detecting of All Zero Blocks in HEVC for RDOQ

Nana SHAN^{a,1}, Henglu WEI^b, Wei ZHOU^c and Zhemin DUAN^c

^a*Taishan University, China*

^b*Tsinghua University, China*

^c*Northwestern Polytechnical University, China*

Abstract. There are a larger number of blocks with all zero-quantized coefficient of the transform and quantization process in HEVC. The coding time of transform and quantization can be greatly reduced by skipping all zero coefficient blocks. All zero coefficient blocks detection algorithms for RDOQ are proposed in this paper. The stair-like thresholds are obtained by statistical analysis, which can speed up all zero coefficient blocks detection for RDOQ to improve the coding efficiency. Experimental results show that it can reduce 40% coding time with negligible loss of BDBR.

Keywords. HEVC, all zero blocks, RDO quantization, SATD

1. Introduction

High Efficiency Video Coding (HEVC) provides better compression performance than H.264/AVC [1]. Transform or quantization (T/Q) are the most time-consuming parts in HEVC [2]. The basic unit for transform and quantization in HEVC is called transform block (TB). If every quantization level (QL) in a TB is zero, this TB will be called an all zero coefficient blocks (AZB). By avoiding T/Q operation of AZB, the complexity of T/Q can be reduced greatly [3].

Many detection algorithms for all zero coefficient blocks were proposed to reduce the coding complexity. Wang use the sum of absolute difference to detect zero quantized discrete cosine transform [4]. Liu adopt the multiple reference frames to reduce the search range [5]. In [6], Hadamard transform (HT)based method was extended into HEVC, and the upper boundary of sum of absolute difference (SAD) was deduced to reduce the overall complexity. Detecting all zero blocks method was proposed in [7] to speed up the encoding process by categorizing the blocks.

AZB detection algorithm for RDOQ is proposed in this paper. The proposed algorithm is based on the conclusion that a TB is an AZB if the QL of the global maximum magnitude (MM) is zero. By statistical analysis of the MM, the stair-like thresholds for the global MM are obtained. AZBs in RDOQ can be determined by comparing the float QL (FQL) of the estimated coefficients with the stair-like thresholds.

¹Corresponding author, Nana Shan; E-mail: helensnn@hotmail.com. This work was supported by the Scientific Research Foundation of Taishan University (Grant No. Y-01-2017003).

2. Global MM Estimation

The proposed DCT coefficients estimation algorithm includes two steps. First, MM estimation and mixed granularity coefficients estimation by using zero mean Gaussian distribution model. Then, the standard deviation of DCT coefficients is adopted to estimate the global MM.

The zero mean Gaussian distribution of prediction residuals can be expressed as,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (1)$$

According to (1), the expectation of $|x|$ is,

$$E[|x|] = \int_{-\infty}^{+\infty} |x| \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = \sqrt{\frac{2}{\pi}}\sigma \quad (2)$$

where N is the size of TB, and SAD is sum of absolute difference,

$$SAD = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |x| \quad (3)$$

And we get the global MM:

$$MM_{est}[SAD] = \gamma_N \sqrt{\frac{\pi}{2}} \frac{SAD}{N^2} \cdot \max_{u,v \in [0,N-1]} \left(\sqrt{\left[DRD^T \right]_{u,u} \left[DRD^T \right]_{v,v}} \right) \quad (4)$$

Table 1 shows the probability of the global MM in the selected low frequency coefficients of HEVC standard test sequences range from Class A to Class F. Most of global MMs are located in the selected low frequency positions.

Table 1. The probability of the global MM in the selected low frequency coefficients (%)

Class	8×8 TB	16×16 TB	32×32 TB
Class A	99.16	98.14	93.12
Class B	92.66	89.95	81.91
Class C	91.77	85.36	81.44
Class D	90.37	81.33	72.80
Class E	97.40	91.82	82.29
Class F	93.85	89.75	87.96
Average	94.20	89.39	83.25

3. AZB detection in HEVC

A TB is an AZB when the QL of the global MM is zero. Rather than estimate the RD-cost, a new method is proposed in this paper. Then, the next step is to decide the QL of the global MM.

As diagonal QL scan pattern is used for inter prediction coding unit in HEVC, the QL of the global MM is related to the position of coefficients matrix. To explore the relationship between zero QL of the global MM and its location, all FQLs of the global MM less than 1.0 have been collected. The statistical result of (0,0) position of 4×4 blocks, which means the global MM is located in the (0,0) position of 4×4 block, is shown in Figure 1 (a). The green line is the number of the FQL quantized to zero QL, and the red line is the number of the FQL quantized to non-zero QL. It can be seen from

the figure that when the FQL is less than 0.5, the corresponding QL is zero, as the red line equals to zero.

FQL can be used as the threshold to distinct zero QL and non-zero QL of the global MM. Figure 1 (b) shows the relationship between FQL and the ratio of zero QL by using the curve of $p/(p+q)$, where p and q are the number of FQLs quantized to zero QL and non-zero QL respectively. The value of $p/(p+q)$ represents the ratio of FQL quantized to zero QL. It can be seen from figure 1 (b) that the probability of the global MM quantized to zero decreases with the increasing of FQL. When $p/(p+q)$ equals to 0.9, the value of FQL is selected as the threshold between zero QL and non-zero QL, which is 0.61. If the global MM is located in (0, 0) position of a 4×4 block, and its FQL is less than 0.61, the corresponding QL is considered as zero.

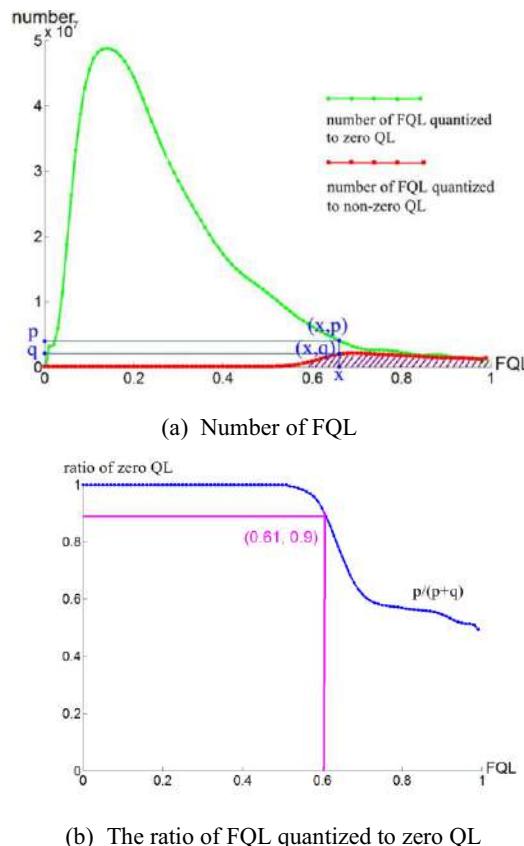


Figure 1. The collection result of the FQLs in (0, 0) position of 4×4 block.

Thresholds for all positions in various TB sizes are shown in figure 2 (b) ~ (e). The threshold is related to the frequency. The main reason is that high frequency coefficients tend to consume more coding bits and they are more likely to be quantized to zero by RDOQ. The above method to detect AZB for DCT can be directly extended to transform skip mode. Thresholds for transform skip mode are shown in figure 2 (a). If the FQL of

every prediction residual is less than the corresponding threshold, the TB is considered as an AZB in transform skip mode.

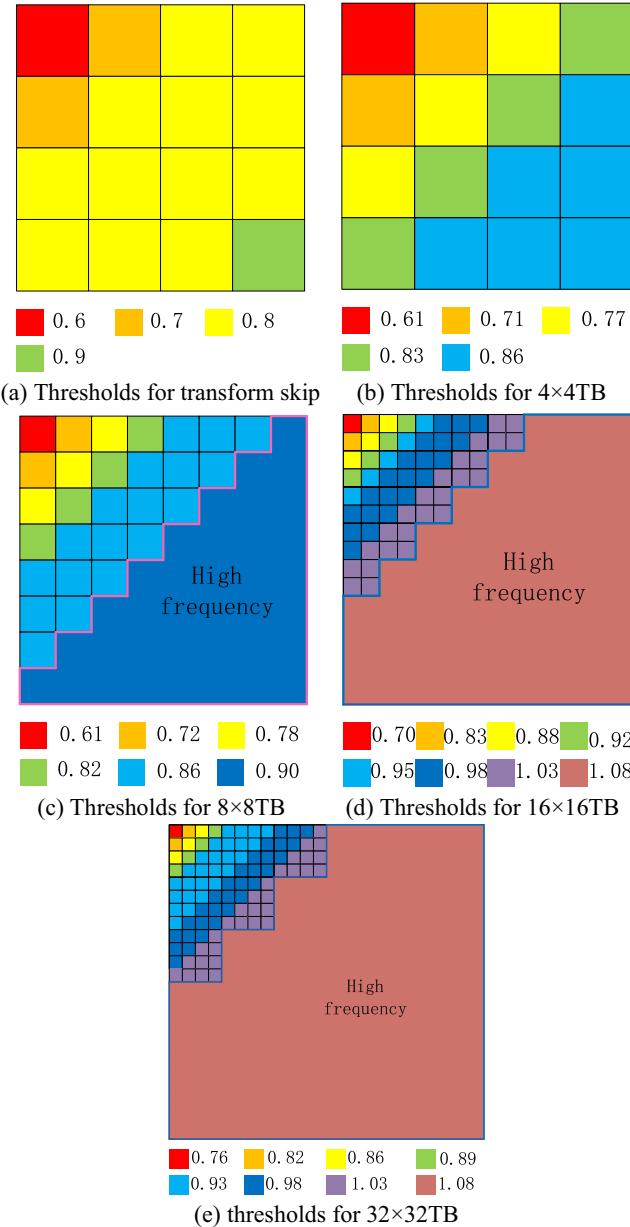


Figure 2. Thresholds for AZB detection in RDOQ

The method to pre-detect AZBs and non-AZBs by the estimated global MM can be used in RDOQ. The flowchart of AZB detection for DCT in RDOQ is shown in figure 3. The thresholds to detect zero QL in RDOQ vary with the position of coefficients. Firstly,

the global MM is estimated to pre-detect AZB and non-AZB. Then, low frequency coefficients are used. Lastly, the MM of high frequency coefficients is estimated to determine AZB and non-AZB.

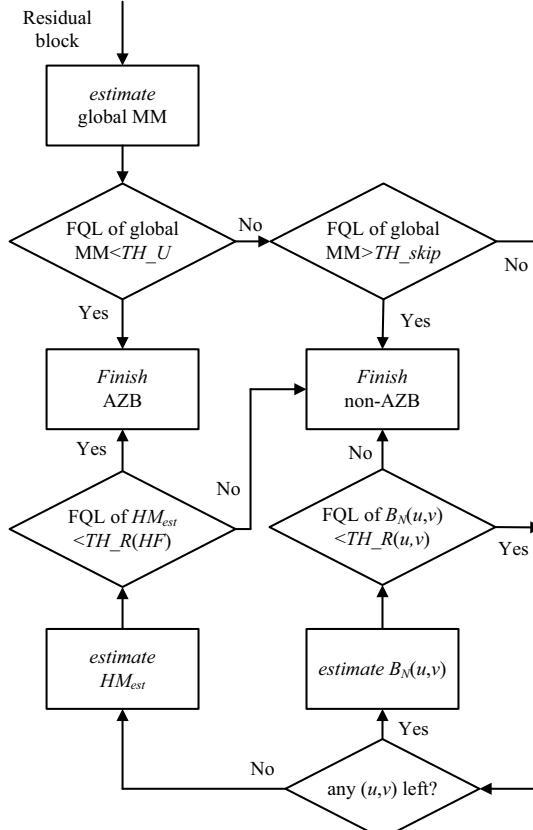


Figure 3. The flowchart of AZB detection in RDOQ

4. Experimental results

The accuracy of AZB detection process can be measured by false acceptance ration (FAR) and false rejection ratio (FRR). They can be calculated as equation (5),

$$\begin{aligned}
 FRR &= \frac{num_{mz}}{num_z} \\
 FAR &= \frac{num_{mn}}{num_n}
 \end{aligned} \tag{5}$$

The num_{mz} means the number of AZB which are classified as non-AZB incorrectly, and num_z represents the real number of AZB. While num_{mn} and num_n represent the number of non-AZB which is misclassified as AZB and the real number of non-AZB. FRR means the ratio of AZB which are classified as non-AZB incorrectly, while FAR means the ratio of non-AZB misclassified as AZB. We want to detect AZB as much as

possible, which means that the FRR and FAR will be very low. In other words, higher detection efficiency means lower FRR and FAR.

The results of the proposed algorithm are shown as Table 2. The average FRR of 4×4, 8×8, 16×16 and 32×32 TBs are 8.2%, 20.0%, 21.0% and 39.3% respectively. The average FAR of 4×4, 8×8, 16×16 and 32×32 TBs are 2.0%, 1.7%, 3.9% and 2.1% respectively. From Table 2, the proposed method is better than Lee [8] and Fan [9].

Table 2. The Results for FRR and FAR

Sequence	Sequence	Lee [8]		Fan [9]		Proposed	
		FRR (%)	FAR (%)	FRR (%)	FAR (%)	FRR (%)	FAR (%)
480P	4×4	9.7	2.0	3.7	7.8	6.65	2.83
	8×8	22.5	1.6	10.0	16.5	16.48	2.28
	16×16	21.2	9.4	15.4	16.0	17.73	5.07
	32×32	52.5	2.6	28.3	12.5	34.15	2.77
1080P	4×4	8.9	2.2	2.7	9.9	9.77	1.24
	8×8	22.5	2.9	16.3	19.0	23.53	1.02
	16×16	13.7	21.6	11.4	20.0	24.19	2.70
	32×32	40.1	8.4	30.1	16.3	44.42	1.35
Average	4×4	9.3	2.1	3.2	8.9	8.2	2.0
	8×8	22.5	2.3	13.2	17.8	20.0	1.7
	16×16	17.5	15.5	13.4	18.0	21.0	3.9
	32×32	46.3	5.5	29.2	14.4	39.3	2.1

The following experimental results are based on HEVC test model version 13.0 (HM 13.0). The standard encoder_lowdelay_main.cfg configuration file is used. 21 typical test sequences are selected from Class A to Class F, with the quantization parameters (QP) set to 22, 27, 32, and 37 respectively. Comparing with the performance of Lee [8] and Fan [9], the change of BDBR [10] and coding time are shown on Table 3. As Table 3 shows the average time saving of the proposed method is 40%, while the number of Lee [8] and Fan [9] is 34% and 43%. For BDBR, the proposed algorithm is better than Lee [8] and Fan [9].

Table 3. The Results for RDOQ AZB Detection

Sequence	Lee [8]		Fan [9]		Proposed	
	BDBR (%)	TS (%)	BDBR (%)	TS (%)	BDBR (%)	TS (%)
PeopleOnStreet	0.60	22	0.61	25	0.11	35
Traffic	0.57	41	0.94	50	0.20	39
BasketballDrive	0.47	32	0.76	38	0.31	34
BQTerrace	0.44	43	0.70	54	0.03	44
Cactus	0.31	34	0.90	44	-0.09	35
Kimono	-0.02	11	0.02	10	-0.11	34
ParkScene	0.55	40	0.89	46	0.27	39
BasketballDrill	0.24	34	0.62	48	0.00	33
BQMall	0.75	37	1.00	53	0.22	33
PartyScene	0.83	34	0.80	41	0.10	26
RaceHorsesC	0.62	20	0.43	34	-0.07	21
BasketballPass	0.86	36	0.80	45	0.13	31
BlowingBubbles	0.68	29	0.79	38	0.20	29
BQSquare	0.80	38	0.48	50	0.22	36
RaceHorsesD	0.41	18	0.34	20	0.19	20
FourPeople	0.30	53	0.36	64	0.16	59
Johnny	0.57	53	0.46	65	0.48	63
KristenAndSara	0.26	48	0.33	58	0.10	61

ChinaSpeed	0.46	23	0.46	30	0.08	31
SlideEditing	0.41	57	0.20	69	0.14	70
SlideShow	0.42	21	1.01	19	0.25	58
Average	0.50	34	0.61	43	0.14	40

5. Conclusion

An AZB detection algorithm for RDOQ is proposed in this paper. The algorithm adopts a global MM strategy based on low and high frequency coefficients to estimate transform coefficients. The proposed algorithms are tested on HEVC test model to compare the results with other methods. Experiment results show that the proposed algorithm can efficiently reduce the computation complexity while keeping nearly the same RD performance with the original algorithm in HEVC.

References

- [1] Ohm, J., Sullivan, G. J., Schwarz, H., Tan, T. K., & Wiegand, T. Comparison of the coding efficiency of video coding standards—including high efficiency video coding (hevc). *Circuits & Systems for Video Technology IEEE Transactions on*, 22(12), 2012, 1669 - 1684.
- [2] Bossen, F., Bross, B., Suhring, K., & Flynn, D. Hevc complexity and implementation analysis. *Circuits & Systems for Video Technology IEEE Transactions on*, 22(12), 2012, 1685 - 1696.
- [3] Jung, S. W., Baek, S. J., Park, C. S., & Ko, S. J. Fast mode decision using all-zero block detection for fidelity and spatial scalable video coding. *IEEE Transactions on Circuits & Systems for Video Technology*, 20(2), 2010, 201-206.
- [4] Wang, H., & Kwong, S. Prediction of zero quantized dct coefficients in h.264/avc using hadamard transformed information. *IEEE Transactions on Circuits & Systems for Video Technology*, 18(4), 2008, 510 - 515.
- [5] Liu, Z., Li, L., Song, Y., Li, S., Goto, S., & Ikenaga, T. Motion feature and hadamard coefficient-based fast multiple reference frame motion estimation for h.264. *IEEE Transactions on Circuits & Systems for Video Technology*, 18(5), 2008, 620-632.
- [6] Sole, J., Joshi, R., Nguyen, J., Karczewicz, M., & Clare, G., et al. Transform coefficient coding in hevc. *Circuits & Systems for Video Technology IEEE Transactions on*, 22(12), 2012, 1765-1777.
- [7] Wei H , Wei Z , Zhang X , et al. All zero block detection for HEVC based on the quantization level of the maximum transform coefficient. *Multimedia Tools and Applications*, 2018, 78.
- [8] Lee, K., Lee, H. J., Kim, J., & Choi, Y. A novel algorithm for zero block detection in high efficiency video coding. *IEEE Journal of Selected Topics in Signal Processing*, 7(7), 2013, 1124-1134.
- [9] Fan, H., Wang, R., Ding, L., Xie, X., Jia, H., & Gao, W. Hybrid Zero Block Detection for High Efficiency Video Coding. *IEEE Transactions on Multimedia*, 18(3), 2016, 537-543.
- [10] Bjontegaard, G. Calculation of average psnr difference between rd-curves. ITU-T VCEG-M33, April, 2001.

Visible Light Wireless Data Center Links with Distinct Beam Configurations

Jupeng DING^{a,1}, Chih-Lin I^b and Lili WANG^c

^a*Key Laboratory of Signal Detection and Processing in Xinjiang Uygur Autonomous Region, School of Information Science and Engineering Xinjiang University, Urumqi, China*

^b*China Mobile Research Institute, Beijing, China*

^c*School of Information and Electrical Engineering, Ludong University, Yantai, China*

Abstract. Visible light communication (VLC) is being explored as one promising approach to enable wireless data centers (WDC). Up to now, the visible light wireless data center links are still limited to the conventional Lambertian beam paradigm. The potential coverage gain relevant to the optical beam space is waiting for sufficient investigation. For addressing this issue, in this paper, the dynamic optical beam based WDC coverage enhancement scheme is introduced, and for each transmitter, the best candidate asymmetrical optical beam is selected to load the data signal. Numerical evaluation shows that, compared with the conventional static beam configuration, up to 6.76 dB peak signal to noise ratio (SNR) gain and 4.46 dB average SNR gain could be provided by the proposed dynamic beam scheme. Moreover, this SNR dynamic range is reduced to 36.65 dB while the counterpart of the static non-Lambertian beam configuration is up to 44.78 dB.

Keywords. Wireless data center, visible light communications, beam effect, non-Lambertian optical beams, wireless link characteristics

1. Introduction

Computation-intensive applications are pushing existing access networks and data center networks (DCN) to their performance limits. Conventional DCN with wired links and finite network interfaces, limit the possible topologies and could result into development and design issues relevant to heat dissipation, power consumption and maintenance [1-4]. Simultaneously, the increasingly emerging disadvantages of wired DCN include overwhelming cabling cost and cost, flexibility lack to tackle the traffic outbursts [5-8]. Thanks to the scalability, flexibility and energy efficiency of emerging visible light communication (VLC) techniques, it has recently attracted attention and could be employed in future-oriented wireless data centers (WDC) to address the cabling complexity, capital expenditures and other inherent restrictions of wired DCN [5-8].

To date, impressive potential performance has been identified for VLC-based WDC via angle diversity receiver, imaging receiver in typical data center scenario [2].

¹ Corresponding Author, Jupeng Ding, E-mail: jupeng7778@163.com.

Moreover, one wavelength division multiple access scheme is also proposed to enhance the transmission capacity of the visible light WDC links [1]. Nevertheless, all above visible light link configurations remain constrained in the sufficiently discussed Lambertian optical source beam pattern [4-10]. Actually, the potential design and optimization exploration via distinct non-Lambertian beams is still absent in the investigation of visible light WDC links.

Based on the above review, the typical non-Lambertian beams are adopted to construct the WDC VLC transmitter. The basic methodology of this design is to dynamically select non-Lambertian beam in order to provide more focused transmission performance. This scheme smartly utilize the radiation pattern of the commercially available LED optical sources without inducing the increase of the whole emitted optical power and the available optical transmitter location resource.

In this paper, the visible light WDC links with static and dynamic beam configuration are presented in Section II. And link performance metric are discussed in Section III. Numerical evaluation is presented in Section IV. Finally, the Section V concludes this paper.

2. Visible Light Wireless Data Center Link Design

To a large extent, the VLC link gain and coverage characteristic are influenced by the optical beam pattern of light emitting diodes (LED) in the involved transmitters. As a matter of fact, the distinct beam patterns objectively provide one novel design space for visible light WDC link performance.

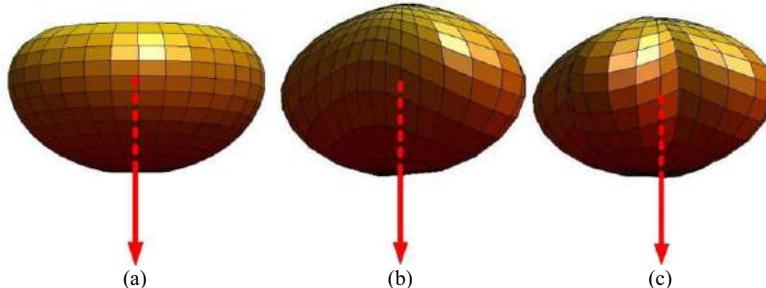


Figure 1. 3D radiation pattern of (a) conventional Lambertian optical beam, (b) typical non-Lambertian optical beam and (c) dynamic beam configuration with two candidate non-Lambertian optical beams.

2.1. Link with Non-Lambertian Beam Configuration

Unlike the generalized Lambertian beam, the spatial radiation characteristic of the non-Lambertian beam could be dependent on the azimuth and elevation angle of the emission direction simultaneously [10].

Typically, this paper discusses the non-Lambertian beam pattern of commercially available NSPW345CS Nichia LEDs. The respective radiation intensity could be given by [10]:

$$I_{\text{NSPW}}(\phi, \alpha + \Delta\alpha) = \sum_{i=1}^2 g1_i \exp \left[-(\ln 2)(|\phi| - g2_i)^2 \left(\frac{\cos^2(\alpha + \Delta\alpha)}{(g3_i)^2} + \frac{\sin^2(\alpha + \Delta\alpha)}{(g4_i)^2} \right) \right] \quad (1)$$

where α is the azimuth angle within the source plane. Specifically, the values of coefficients in this expression are: $g1_1 = 0.13$, $g2_1 = 45^\circ$, $g3_1 = g4_1 = 18^\circ$, $g1_2 = 1$, $g2_2 = 0$, $g3_2 = 38^\circ$ and $g4_2 = 22^\circ$. From the sideview angles, $\Delta\alpha$ denotes the original azimuth offset and equal 90° in this static configuration. Fig. 1 show the 3D beam pattern of this non-Lambertian optical beam and red arrow presents the normal direction of the optical source. Accordingly, the channel gain expression must be renewed and given as:

$$H(S^{\text{NSPW}}, R) = \begin{cases} \frac{A_R I_{\text{NSPW}}(\phi, \alpha + \Delta\alpha)}{P_{\text{norm}} d_0^2} \cos \theta_0 G_{\text{of}} G_{\text{oc}}, & 0 \leq \theta_0 \leq \theta_{\text{FOV}} \\ 0, & \theta_0 > \theta_{\text{FOV}} \end{cases} \quad (2)$$

where P_{norm} denotes the power normalization factor of this non-Lambertian beam, to ensure that the emission power accumulation in all spatial directions equals 1W.

2.2. Link with Dynamic Beam Configuration

In the proposed dynamic WDC beam configuration scheme, the transmitter include two mentioned NSPW345CS Nichia candidate beams with different original azimuth offset. For the i th candidate beam, the respective radiation intensity could be given by

$$I_{\text{NSPW}}(\phi, \alpha + \Delta\alpha_k) = \sum_{i=1}^2 g1_i \exp \left[-(\ln 2)(|\phi| - g2_i)^2 \left(\frac{\cos^2(\alpha_0 + \Delta\alpha_i)}{(g3_i)^2} + \frac{\sin^2(\alpha_0 + \Delta\alpha_i)}{(g4_i)^2} \right) \right] \quad (3)$$

where $\Delta\alpha_k$ is the azimuth offset of the k th candidate beam. In this work, for clarity, the specific offset value is set as 45° and 135° to both beams, respectively.

It should be noted that, the signal is merely loaded to one selected beam, and for the left candidate beam in each transmitter, only direct current is set to drive the sources lighting, such that the fundamental luminaire function of the optical sources is not affected by the proposed dynamic configuration.

Similar to equation (2), for one receiver on the data center racks top, the visible light line of sight channel gain between the selected beam and the receiver could be given as:

$$H(S_{\text{Beam}}^{\text{NSPW}}, R, \Delta\alpha_k) = \begin{cases} \frac{A_R I_{\text{NSPW}}(\phi_0, \alpha_0 + \Delta\alpha_k)}{d_0^2} \cos \theta_0 G_{\text{of}} G_{\text{oc}}, & 0 \leq \theta_0 \leq \theta_{\text{FOV}} \\ 0, & \theta_0 \geq \theta_{\text{FOV}} \end{cases} \quad (4)$$

For clarity, Fig. 1 describe the potential 3D radiation pattern including two candidate beams.

3. Link Performance Metric

In this paper, at the receiver end, signal to noise ratio (SNR) is utilized as the link performance key metric to numerically compare the conventional and proposed dynamic configuration scheme. As for the discussed static non-Lambertian beam configuration, the SNR could be given by:

$$\text{SNR}_{\text{NSPW}} = \frac{\left(P_{\text{T}} \sum_{j=1}^N H(S_j^{\text{NSPW}}, R) r \right)^2 F_{\text{OE}}}{\delta^2} \quad (5)$$

where r is the PD responsivity, P_{T} is the emitted optical power, $H(S_j^{\text{NSPW}}, R)$ is the optical channel gain between the j th NSPW345CS Nichia non-Lambertian source S_j^{NSPW} and the receiver R , N is the amount of the visible light transmitter, F_{OE} is optical - electrical conversion faxtor, and δ^2 denotes the additive noise variance at the receiver. Accordingly, this noise variance at the receiver could be given by [9, 14]:

$$\delta^2 = 2qI_{\text{bg}}B + \frac{4K_{\text{b}}TB}{R_{\text{f}}} \quad (6)$$

where q is the electronic charge, I_{bg} is the background light current, K_{b} is Boltzmann constant, T is the absolute temperature, B is modulation bandwidth and R_{f} is the feedback resistance of transimpedance amplifier (TIA) of the visible light receiver.

As for the proposed dynamic beam configuration scheme, in each visible light transmitter, assuming the channel state information (CSI) of the receiver is perfectly known to the transmitters. The candidate optical beam that could provide the best coverage SNR, is dynamically selected to transmit the information signal. For the j th visible light transmitter, the estimation of the selected candidate beam index k could be identified as:

$$\hat{k} = \arg \max_{k \in K} \left[H(S_{jk}^{\text{NSPW}}, R, \Delta\alpha_{jk}) \right]^+ \quad (7)$$

where $\Delta\alpha_{jk}$ is the azimuth offset of the k th candidate beam in the j th transmitter, K is the amount of candidate optical beams. Such that, the SNR expression for this dynamic beam scheme can be given by:

$$\text{SNR}_{\text{dyna}} = \frac{\left(P_{\text{T}} \sum_{j=1}^N H(S_{jk}^{\text{NSPW}}, R, \Delta\alpha_{jk}) r \right)^2 F_{\text{OE}}}{\delta^2} \quad (8)$$

4. Numerical Evaluation

In this section, the numerical evaluation is made between the conventional static optical beam configuration and dynamic optical beam configuration in visible light WDC link performance. Specifically, one typical data center scenario is considered. In envisioned WDC scenario with two transmitters, the receivers are located on the top of the WDC racks, such that the height of receiver working plane equals to 2 m. In addition, the main parameters are presented in Table 1.

Table 1. The main parameters configuration

Parameters	Values	Parameters	Values
Room size (W×L×H)	$3 \times 6 \times 3 \text{ m}^3$	LED Lambertian index	10.2
Emitted power of each transmitter	14.7 W	Receiver field of view	70°
Amount of VLC transmitters	2	Height of receiving plane	2 m
Boltzmann constant	1.38×10^{-23}	Physical area of PD	1.5 cm ²
Electronic charge	1.6022×10^{-19}	Responsivity of PD	0.28 A/W

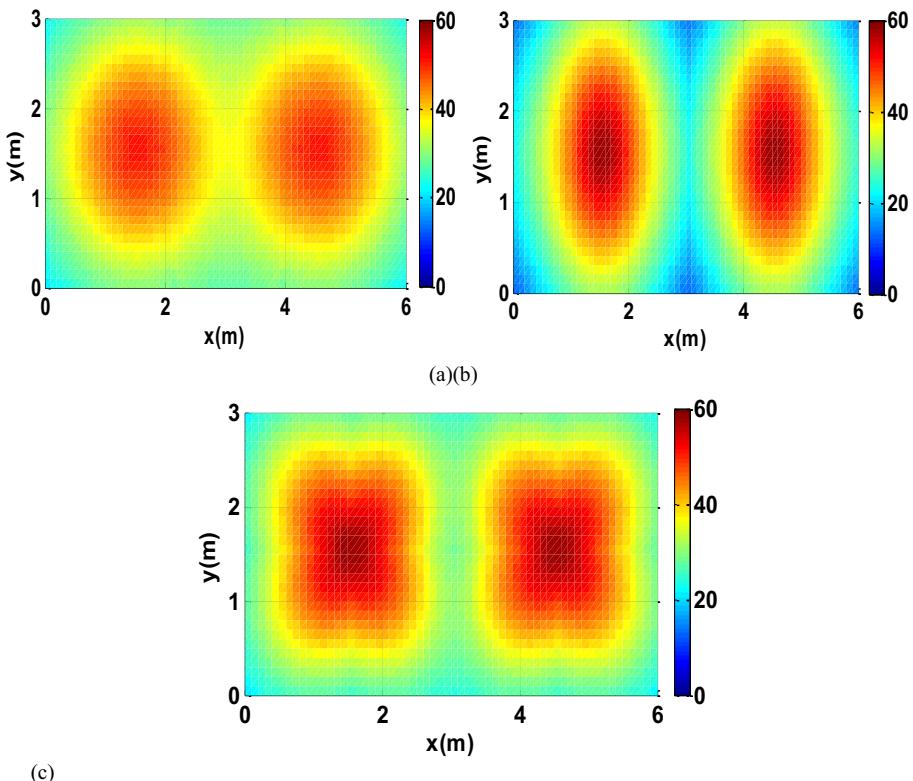


Figure 2. SNR spatial distribution in envisioned wireless data center scenario of (a) Lambertian beam configuration; (b) typical non-Lambertian beam configuration; (c) dynamic non-Lambertian beam configuration and (d) respective cumulative distribution function.

In Fig. 2, the SNR spatial distribution in envisioned visible light WDC scenario is illustrated for conventional Lambertian beam configuration, typical non-Lambertian beam configuration and dynamic non-Lambertian beam configuration, respectively. In

the first case, the SNR range between 21.56 and 51.18 dB while the average SNR is about 41.72 dB. As for the non-Lambertian beam configuration, the average SNR is increased to 45.21 dB while the dynamic range is varied to between 13.16 and 57.94 dB. Moreover, when the dynamic non-Lambertian beam configuration is adopted, the average SNR is further recovered to 46.18 dB. And the respective dynamic range lies between 21.29 and 57.94 dB. Accordingly, the cumulative distribution function of SNR spatial distribution is shown in Fig. 3.

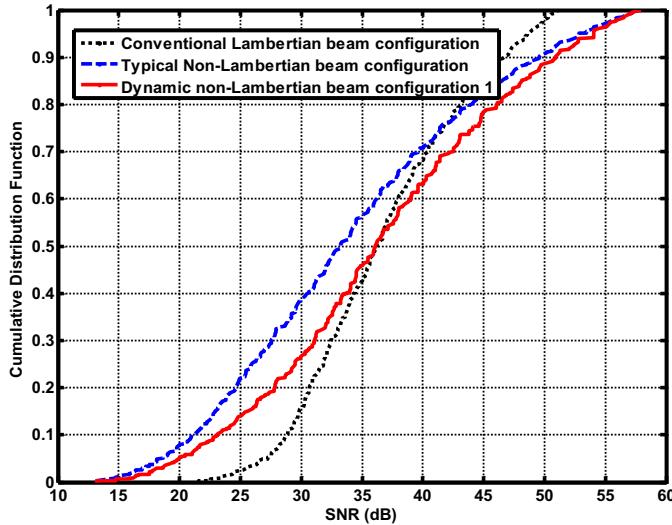


Figure3. Cumulative distribution function of SNR spatial distribution.

For the dynamic configuration, the percentage of SNR less than 30 dB is kept to about 23.74%, while the counterpart of the static Lambertian link configuration and the static non-Lambertian link configuration is 15.76% and 38.97%, respectively. Moreover, this peak SNR will be further up to 57.94 dB from the original 51.18 dB of the Lambertian link configuration, up to 6.76 dB peak SNR gain could be provided by the proposed dynamic beam scheme.

5. Conclusion

One wireless data center link design that utilize distinct optical beam configurations for downlink communications is proposed in this paper. This dynamic beam scheme is motivated by potential optical beam space, to provide coverage flexibility. The proposed dynamic scheme could take the peak SNR, fluctuation range into consideration simultaneously. For the dynamic configuration, the percentage of SNR less than 30 dB is reduced to about 23.74%, while the counterpart of the static non-Lambertian link configuration is more than 38.97%. Moreover, this peak SNR will be further up to 57.94 dB from the original 51.18 dB of the Lambertian link configuration.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grants No. 62061043), Tianshan Cedar Project of Xinjiang Uygur Autonomous Region (Grants No. 2020XS27), the Natural Science Foundation of the Xinjiang Uygur Autonomous Region (Grants No. 2019D01C020) and High-level Talents Introduction Project in Autonomous Region (Grants No. 042419004).

References

- [1] O. Z. Alsulami, S. O. M. Saeed, S. H. Mohamed, T. E. H. El-Gorashi, M. T. Alresheedi and J. M. H. Elmirghani, "Data Centre Optical Wireless Downlink with WDM and Multi-Access Point Support," 2020 22nd International Conference on Transparent Optical Networks (ICTON), 2020, pp. 1-5.
- [2] O. Z. Alsulami, M. O. I. Musa, M. T. Alresheedi and J. M. H. Elmirghani, "Visible Light Optical Data Centre Links," 2019 21st International Conference on Transparent Optical Networks (ICTON), 2019, pp. 1-5.
- [3] M. Kavehrad, "Reconfigurable optical wireless applications in data centers," 2016 IEEE Photonics Society Summer Topical Meeting Series (SUM), 2016, pp. 6-7.
- [4] Kupferman, J.S. Arnon. "Energy saving for data centers using spatial multichannel optical wireless communication." *Journal of Communications & Information Networks*, vol. 2, no. 2, pp. 88-99, 2017.
- [5] Abdulkadir C, Basem S, and Mohamed-Slim Alouini "Optical wireless data center networks: potentials, limitations, and prospects", *Proc. SPIE 10945, Broadband Access Communication Technologies XIII*,2019, pp.109450L.
- [6] Shlomi A "Review of optical wireless communications for data centers ", *Proc. SPIE 10437, Advanced Free-Space Optical Communication Techniques and Applications III*, 2017, pp.1043703.
- [7] E. CalvaneseStrinati et al., "6G: the next frontier: from holographic messaging to artificial intelligence using subterahertz and visible light communication," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 42-50, Sept. 2019.
- [8] T. Komine, M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100-107, 2004.
- [9] J. Ding, Z. Xu and L. Hanzo, "Accuracy of the Point-Source Model of a Multi-LED Array in High-Speed Visible Light Communication Channel Characterization," in *IEEE Photonics Journal*, vol. 7, no. 4, pp. 1-14, Aug. 2015.
- [10] I. Moreno, C. Sun, "Modeling the radiation pattern of LEDs," *Optics Express*, vol. 16, no. 3, pp. 1808-1819, 2008.

A Study on Fault Tolerance Technology of Flight Control Computer for Unmanned Aerial Vehicle

Chen ZHANG^{a,1} and Jihui PAN^b

^a*Major Laboratory of Space Physics, Beijing, 100076, China*

^b*Northwestern Polytechnical University, Xian, Shaanxi, 710072, China*

Abstract: According to the reliability requirement of the Flight Control Computer for Unmanned Aerial Vehicle (UAV), a design scheme is proposed to ensure its reliability by using tri-redundancy technology. Further, by selecting appropriate redundant mode and the architecture model of the triple redundant flight control computer is established in this paper. The multi-channel security level method can give full play to the error tolerance ability of the system and improve the fault tolerance performance of the aircraft. After an extensive analysis and study of the structure of each module, the hardware circuit and software flow chart of the key technologies, such as redundancy strategy and synchronization method are suggested. A channel selection method based on channel security level is proposed. Combined with the comparison technology between channels, the selection of the optimal safe channel is realized.

Keywords: Flight control system; synchronization; security level

1. Introduction

Flight Control Computer is one of the core components of UAV, and the reliability of its design directly affects the flight safety of the aircraft [1,2]. Redundancy technology however, can effectively improve the reliability and fault tolerance of the flight control computer system. The number of redundancy has a direct impact on the reliability, it is obvious that the more the number of redundancy, the higher is the system mission reliability. But increasing the number of redundancy will in turn, reduce the basic reliability, increase the volume, weight and cost, so the number of redundancy should be balanced according to the target of product development and constraints. Therefore, it is necessary to start with the design of the flight control computer architecture and adopt redundancy technology to fundamentally improve the fault tolerance and the survivability of the system [3-5].

In this paper, a tri-redundancy computer is designed and a channel selection method based on the channel security level is proposed. The hardware circuit and software flow chart of the key technologies such as redundancy strategy and synchronization method are given.

¹ Corresponding Author: Zhang Chen, Major Laboratory of Space Physics, Beijing, 100076, China; E-mail: 925819259@qq.com.

2. Flight Control Computer Architecture

Commonly used redundancy architectures include n-mode redundant structures and comparative monitoring structures. Most flight control computers of large manned aircraft use the above two redundant structures, or a combination of two redundant structures [6,7].

The advantage of n-mode redundancy structure is the majority voting structure. When the redundancy is more than 3, the system is reliable, but when it is less than 3, the system can only select channels through self-monitoring, which reduces the reliability [8,9].

The advantage of the redundant monitoring structure is its high reliability. It selects channels by comparing two computers in each channel, but the disadvantage is that the hardware is rather more complex [10-13].

A tri-redundancy computer scheme is presented in this paper. By comparing the security levels of multiple channels, the selection of the optimal security channel is realized, the error tolerance ability of the system is brought into full play, and the fault tolerance performance of the aircraft is improved.

The flight control computer uses a tri-redundancy structure, the internal structure of which is shown in Figure 1. The flight control computer has three identical channels, each of which includes a central processing unit (CPU), input and output interface (DIO), analog quantity processing module (AIO), serial port transceiver module (Sio), power supply module (PS) and other functional modules. The flight control computer only has one set of I/O interface, and the I/O interface is pseudo triplex. In the process of flight control computer working, 3 channels receive external input data at the same time, and one of them is chosen as the control computer by voting mechanism.

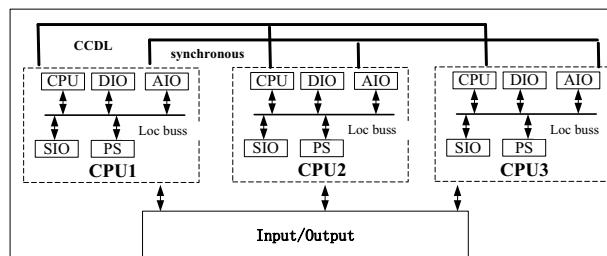


Figure 1 Tri-redundancy flight control architecture

3. Channel selection method based on security level

In order to improve the reliability and the fault tolerance of the system, a channel selection method based on channel security level is proposed. Based on the comparative technology between the channels, the selection of the optimal safe channel is realized [6].

First, the flight parameters of the channel are obtained and synthesized to obtain the safety level of the flight parameters of the channel. Then, they are combined with the internal parameters of the channel, the security level of the channel is synthesized, and the security level of other channels of the redundancy computer is obtained through

inter channel transmission. On comparing the security levels of the multiple channels of redundancy computer, the channel effectiveness of redundancy computer is obtained and one of the computers is selected as the control computer through the voting arbitration circuit.

3.1. Safety level of channel flight parameters

The input parameters of channel are divided into three parameter sets: the parameter set for the safe execution of mission, the parameter set for the safe flight and guidance of aircraft, and the parameter set for the stability of flight attitude.

The first parameter set is to ensure the flight attitude stability of the aircraft, including pitch angle, tilt angle, heading angle, pitch angle rate, tilt angle rate, heading angle rate, barometric altitude, indicated airspeed and effective parameters of steering gear.

The second parameter set is the parameter set to ensure the safe flight and guidance of the aircraft. In addition to the first parameter set, it also includes the position parameters of the aircraft. The second parameter set includes the pitch angle, tilt angle, heading angle, pitch angle rate, tilt angle rate, heading angle rate, air pressure altitude, indicated airspeed, steering gear communication parameters and position parameters.

The third parameter set ensures the safe execution of the mission by the aircraft. In addition to the second parameter set, it also includes the parameters of the aircraft that affect the mission execution. The third parameter set includes pitch angle, tilt angle, heading angle, pitch angle rate, tilt angle rate, heading angle rate, air pressure altitude, indicated airspeed, steering gear communication parameters, position parameters and mission load parameters.

The safety level of channel flight parameters is integrated, and the external input parameters are classified into three parameter sets according to the critical difference of parameter. They include the parameter set of flight attitude stability, the parameter set of aircraft safe flight and guidance, and the parameter set of aircraft safe mission execution. The status is obtained according to the external input parameters of the channel, and synthesizes the flight parameter safety level of the channel. It is characterized in that the safety level of the channel flight parameters is divided into four levels: level 0: mission parameter safety, level 1: flight parameter safety, level 2: flight parameter degradation safety, and level 3: flight parameter insecurity.

Level 0: mission parameter safety, when the parameter is set for the aircraft to safely perform the mission, including all the parameters that can safely perform the mission.

Level 1: flight parameter safety, when the parameters are set for aircraft safe flight and guidance, aircraft safe flight and guidance parameters, but not the mission load parameters.

Level 2: flight parameters are degraded safely, when parameter is set for stable flight attitude and safe flight parameters after degradation.

Level 3: flight parameters are unsafe, and there is no parameter set for stable flight attitude.

3.2. Channel internal parameters

On the other hand, in channel security level synthesis method, the channel memory, the channel power supply, the channel watchdog, and the channel internal parameters are valid, otherwise the channel internal parameters become invalid.

3.3. Comprehensive channel safety level

According to the safety level of the channel flight parameters and combined with the effectiveness of the internal parameters of the channel, the comprehensive safety level of the channel is realized.

The channel safety level comprehensive method is characterized when the channel safety level is divided into four levels: level 0: channel mission safety, level 1: channel flight safety, level 2: channel flight degradation safety, and level 3: channel flight insecurity.

In the channel safety level synthesis method, when the internal parameters are valid, the channel safety level is equal to the channel flight parameter safety level; otherwise, the channel security level equals to level 3.

3.4. Channel selection circuit design

Channel flight parameter safety level Ex_Pa_Level is one of the above four values, Ex_Pa_Level consists of Bit1, bit0 and 2 bits, where, 00 represents level 0, 01 represents level 1, 10 represents Level 2 and 11 represents Level 3.

The validity of the channel internal parameters depends on the validity of channel CPU, channel memory, channel power supply and channel watchdog. When Cpu_V is CPU effectiveness, Ram_V is memory validity, $Power_V$ is power supply effectiveness, $Wachdog_V$ is the internal parameter inside the channel when the watchdog validity is valid at the same time, and $Inside_Pa_V$ is valid. Otherwise, $Inside_Pa_V$ is invalid (the signal is valid at low level and invalid at high level). 4 or gates are used to realize the synthesis of internal parameters of the channel.

Safety level according to the channel flight parameters Ex_Pa_Level , is combined with the internal parameter validity of the $Inside_Pa_V$ to achieve comprehensive channel security level. Internal parameter $Inside_Pa_V$ is valid when the channel safety level is Chx_Pa_Level and is equal to the safety level of channel flight parameters; otherwise, the channel security level is equal to level 4. Chx_Pa_Level consists of 2 bits where, 00 represents level 0, 01 represents level 1, 10 represents Level 2 and 11 represents Level 3.

The channel effectiveness comparison comprehensive method is characterized when the input has three channel security levels Chx_Pa_Level and output has three channel validity Chx_V . If the number of safety levels of the current channel is not greater than the other two channels, the channel is valid, otherwise the channel is invalid (Chx_V signal high level is valid and low level is invalid).

According to the priority and effectiveness of the three channels, the channel with the best health is voted as the control channel.

The three channel control right arbitration circuit is characterized in that the input is three channel validity Chx_En . The output is channel enable Chx_En (high level of Chx_En signal is valid and low level is invalid).

The channel with the best health is used as the control method, which is characterized in that if channel 1 is effective, channel 1 enable output is effective; if channel 1 fails and channel 2 is valid, channel 2 enable output is valid; if channel 1 fails, channel 2 fails and channel 3 is valid, channel 3 enable output is valid; if all three channels fail, the enable output of channel 1 is valid.

The logic diagram is shown in Figure 2.

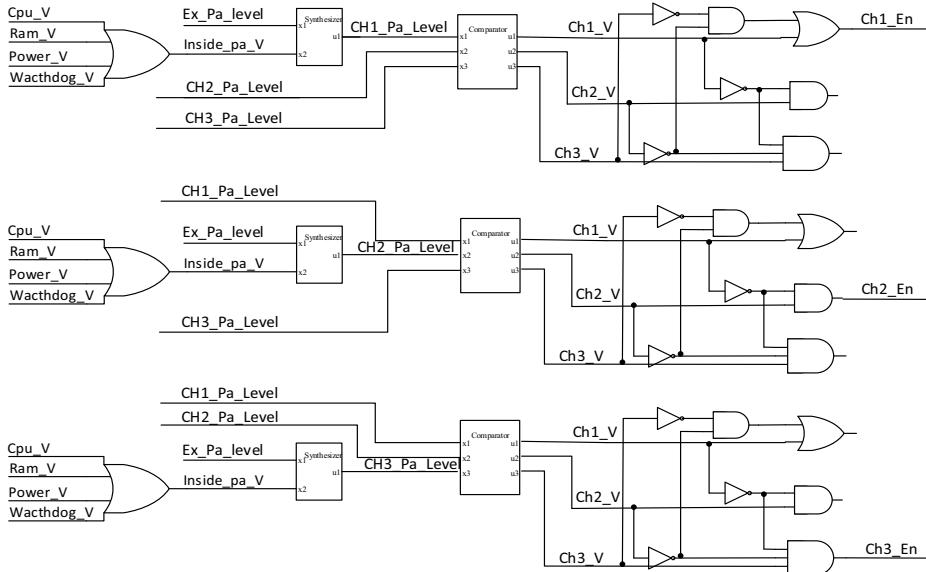


Figure 2 Channel selection circuit design graph

4. Synchronization method of the three channels

In order to meet the requirement of seamless switching during flight, the three channels must work synchronously. Synchronization means that the 3-channel periodic tasks are carried out in the same cycle, and the same tasks are executed at the same time as possible among the 3-channel periodic tasks. If it cannot work synchronously, the periodic tasks between the three channels cannot be completed in the same beat, which means that both the sampling and controlling of the three computers cannot be carried out simultaneously, and the seamless switching between the three channels cannot be realized, which could cause the craft to become unstable in an instant.

4.1. Synchronization method of the three channels

The main causes of the three-channel synchronization include the difference of channel start-up time and the accumulation of crystal oscillator errors. The difference of start-up time refers to the difference of start-up time of the operating system, which causes the three channels not to enter the flight control program at the same time. The accumulation of crystal oscillator errors is another reason why the three channels cannot be synchronized. The error accumulation test is carried out with two channels. After power-on, the two channels get synchronized and the periodic (period is 10ms)

pulses are sent out. From the oscillator, the clock errors accumulate and one channel is 10ms more than the other at 15 minutes. However, this error is not tolerated during the flight. Therefore, in addition to the start-up synchronization between the three channels, it is necessary to carry out periodic synchronization.

4.2. Hardware design of synchronous circuit

In order to synchronize the three channels, a simple handshake protocol is proposed to synchronize the flight control computer. As shown in figure 3, each channel has an output DO for outputting a synchronous handshake signal to the other two channels, and two inputs DI are used to receive synchronous input signals from the other two channels. The synchronization method uses handshake mode: First turn off the interrupt, then DO outputs a "logical low" synchronous handshake signal, and then inquire within a limited period whether the two DI generate "logical low" handshake signal; after the handshake is successful, open the interrupt, DO output "logic high", all 3 channels keep synchronization signal out for logic high preparation next synchronization.

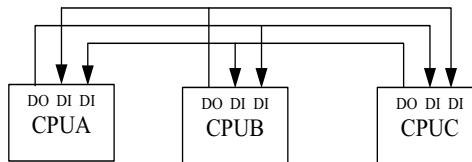


Figure 3 Schematic diagram of synchronous signal hard connection

4.3. Software design of synchronization method

Aiming at the difference of start-up time, the concept of start-up synchronization and 10ms period synchronization are put forward in the software design. The synchronization software is interspersed in the whole Flight Control Program as shown in Figure 4. The start-up synchronization which is called only once after booting, guarantees the elimination of three channel start-up time-consuming errors, and is encapsulated as a function void power on synchronize. Start-up synchronization method: After the start-up, through the DO pin, the other two processors send synchronous signal, while collecting the synchronous signal sent by the other two processors, and after collecting the synchronous signal of the two machines, switch to 10ms mission cycle. If the synchronous signals of the other two computers have not been collected after waiting for 3S, the other two computers are considered to be out of order. The remaining two processors are set to permanent failure. The 10ms cycle synchronization, which is encapsulated as a function, void period 10ms synchronize (void), is called in a 10ms interrupt service routine. The function first disables all interrupts, determines if the other two processors are available, if not, goes into stand-alone mode, and if available, goes into synchronization. Synchronous method: The DO pin sends out the synchronous signal to the other two processors, and collects the synchronous signal from the other two processors at the same time. If the synchronous signal of the other two computers is not collected within 50 micro-seconds, it can be interrupted to judge whether it cannot be synchronized for 10 consecutive times, and if it is, it can be put into single-machine working mode.

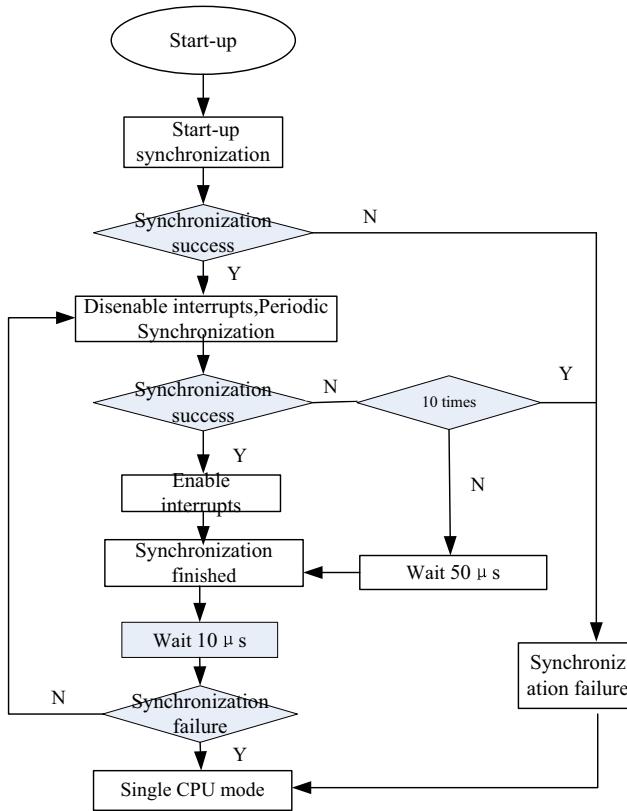


Figure 4 Synchronous Flow Chart in Flight Control Program

Figure 5 is a diagram of narrow-amplitude pulse signals sent periodically (10ms period) by three channels of the flight controller. After synchronization, the narrow-amplitude pulses can be sent out at the same time the handshake protocol completely solves the problem of synchronization.

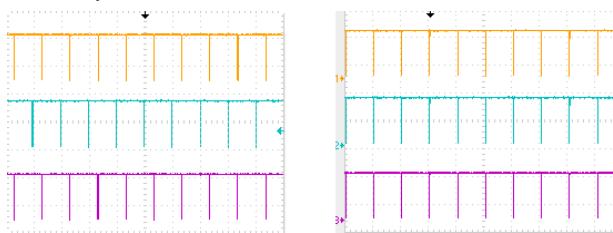


Figure 5 Asynchronous/synchronous periodic pulse graph

5. Conclusion

After hundreds of simulation tests significant results were obtained in fault tolerance and reliability. The flight control aircraft adopts the redundancy strategy under this configuration to meet the requirements of small size and low weight of the UAV, and the system reliability is guaranteed. The multi-channel security level method can give

full play to the error tolerance ability of the system and improve the fault tolerance performance of the aircraft.

References

- [1] LI Hailiang, CAI Jing, KANG Tingwei. Study on design of visual intelligent detection instrument for aviation cable fault [J]. Journal of NorthwesternPolytechnical University, 2021, 39(4): 770-775.
- [2] Huang Y, He J H, Zhou M L, et al. Researching Integrated Flight / Fire Control System of Air-to-Ground Guided Bombs. Journal of NorthwesternPolytechnical University, 2016, 34(2): 275-280.
- [3] Pan J H, Zhang S B, Zhang X L. Fault Diagnosis and Adaptive Reconfiguration for Multi -Redundancy Flight Control System [J]. Computer measurement and control, 2015. 23(48): 441-447.
- [4] ZHANG Yang, ZHOU Zhou, LI Xu. Effect of turbulence intensity and gradient of turbulence intensity on airfoil aerodynamic characteristics at low Reynolds number [J]. Journal of NorthwesternPolytechnical University, 2021, 39 (4): 721-730
- [5] Pan J H, Zhang S B, Zhang X L. Design and Realization of Treble-Redundancy Management Method of Flight Control System [J]. Journal of NorthwesternPolytechnical University, 2013, 31(5): 798-802.
- [6] Suo X J, Kang X D, Zhou Q. Application of Fault Injection Technique in Analog Signal Acquisition of Airborne Computer [J]. Computer Measurement & Control, 2016, 24(1) :304 -312.
- [7] YANG Lei, ZHANG Bainan, GUO Bin, ZUO Guang, SHI Yong, HUANG Zhen. Concept definition of new-generation multi-purpose manned spacecraft [J]. Acta Aeronautica et Astronautica Sinica. 2015 ,36(3): 703-713.
- [8] LEE J, KWON D, KIM N, et al. PHM-based wiring system damage estimation for near zero downtime in manufacturing facilities [J]. Reliability Engineering and System Safety, 2019, 184: 213-218.
- [9] WANG Danyang, TANG Jianjun, CHEN Ou, et al. Aerospace-cable fault location technology research based on time domain reflectometry (TDR) [J]. Aeronautical Manufacturing Technology, 2019, 62(Z2): 84-88, 96.
- [10] LI Lu. Simulation study on fault diagnosis of marine power cable [D]. Dalian: Dalian University of Technology, 2017.
- [11] CHUNG C C, LIN C P. A comprehensive framework of TDR landslide monitoring and early warning substantiated by field examples [J]. Engineering Geology, 2019, 262(28): 1-11.
- [12] LIU Yandong. Analysis of airworthiness requirement in EWIS design [J]. Journal of Shenyang Aerospace University, 2019, 30(4): 23-27.
- [13] MEI Z P, LI Q, WEN J Q. Research on optimization of wiring paths in airplane harness process [C]/2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, 2019: 485-488.

Ultra-HD Video Streaming in 5G Fixed Wireless Access Bottlenecks

Koffka KHAN¹, Wayne GOODRIDGE

Department of Computing and Information Technology (DCIT), The University of the West Indies, St. Augustine, Trinidad and Tobago, West Indies

Abstract. 5G Fixed Wireless Access (FWA) is an enabling technology in intelligent systems (IS) that may provide Ultra-HD (UHD) video streaming services with high Quality of Experience (QoE) in a small business use case setting. However, UHD streaming over 5G FWA is difficult in terms of latency and dependability due to numerous network factors. Due to this there may be multiple video players competing for network resources when streaming a UHD video. To date there has been very little work of 5G 'last mile access' streaming over bottleneck FWA Local Area Networks (LANs) under congested network conditions. The bottleneck link is the 5G FWA gateway. In these networks viewers may get sub-optimal QoE. Adaptive bitrate (ABR) algorithms are used to select the near optimal bitrates during a streaming session. To obtain the QoE of viewers in 5G FWA bottleneck networks we study the performance of four DASH-based adaptive video streaming algorithms (MPC, BOLA, Oboe and Pensive). BOLA performs the best and Pensive the worst. However, BOLA's overall performance is sub-optimal. This work supports the need for developing new ABR algorithms for the 5G FWA environment.

Keywords. 5G, FWA, Ultra-HD, video, streaming, QoE, last mile access, gateway, bottleneck

1. Introduction

Fixed Wireless Access (FWA) is one of the first planned technologies that could transform the digital environment during the early stages of 5G deployment. FWA delivers Internet services to end consumers with lower infrastructure costs by using both dedicated fixed networks and shared mobile networks (e.g., 4G and 5G networks), see Figure 1. FWA will offer innovative and more flexible wireless broadband options for homes and businesses. Many systems are considering bands at 'higher' millimeter wave frequencies due to the need for high bandwidth and new spectrum. Current 5G FWA devices can attain 150 Mbps on 5G New Radio (5G-NR) bands below 7 GHz [1]. Future 5G-NR deployments in the millimeter wave (mmWave) spectrum will allow the FWA to attain multi-gigabit per second (Gbps) rates, equivalent to super fast fiber networks. However, a clear LoS (Line of Sight) is required from the base station to consumer premises equipment (CPE) [2]. Practically this has its challenges. To address this challenge, a number

¹Corresponding Author: Lecturer in Computer Science, Department of Computing and Information Technology (DCIT), The University of the West Indies, St. Augustine, Trinidad and Tobago, West Indies; E-mail: koffka.khan@sta.uwi.edu

of options are being considered. This includes the placement of the base station and CPE in a direct LoS which gives the best coverage. Multiple input, multiple output (MIMO) and beamforming are used to overcome the high data loss incurred when the transmitter and receiver are not in a direct LoS [3].

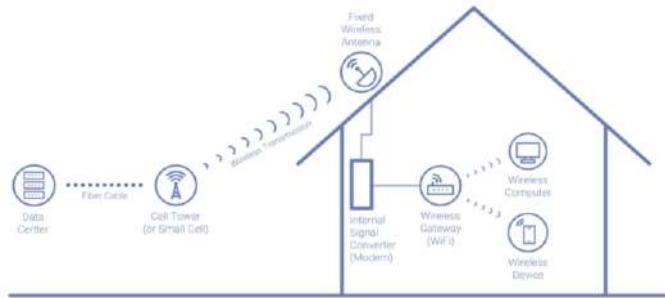


Figure 1. Fixed Wireless Access (Source: cbinsights.com).

HTTP Adaptive Streaming (HAS) is the de-facto solution for Internet video distribution and is used by prominent providers such as YouTube and Netflix. DASH's main idea is that video sequences are broken into chunks that are encoded with varied qualities (e.g., resolutions and bit rates) and provided through Hyper Text Transfer Protocol (HTTP) and Transmission Control Protocol (TCP) from web servers [4,5]. During streaming sessions, a DASH client requests the best quality for each of these chunks based on its local conditions (e.g., available bandwidth and buffer level). Given that the DASH standard does not specify the implementation details of the adaptation logic, a great deal of effort has gone into developing effective solutions. This logic can be implemented at the client [6,5], in-network [7] or server [7].

UHD has a resolution of 3840 x 2160 pixels. UHD video streaming is difficult for a variety of reasons [8]. Your connection speed must be at least 25Mbps in order to watch a UHD video stream a small business [9]. Thus, our first reason is UHD video streams consume a lot of bandwidth. Our second reason is DASH-based video streaming systems often utilize bitrate adaptation, buffer control, Quality of Experience (QoE) [10] inference in their optimization strategies which are not suited to the 5G environment [8]. Our third reason is TCP's [11] transitory behavior via its congestion control mechanism significantly underutilizes network bandwidth. Finally, mmWave 5G experiences [12] dramatic and frequent performance variations which results in highly varying network speeds [8], see Figure 2. This might possibly confound network and application layer logic, such as ABR video streaming, and result in underutilization of the carrier's channel capacity and resources.

The contributions of this paper are: (1) First time the performance of BOLA [13], MPC [14], Pensieve [15] and Oboe [16] DASH ABR algorithms are studied under a UHD DASH video streaming testbed mimicking 5G FWA. (2) Subjective QoE metrics have never been used to evaluate the performance of these streaming algorithms in this environment. (3) Experimental work involving varying number of UHD DASH players are performed and analyzed. We present our work in five Sections. MPC, BOLA, Oboe and Pensieve DASH algorithms are described in Section 2. These algorithms had superb performance in 4G video streaming. Our experimental 5G FWA Testbed is described in

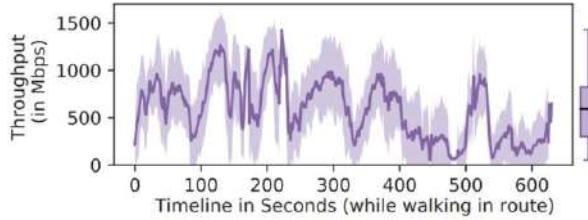


Figure 2. Variation in 5G Throughput [8].

Section 3. Objective and subjective QoE findings are given in Section 4 with a discussion on how sharing 5G DASH UHD video using ABR algorithms in a FWA environment impacts viewer QoE. Finally, we emphasize the need to develop new 5G ABR algorithms in our conclusion, Section 5.

2. DASH approaches

2.1. MPC: Model Predictive Control

MPC [14] selects bitrates that use the best QoE parameters across the projected segment range using both throughput estimations and buffer occupancy data. The characteristics that are generally evaluated include average video quality, average video quality fluctuations, rebuffer or startup delay. However, because viewers may have various priorities regarding which of these factors is more important, the weighted total of these four metrics is used to define the QoE of the video segment. In this paper we assume equal weightings for the four metrics. The MPC method makes appropriate use of both bandwidth estimation and buffer levels. MPC is resistant to prediction error because at any step it reduces the inaccuracy in prediction, see Figure 3 for the MPC adaptation workflow algorithm and [14] for accompanying details.

```

1: Initialize
2: for  $k = 1$  to  $K$  do
3:   if player is in startup phase then
4:      $\hat{C}_{[t_k, t_{k+N}]} = \text{ThroughputPred}(C_{[t_1, t_k]})$ 
5:      $[R_k, T_s] = f_{mpc}^{st}(R_{k-1}, B_k, \hat{C}_{[t_k, t_{k+N}]})$ 
6:     Start playback after  $T_s$  seconds
7:   else if playback has started then
8:      $\hat{C}_{[t_k, t_{k+N}]} = \text{ThroughputPred}(C_{[t_1, t_k]})$ 
9:      $R_k = f_{mpc}(R_{k-1}, B_k, \hat{C}_{[t_k, t_{k+N}]})$ 
10:  end if
11:  Download chunk  $k$  with bitrate  $R_k$ , wait till finished
12: end for

```

Figure 3. Video adaptation workflow using MPC [14].

2.2. BOLA: Buffer Occupancy based Lyapunov Algorithm

The BOLA [13] design is purely buffer-based. The DASH session is modeled as a stochastic optimization problem with a time-average objective over a finite horizon. Dynamic programming (DP) is used to solve it. BOLA maximizes the playback's utility as determined by a weighted sum of the average quality bitrate and average stalling time. BOLA solves the chunk quality choice problem with a Lyapunov optimization approach for each chunk request, that is whether to avoid downloading a new chunk or to download a new utility-maximizing chunk. BOLA avoids the overheads of more complicated bandwidth prediction ABR schemes and is more reliable in the face of bandwidth variations, see Figure 4 for the BOLA algorithm and [13] for accompanying details.

```

1: for  $n$  in  $[1, N]$  do
2:    $t \leftarrow \min[\text{playtime from begin, playtime to end}]$ 
3:    $t' \leftarrow \max[t/2, 3p]$ 
4:    $Q_{\max}^D \leftarrow \min[Q_{\max}, t'/p]$ 
5:    $V^B \leftarrow (Q_{\max}^D - 1)/(v_M + \gamma p)$ 
6:    $m^*[n] \leftarrow \arg \max(V^D v_m + V^D \gamma p - Q)/S_m$ 
7:   if  $m^*[n] > m^*[n - 1]$  then
8:      $r \leftarrow \text{bandwidth measured when downloading segment } (n - 1)$ 
9:      $m' \leftarrow \max m \text{ such that } S_m/p \leq \max[r, S_1/p]$ 
10:    if  $m' \geq m^*[n]$  then
11:       $m' \leftarrow m^*[n]$ 
12:    else if  $m' < m^*[n - 1]$  then
13:       $m' \leftarrow m^*[n - 1]$ 
14:    else if some utility sacrificed for fewer oscillations then
15:      pause until  $(V^D v_{m'} + V^D \gamma p - Q)/S_{m'} \geq (V^D v_{m'+1} + V^D \gamma p - Q)/S_{m'+1}$   $\triangleright$  BOLA-O
16:    else
17:       $m' \leftarrow m' + 1$   $\triangleright$  BOLA-U
18:    end if
19:     $m^*[n] \leftarrow m'$ 
20:  end if
21:  pause for  $\max[p \cdot (Q - Q_{\max}^D + 1), 0]$ 
22:  download segment  $n$  at bitrate index  $m^*[n]$ , possibly abandoning
23: end for

```

Figure 4. The BOLA Algorithm [13].

2.3. Oboe

Oboe [16] calculates the best potential parameters for a particular ABR algorithm under various network situations. At runtime, it proactively updates the settings for real-world network scenarios. Oboe takes advantage of piecewise-stationary network connections. It addresses the problem of network configuration sensitivity to changing network environments. It begins with an offline stage in which the best configuration for each network state is calculated in advance. Then, it employs an online stage. This stage monitors changes in network status throughout a session and sets the optimal pre-calculated configuration to the present (stationary) state. Oboe can also include other parameters such as session type (live and video on demand (VOD)), streaming rate thresholds and QoE measures (e.g., choice between frequent video stalls vs. UHD quality level). Oboe tweaks robustMPC [14], an MPC version developed to employ a 5-segment horizon, in our 5G trials. It employs a conservative throughput estimate, which normalizes the de-

fault throughput prediction by taking the highest estimation error over the preceding five chunks. Robust MPC efficiently optimizes the worst-case QoE. Oboe's offline and online pipeline is shown in Figure 5 with [16] for accompanying details.

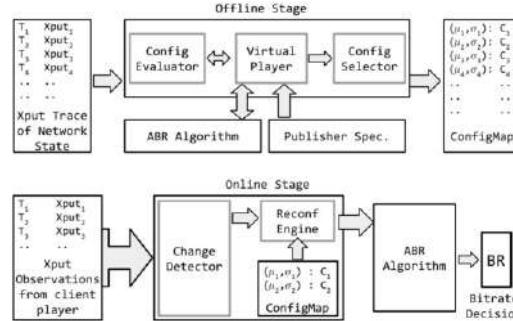


Figure 5. Logical diagram of Oboe's offline and online pipeline [16].

2.4. Pensieve

Reinforcement learning (RL) is used by Pensieve [15] to solve difficulties that video players have during DASH session. It creates a neural network model that uses receiver's video player input to pick bitrates for potential video chunks. Only through analyzing the outcomes of previous decisions can Pensieve develop the capacity to make adaptive bitrate (ABR) decisions. Pensieve constructs its policy with the actor-critical algorithm (see Figure 6 and [15] for accompanying details), a policy gradient method. By looking at the execution paths, policy gradient techniques can quantify the gradient of the projected cumulative reward. Once Pensieve has created an ABR algorithm using our 5G simulations, the model's concepts must be applied to real DASH sessions. When a client requests for individual chunks arrive at the video provider, Pensieve feeds the incoming data into its neural network model, which responds to the video client with the bitrate level to be used for the next chunk update.

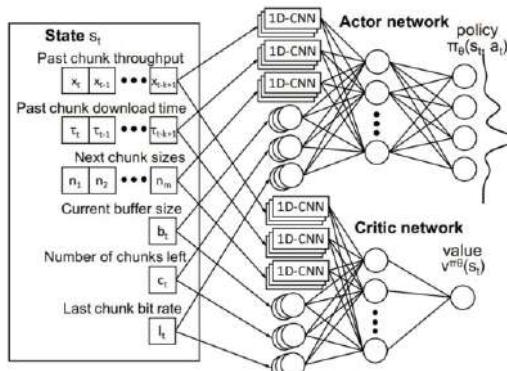


Figure 6. The Actor-Critic algorithm that Pensieve uses to generate ABR policies [15].

3. Empirical 5G FWA Testbed

3.1. 5G FWA Business viewer environments

3.1.1. 5G FWA Business Environment 1: Worker added when viewing video

All 10 workers of the business begin watching their (separate) video at the same time until it ends.

3.1.2. 5G FWA Business Environment 2: Worker removed when viewing videos

One of the workers begin to watch their video. In addition, each worker begins watching their (separate) video at 60 second intervals. Thus, all 10 workers will be watching a video during the last minute.

3.2. 5G FWA small business network

Two-second Group of Pictures (GOP) are used to create a two-minute and twenty-two-second clip in 3840x2160 (4K) with thirteen bitrates ranging from 1.8Mbps to 18Mbps and on demand DASH profiles [17]. There are three 4K bitrates, two Full HD bitrates, two HD bitrates, and six lower resolution bitrates. The two second GOP is repeated five times with different segment lengths (2, 4, 6, 10 and 20 seconds). We used 2 second chunks in experiments. The video clips (Big Buck Bunny (BBB), Sintel, and Tears of Steel (TOS)), each had a duration of more than 10 minutes.

One laptop serves as the UHD video streamer, while the remainder of the system (including wireless gateway) is hosted on a high-end PC. The testbed is set up to be able to stream UHD video. The Streamer machine runs Ubuntu 18.04.5 LTS 64-bit and comes with 32 GB of RAM, an AMD Ryzen 9 5900X 12-core, 24-Thread CPU and a 2.0 TB hard drive. In our FWA environment the laptop acts as the base station. The high-end PC runs Ubuntu 16.04.7 LTS 64-bit and comes with 64 GB of RAM, a Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz processor, and a 3.0 TB hard drive. The wireless gateway (CPE in our FWA environment) is hosted on a FreeBSD dummynet virtual machine on the PC. It is connected to the UHD players (business workers in our FWA environment) who are also hosted on the PC. Network parameters were set to follow the SPEED-5G [9] specifications.

We set the wireless gateway at 120Mbps. There are 10 workers streaming UHD video. Each video requires 18Mbps download speeds to be viewed at the highest quality. Thus, when the sixth worker joins in experiment 2's 5G FWA Business Environment 1 (see Table 1) or the entire experiment 1's 5G FWA Business Environment 1 the gateway will act as a bottleneck link. UHD players will have to compete for network resources in order to serve high quality to their users. We explore the results of this competition in our experiments, see Section 4.

Time (s)	60	120	180	240	300	360	420	480	540	600
Add Worker	1	2	3	4	5	6	7	8	9	10
UHD Bandwidth Usage (Mbps)	18	43	68	93	118	143	168	193	218	243

Table 1. Bottleneck: Bandwidth usage exceeds 120Mbps

4. Experimental results

4.1. 5G FWA Business Environment 1

Except for BOLA, all other ABR algorithms suffer from more than 2 video stalls while operating over 5G FWA, see Table 2. Video stalls are the most serious issue with video streaming over 5G. BOLA performs best across four objective QoE metrics (switches, stalls, stall duration and average utilization) followed by MPC, Oboe and Pensieve.

ABR	Switches	Stalls	Avg. Stall Duration (s)	Avg. Utilization
BOLA	75	2	0.5	23.27
MPC	87	5	1.5	18.59
Oboe	108	11	3	14.38
Pensieve	116	16	4	12.59

Table 2. QoE Objective metric performance of DASH ABR in 5G FWA Business Environment 1

4.2. 5G FWA Business Environment 1

Video stalls are still numerous in this scenario but are less than the experiment in Section 4.1, see Table 3. All players compete for bandwidth throughout the experiment in Section 4.1. However, bottleneck conditions only occur after 360 s in this experiment so there is more overall bandwidth available for players initially and up to the last minute where all workers are using the network resources. Again, BOLA performs best across the objective QoE metrics followed by MPC, Oboe and Pensieve.

ABR	Switches	Stalls	Avg. Stall Duration	Avg. Utilization
BOLA	37	1	0.3	24.13
MPC	48	3	1.2	19.83
Oboe	64	8	2.5	15.26
Pensieve	25	10	3.3	13.79

Table 3. QoE Objective metric performance of DASH ABR in 5G FWA Business Environment 2

4.3. Subjective QoE tests

The MPC, BOLA, Oboe, and Pensieve ABR algorithms were evaluated against the identical 5G bandwidth traces that we obtained during our study (see Section 4.1). This enabled fair comparisons to be made. The video chunks were then merged into a single UHD file. For each ABR algorithm, this procedure was repeated. We recruited 76 people to take part in the subjective tests in our research. All of the volunteers were not color-blind and had normal vision acuity, and they had no prior awareness of the DASH ABR algorithm utilized in the experiment. Twenty sequences were assessed by the volunteers. They were only required to watch five sequences of a specific UHD video to retain their focus on the evaluation. As a result, 19 people assessed every test sequence. This was

judged sufficient to guarantee that the results were not skewed by the presence of a few volunteers. The subjective assessments took place in a lab with controlled lighting. The sequences were presented on a 31.5 inches monitor. A resolution of 3840 x 2160 and aspect ratio of 16:9 were used.

BOLA promotes lower number of switches, fewer stalls, small stall duration times and better bandwidth utilization among multiple UHD players, thereby leading to a better subjective user experience, as justified by its MOS rating [18] in all evaluation environments, see Figure 7 and Figure 8.

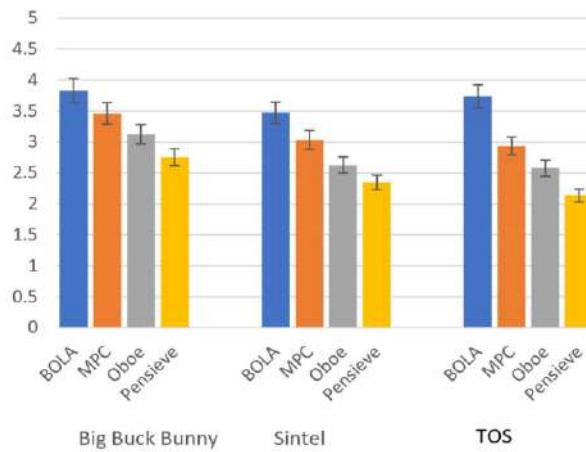


Figure 7. 5G FWA Business Environment 1: Subjective visual quality comparison of UHD DASH - an average MOS score ratings.

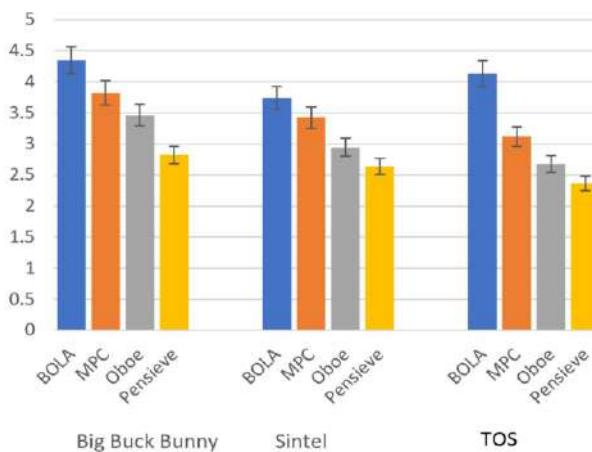


Figure 8. 5G FWA Business Environment 1: Subjective visual quality comparison of UHD DASH - an average MOS score ratings.

4.4. Discussion

BOLA utilizes buffer information to make its next segment streaming decisions. This metric proves more successful than the metrics used by the other ABR DASH-based algorithms explored in this paper. MPC utilized incoming bitrate and buffer metrics in its adaptation strategy. However, the incoming rates are difficult to predict due to highly varying nature of 5G traffic. This fact interferes with the MPC players ability to make proper inferences on the incoming bitrates. Thus, the MPC player performance is lower than the fully buffer-based algorithm, BOLA. Oboe utilizes throughput estimate over the previous five chunks. Its performance suffers and again, we see sub-optimal performance with an ABR using incoming rate estimates as part of its adaptation strategy. Pensieve performs the worst. A possible explanation is that, in order to train the model to understand 5G FWA specific features and make better judgments, a larger dataset is required. However, this merits more investigation. In addition, coping with the high bandwidth variations in 5G FWA may cause the player to sometimes run out of data chunks due to for example a prolonged spell of low bandwidth which could result in more frequent video stalls.

5. Conclusion

There has not been enough research conducted on video streaming in 5G FWA networks. Most researchers focus on mobile 5G video streaming. To bridge this research gap we conducted DASH-based video streaming experiments mimicking a 5G FWA network with a bottleneck gateway. We used the MPC, BOLA, Oboe and Pensieve ABR DASH algorithms. Experiments consisted of all small business workers viewing UHD videos or joining the streaming session. Most algorithms perform poorly in the 5G FWA experiments, except BOLA. Many ABR algorithms use a throughput predictor to factor network throughput into their decisions, and their success is highly reliant on predictive accuracy. ABR 5G throughput prediction is poor due to high 5G bandwidth variability. BOLA is a strictly buffer-based ABR algorithm thus does not suffer from inaccurate predictions as the others (MPC, Oboe) or lack of bigger training dataset in the case of Pensieve. Current 5G covers several bands and has a wide range of network performance. Thus, developing improved throughput prediction methods is critical not just for making ABRs function well over 5G FWA, but also for making ABRs work well in general. New adaptive techniques will have to be developed to cope with the ever expanding 5G FWA environment.

References

- [1] Asplund H, Astely D, von Butovitsch P, Chapman T, Frenne M, Ghasemzadeh F, et al. Advanced Antenna Systems for 5G Network Deployments: Bridging the Gap Between Theory and Practice. Academic Press; 2020.
- [2] Topyan K, Ulema M. Architectural and Financial Considerations for Deploying 5G Based Fixed Wireless Access. In: 2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom). IEEE; 2020. p. 1-6.
- [3] Bechta K, Ziolkowski C, Kelner JM, Nowosielski L. Modeling of downlink interference in massive MIMO 5G macro-cell. Sensors. 2021;21(2):597.

- [4] Tuysuz MF, Aydin ME. QoE-based mobility-aware collaborative video streaming on the edge of 5G. *IEEE Transactions on Industrial Informatics*. 2020;16(11):7115-25.
- [5] Khan K, Goodridge W. S-MDP: Streaming with markov decision processes. *IEEE Transactions on Multimedia*. 2019;21(8):2012-25.
- [6] Khan K, Goodridge W. B-DASH: broadcast-based dynamic adaptive streaming over HTTP. *International Journal of Autonomous and Adaptive Communications Systems*. 2019;12(1):50-74.
- [7] Khan K, Goodridge W. Server-based and network-assisted solutions for adaptive video streaming. *International Journal of Advanced Networking and Applications*. 2017;9(3):3432-42.
- [8] Ramadan E, Narayanan A, Dayalan UK, Fezou RA, Qian F, Zhang ZL. Case for 5G-aware video streaming applications. In: *Proceedings of the 1st Workshop on 5G Measurements, Modeling, and Use Cases*; 2021. p. 27-34.
- [9] Navarro-Ortiz J, Romero-Diaz P, Sendra S, Ameigeiras P, Ramos-Munoz JJ, Lopez-Soler JM. A survey on 5G usage scenarios and traffic models. *IEEE Communications Surveys & Tutorials*. 2020;22(2):905-29.
- [10] Kimura T, Kimura T, Matsumoto A, Yamagishi K. Balancing quality of experience and traffic volume in adaptive bitrate streaming. *IEEE Access*. 2021;9:15530-47.
- [11] Le HD, Nguyen CT, Mai VV, Pham AT. On the throughput performance of tcp cubic in millimeter-wave cellular networks. *IEEE Access*. 2019;7:178618-30.
- [12] Sim MS, Lim YG, Park SH, Dai L, Chae CB. Deep learning-based mmWave beam selection for 5G NR/6G with sub-6 GHz channel information: Algorithms and prototype validation. *IEEE Access*. 2020;8:51634-46.
- [13] Spiteri K, Urgaonkar R, Sitaraman RK. BOLA: Near-optimal bitrate adaptation for online videos. *IEEE/ACM Transactions on Networking*. 2020;28(4):1698-711.
- [14] Yin X, Jindal A, Sekar V, Sinopoli B. A control-theoretic approach for dynamic adaptive video streaming over HTTP. In: *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*; 2015. p. 325-38.
- [15] Mao H, Netravali R, Alizadeh M. Neural adaptive video streaming with pensieve. In: *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*; 2017. p. 197-210.
- [16] Akhtar Z, Nam YS, Govindan R, Rao S, Chen J, Katz-Bassett E, et al. Oboe: Auto-tuning video ABR algorithms to network conditions. In: *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*; 2018. p. 44-58.
- [17] Quinlan JJ, Sreenan CJ. Multi-profile ultra high definition (UHD) AVC and HEVC 4K DASH datasets. In: *Proceedings of the 9th ACM Multimedia Systems Conference*; 2018. p. 375-80.
- [18] Kaipio A, Ponomarenko M, Egiazarian K. Merging of MOS of Large Image Databases for No-reference Image Visual Quality Assessment. In: *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE; 2020. p. 1-6.

Design and Implementation of Highly Integrated Electronic System for Small-Size Spacecraft

Xinfeng YAN¹, Wei LIU, Gaosi LI and Jian GENG

Beijing Institute of Space Long March Vehicle, Beijing, 10007, China

Abstract. Due to the tight installation space, strict weight requirements, numerous functional units and dense wireless equipments of small-size spacecraft, the traditional electronic system design of aircraft cannot meet the requirements of miniaturization and lightweight for small-size spacecraft. In this paper, a highly integrated electronic system suitable for small-size spacecraft is designed, in which many key technologies such as miniaturization and lightweight design technology, multi frequency wireless electromagnetic compatibility design technology, small aircraft low delay cooperative networking and high-precision differential positioning are adopted. The proposed highly integrated electronic system makes timing control, attitude and orbit control, wireless telemetry, networking communication, satellite navigation, power distribution and other functions integrated. The total weight of the system is about 7.62kg, which is much lighter than existing electronic system. The experiment results that the highly integrated electronic system achieves good effect. This technology has broad application prospects in small-size aircraft with strict weight and space requirements.

Keywords. Spacecraft, electronic system, small-size, highly integrated

1. Introduction

The electrical system on the aircraft mainly is supposed to realize various functions such as timing control, attitude and orbit control, wireless telemetry, networking communication, satellite navigation, power distribution and so on [1].

The traditional aircraft electronic system is generally composed of several subsystems such as control subsystem, telemetry subsystem, networking communication subsystem, power distribution subsystem and satellite navigation subsystem. Each subsystem is composed of several single machines. The functions of the electronic system realize through cable network interconnection within and between the subsystems. The traditional aircraft electronic system has complex equipment composition and connection relationship, large space occupation and heavy weight.

Small-size spacecraft have many features such as light weight, small size, multiple components of electronic system and complex function. The traditional electronic system design cannot meet the requirements of small space vehicle for miniaturization and lightweight.

¹ Corresponding author: Xinfeng YAN, yanxf185@163.com

Integrated electronic technology can realize the functional comprehensive design of the aircraft electronic system. Integrated electronic technology has been widely used in foreign spacecraft [2-6]. As early as the mid-1990s, the APL laboratory of American space department established the advanced satellite technology committee to study the advanced electronic system architecture for future space projects. The "Opportunity", "Courage" and "Curiosity" Rovers in the Mars exploration project adopted integrated electronic technology [7].

Domestic spacecraft such as the firefly-1 Mars probe [8], the chang'e-3 patrol [9], and the beidou-3 satellite [10] have adopted integrated electronic technology.

This paper provides an integrated electronic system design suitable for small-size spacecraft. It adopts an integrated design, has the advantages of simple electronic system equipment composition and connection relationship, small occupied space and light weight, and can meet the functional requirements of small-size spacecraft for complex electronic systems.

2. General system formula design

2.1 System composition and function

The highly integrated electronic system of spacecraft is composed of control system, telemetry system, networking communication system, satellite navigation system and power distribution system. The system composition block diagram is shown in Figure 1.

2.2 Control system

The control system mainly completes the timing control function according to the overall timing flow. According to the information from inertial measurement module, satellite navigation and star sensor, the dynamic system is controlled to work to realize

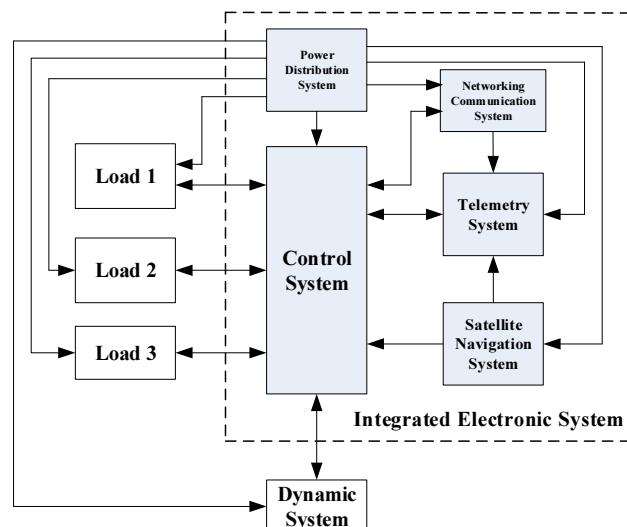


Figure 1. System composition block diagram

the attitude and orbit control function of spacecraft by solving the control equation. The control system is composed of integrated control computer, inertial measurement unit, star sensor and on-board cable network.

2.3 Telemetry system

The telemetry system receives and frames sensor image data and digital data, collects and frames analog parameters such as voltage, temperature and pressure on the aircraft, encodes the framed data by TPC, encrypts it, and then transmits the output PCM code stream with transmitter modulation and power amplification via the telemetry antenna. Due to the miniaturization design requirements of spacecraft, the functions of telemetry system such as data receiving, collecting, framing, channel coding and encryption are integrated in the control computer, which reduces the number of equipment and saves weight and size. The telemetry system is mainly composed of integrated control computer, transmitter, telemetry antenna and high-frequency cable network.

The telemetry frame format is designed as 208×64 , two subframes, each channel word length is 1 byte, the major frame sampling rate is 5.8323kHz, the subframe sampling rate is 91.1Hz, and the wireless telemetry code rate is 9.704969Mbps.

2.4 Networking communication system

The networking communication system completes the information exchange between the two spacecraft. The main interactive information includes the attitude, orbit data and satellite navigation data of the spacecraft. The networking communication system is mainly composed of networking communication components, networking transceiver antennas and high-frequency cable network.

2.5 Satellite navigation system

The satellite navigation system completes the satellite navigation and positioning of the two spacecraft, performs differential decomposition through the satellite navigation information exchanged by the networking communication system, and outputs the differential positioning results to the control system. The satellite navigation system is mainly composed of satellite navigation components, satellite navigation antennas and high-frequency cable network.

2.6 Power distribution system

The power distribution system provides 48V and 28V power supply for the highly integrated electronic system and other external systems, isolates and converts 28V, outputs one current of isolated 15V and three currents of isolated 5V voltage to other modules of the highly integrated electronic system, and completes the functions of power transfer and emergency power off. The power distribution system is mainly composed of power distribution components, control batteries and load batteries.

3. Key technologies and solutions

3.1 Miniaturization and lightweight design technology

In this paper, the miniaturization and lightweight design of highly integrated electronic system for spacecratfs is realized through the design of electronic structure integration, control computer and telemetry equipment integration, unified power supply and distribution design and so on.

3.1.1 Electronic structure integration

The power distribution components, control computer, image information processing computer, networking communication components and satellite navigation components are designed for electrical structure integration. The image information processing computer is the component equipment of load 1. The above equipments are uniformly installed in an integrated electronic cabin, and each equipment exists in the form of module components. The modules are connected through the inter board connectors. The upper and lower structure shells of integrated control computer and image information processing computer are canceled with the frame structure retained. Considering the adaptability of electromagnetic compatibility, the structure shells of power distribution components, networking communication components and satellite navigation components are also retained. In order to further reduce the weight of the equipment, the structural shell is made of magnesium lithium alloy. The schematic diagram of the integrated electronic cabin is shown in Figure 2.

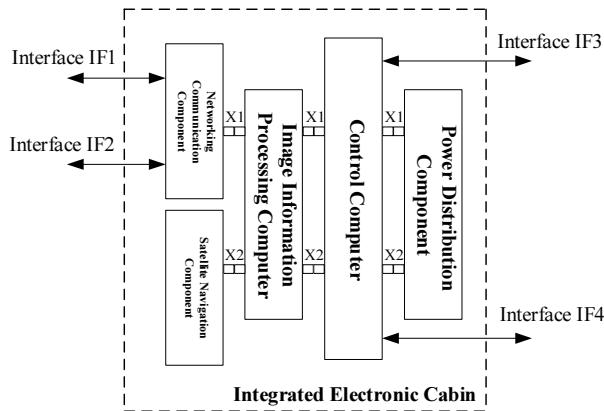


Figure 2. Schematic diagram of integrated electronic cabin

X1 and X2 are inter board connector interfaces, and each component achieves power supply and data interaction through X1 and x2. The interactive signals between the integrated electronic cabin and the equipment outside are collected on the networking communication component and led out through the external interfaces IF1 and IF2. The interactive signals between the integrated electronic cabin and the ground are uniformly led out by the external interface IF3 on the control computer. The test signals of the integrated electronic cabin are uniformly led out through the test interface IF4 on the control computer.

3.1.2 Control computer and telemetry equipment integration

The traditional control computer and telemetry equipment are integrated and designed to share processing chip and interface resources. After the integrated design, the serial port circuit is reduced from 18 to 8. The functional block diagram of integrated control computer is shown in Figure 3, which is composed of main control module, interface module, power conversion module, encryption module and crystal oscillator circuit. The integrated control computer do not only realize the functions of timing control, solving control equations and controlling the power system, but also realizes the functions of data receiving, collecting, framing, channel coding and encryption of telemetry data. Among, the main control module adopts Xilinx SOC chip XC7Z020-1CLG484, together with two DDR3 chips and one nor flash chip.

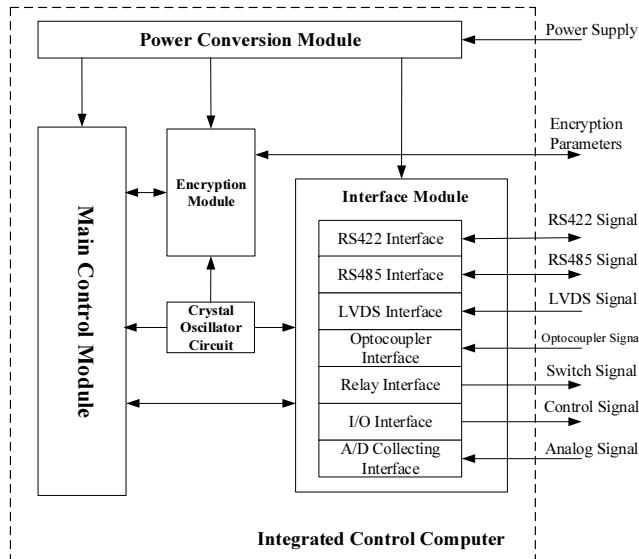


Figure 3. Functional block diagram of integrated control computer

3.1.3 Unified power supply

The power supply interface of each component is integrated and designed as a power distribution component. The power distribution component is composed of power supply and distribution circuit, input EMI filter circuit, DC/DC power circuit, output filter circuit and other units. The circuit block diagram is shown in Figure 4. Among, the power supply and distribution circuit includes power transfer circuit and emergency power-off circuit, which is realized by solid-state relay. Discrete devices are used to build the input EMI filter circuit, and transient voltage suppression diodes are used to suppress the maximum 200V voltage peak of the bus. Common mode filter inductor and differential mode capacitor are selected to suppress common mode and differential mode noise. RC noise suppression circuit is adopted to suppress absorption noise. High frequency magnetic beads are used to further suppress EMI high frequency noise at the input side.

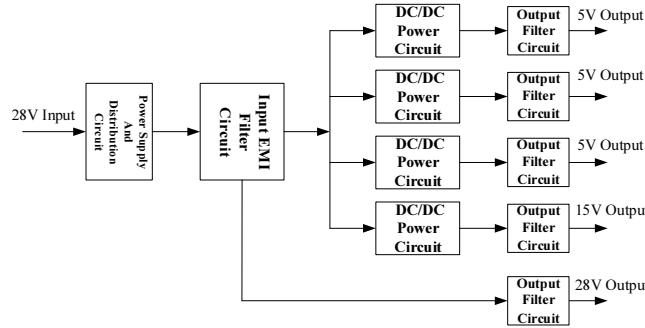


Figure 4. Circuit block diagram of power distribution component

3.2 Multi-frequency wireless electromagnetic compatibility design

The size of spacecraft is small, and the composition of wireless system on the aircraft is complex, including networking communication system, satellite navigation system and wireless telemetry system. The antennas of the three wireless systems are installed close to each other and wireless frequency points are close to each other partly. Multiple methods such as local shielding of metal shell, determination of specific frequency points through third-order intermodulation analysis, use of different antenna polarization methods to increase spatial isolation, and extra filtering of transmitting sources in sensitive frequency bands are used to jointly realize multi frequency point wireless electromagnetic compatibility in a narrow space.

3.3 Communication design of low-delay high-speed anti-interference networking

The networking communication system adopts TDMA and frequency hopping technology. Considering the minimization of transmission and processing delay, each time frame is designed to contain three time slots, as shown in Fig. 5, and each time frame is 10ms. The primary node occupies the first time slot (TS1), and the primary node has the function of network management. However, the other non-primary node occupies the second time slot (TS2). The length of time slots TS1 and TS2 is 4.5ms. In the last 1ms of each time frame (corresponding to time slot TS3), each node remains silent. The primary node senses the channel and detects the interference on each frequency point.

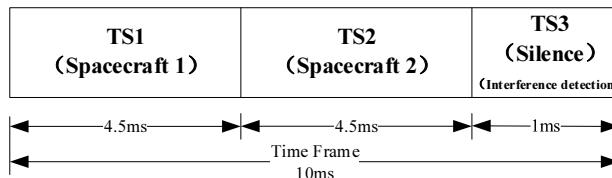


Figure 5. Diagram of time slot allocation

As the duration of each time slot is very short, in order to support fast reception synchronization, the system adopts several methods, such as, using Barker preamble with good correlation in the link layer, selecting MSK differential decomposition and LDPC coding to improve the transmission rate in the physical layer, and finally realizes

the effect of receiving synchronization in 0.5ms. In the physical layer, in order to reduce space loss and lower transmission power, LDPC(1536, 1024) coding is combined with MSK modulation. The bit sequence after channel coding is interleaved for MSK modulation and the receiver uses orthogonal differential demodulation. According to the antenna pattern simulation result and link calculation, within the range of $\pm 80^\circ$ antenna pattern, the bit transmission rate capability is not less than 1.33Mbps at 5km, the link margin is greater than 6dB, and the required transmission power is only 3W. Meanwhile, according to the measured results of the physical prototype, the maximum end-to-end delay of the two networking communication nodes will not exceed 14ms.

3.4 Design of high precision satellite navigation and differential positioning

The satellite navigation component is mainly composed of RF channel, navigation signal processing unit, information processing unit, interface unit, power supply unit, clock unit, etc. The composition block diagram is shown in Figure 6. The RF channel receives the GPS/B1 signal input by the antenna and outputs the IF signal to the navigation signal processing unit after amplification, mixing and filtering. The signal is captured, tracked and measured the pseudo range by the navigation signal processing unit. The information processing unit calculates the position, speed and time information of the spacecraft according to the measured pseudo range and ephemeris original information, and sends the obtained information to the control system through the interface unit.

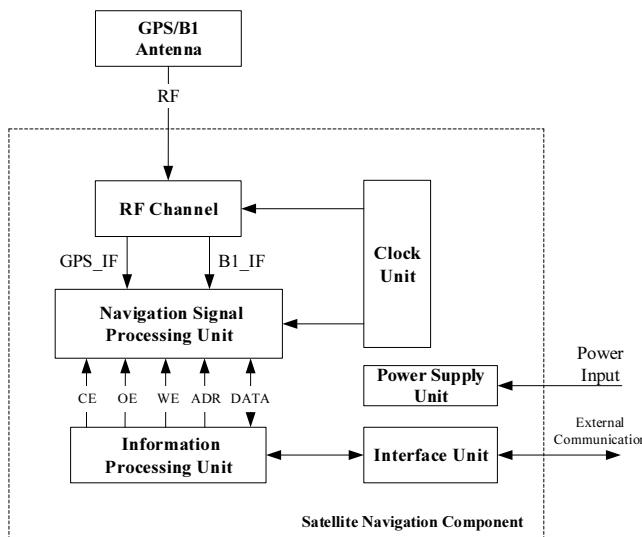


Figure 6. Composition block diagram of satellite navigation component

In order to improve the relative position accuracy of two spacecraft, the satellite navigation differential positioning design is adopted. Each spacecraft first performs satellite receiving and positioning respectively, and then the two sides use the networking communication system to interact received the ephemeris numbers. According to the similar ephemeris numbers, the positioning results are calculated by

difference decomposition, which can offset the same position error and improve the relative position accuracy between the two spacecraft.

The error sources of differential positioning mainly include satellite clock error, satellite ephemeris, ionospheric delay error, tropospheric delay error, multipath error and receiver noise. The flight process of the spacecraft is relatively open, the number of visible satellites is more than 12 (GPS + BDS), and the PDOP value is basically less than 3. The distance between the two spacecraft is about 3km so that the included angle of the two receivers relative to the same navigation satellite is very small, and the connection between the two receivers and the navigation satellite can be considered to be parallel. After differential calculation, except multipath error and receiver noise, the positioning error caused by other error sources tends to the same side. In the positioning results, it can be considered that the relative error of the two receivers is almost equal to the positioning error of a single receiver. According to the calculation, the relative error of the dual receiver is not more than 15m.

3.5 Centralized data interaction and management architecture based on integrated control computer

The integrated control computer is the general controller of the spacecraft, which is used to collect the information of inertial measurement unit, image information processing computer, satellite navigation component and star sensor, control the power system, load, networking communication component and other equipment according to the control equation and control sequence. The integrated control computer is also used to communicate with the rocket and the ground test bench and make the two spacecraft perform the corresponding sequential actions. The communication interface between devices is the byte-by-byte transmission interface, such as isolated RS422 or RS485. The communication connection relationship between devices is shown in Figure 7.

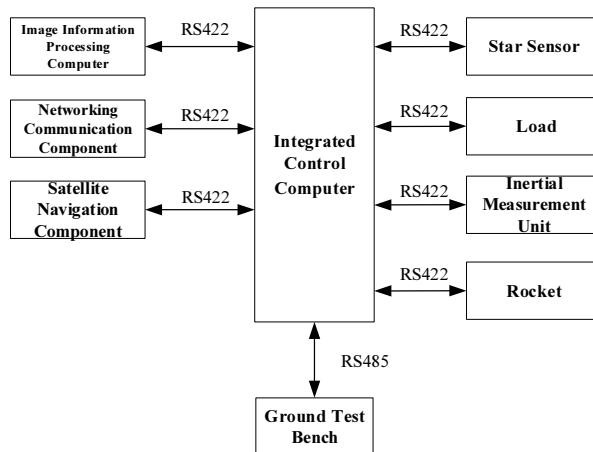


Figure 7. Communication connection relationship between devices

The data frame format consists of frame header, command byte, data, fill byte, check code, data length and frame tail, as shown in Table 1.

Table 1. Definition of data frame format

Order	1	2	3	4	5	6	7
Field	Frame header	Command byte	Data	Fill byte	Check code	Data length	Frame tail
Number of bytes	2	1	Floating length (0 ~ 249)	1	1	1	2

In order to improve the reliability of data interaction, redundancy design is adopted. For the instruction with response, if no reply is received within 200ms after the instruction is sent, the instruction will be retransmitted, and the above process can be repeated twice. As long as the correct reply is received, the process will be stopped. If no reply is received for three times, the function of the instruction is considered failed.

For the instruction without response, the send will be stopped after three times continuously at an interval of 10ms. If the receiver captures the correct instruction once, it will respond to the instruction.

4. Experiment results

The highly integrated electronic system is applied to a spacecraft. After a number of test assessments and the flight test successfully completed with a project, the system works as expected and achieves good results, which verifies the correctness and matching of the design of the highly integrated electronic system. The total weight of the highly integrated electronic system is 7.62kg, and, however, the weight of existing electronic system is normally more than 15kg because of the independent design of functional modules. In terms of power consumption, the overall power consumption of the highly integrated electronic system is about 160W. In contrast, the overall power consumption of the existing electronic system is about 280W.

The telemetry subsystem in the highly integrated electronic system obtains complete telemetry data such as image, control and analog data. The data are detected, stripped, subcontracted and processed. The analysis results show that the working sequence of the spacecraft is correct, the flight attitude and orbit control works well, the networking interaction between the two spacecraft is fluent, and the satellite receiving and positioning and differential resolution works as expected. The temperature curve of a test is shown in Figure 8, the distance curve of two spacecraft is shown in Figure 9, and the change curve of spacecraft PDOP value is shown in Figure 10.

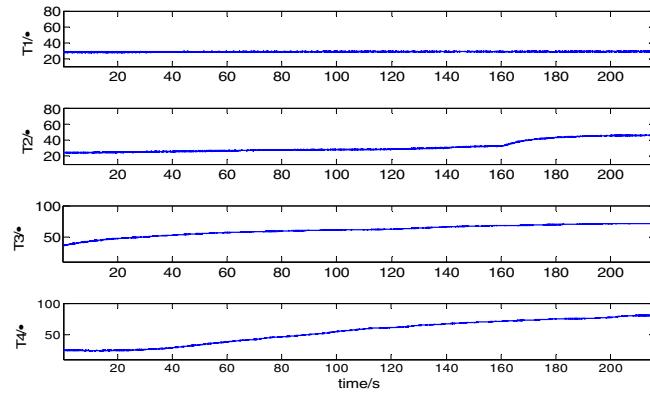


Figure 8. Temperature curve of a test

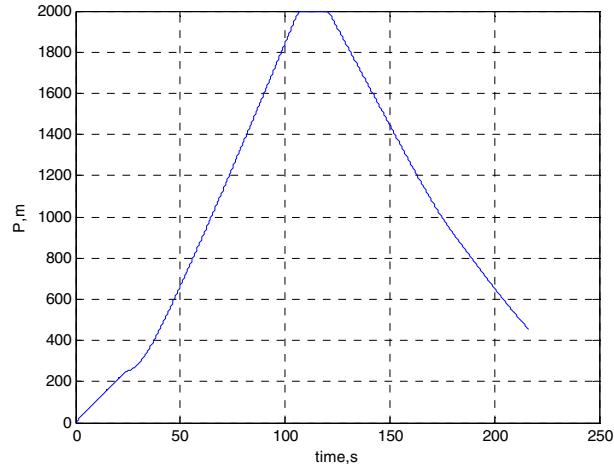


Figure 9. Distance curve of two spacecraft

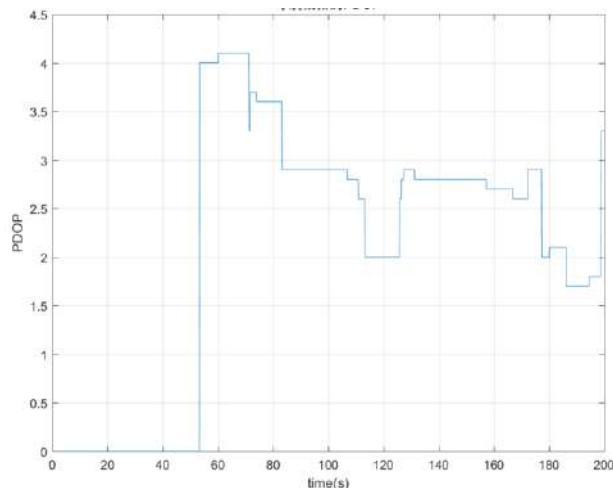


Figure 10. Change curve of spacecraft PDOP value

5. Conclusion

This paper introduces a highly integrated electronic system suitable for small-size spacecraft. Under the constraints of tight installation space, strict weight requirements, numerous functional units and dense wireless equipment, it realizes the integration of timing sequence, attitude and orbit control, wireless telemetry, networking communication, satellite navigation, power distribution and other functions. The miniaturized highly integrated electronic system breaks through the key technologies such as miniaturization and lightweight design, multi-frequency wireless electromagnetic compatibility design, small aircraft low-delay cooperative networking and high-precision differential positioning. The experiment results that the highly integrated electronic system achieves good effect. This technology has broad application prospects in small-size aircraft with strict weight and space requirements.

Reference

- [1] CHEN Shimiao1, NI Shuyan, LIAO Yurong. Summary of Integrated Electronic Systems of Microsatellite [J]. SPACE ELECTRONIC TECHNOLOGY, 2020,5:82-87.
- [2] BLUE R C, BOLOTIN G S, Development of a new generation of avionics for space applications[C], X2000 Advanced Avionics Project. Citeseer, 2003:1-5.
- [3] Nguyen Q, Yuknis W, Pursley S, et al. A high performance command and data handling system for NASA's lunar reconnaissance orbiter. AIAA Space 2008 Conference & Exposition, San Diego, California, 2008. 1-8.
- [4] Berger R, Burcin L, Hutcheson D, et al. The RAD6000MCTM system-on-chip microcontroller for spacecraft avionics and instrument control. Proceedings of IEEE Aerospace Conference, 2008. 1-14.
- [5] Rufino J, Craveiro J, Verissimo P. Building a time-and space-partitioned architecture for the next generation of space vehicle avionics. IFIP International Federation for Information Processing, 2010. 179-190.
- [6] ZHAI J Q, Li X F, Introduction of oneWeb system and domestic LEO internet satellite system[J]. Space Electronic Technology ,2017, 14(6):1-7.
- [7] Bajracharya M, Maimone M W, Helmick D. Autonomy for Mars rovers: Past, present, and future. IEEE Computer Society, 2008. 44-50.
- [8] GAI Jian-ning, FAN You-cheng, SHEN Li, et al. Design of Integrated Electronic Subsystem of YH-1 Mars Probe[J]. AEROSPACE SHANGHAI, 2013,30(4):139-146.
- [9] Chen J X, Zhang Z, Wang L, et al. Design and realization of the integrated electronic system for the Chang'E 3 lunar rover (in Chinese). Sci Sin Tech,2014, 44: 450-460.
- [10] Jia W S, Zeng L L, Li L M, et al. Design of Beidou-3 satellite avionics system[J]. Astronautical Systems Engineering Technology,2020,4(6):50-55.

SDN-Enabled 3C Resource Integration in Green Internet of Electrical Vehicles

Handi CHEN^a, Xiaojie WANG, ^{b,1}, Zhaolong NING^{a,b}, and Lei GUO^b

^a*School of Software, Dalian University of Technology, Dalian, Liaoning, China*

^b*School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China*

Abstract. With the advocacy of green renewable energy, Electric Vehicles (EVs) have gradually become the mainstream in the automobile market. Due to the finite edge resources of the Internet of EVs, this paper integrates idle communication, caching and computational resources of EVs to enrich the available resources for vehicular task migration. Considering the limited capacity and resources of EVs, a distributed lightweight imitation learning-based efficient Task cOoperative migration Policy Integrating 3C resource policy, named TOPIC, is proposed to maximize the obtained quality of service. The experimental results based on the real-world traffic dataset of Hangzhou (China) demonstrate the QoS obtained based on the expert policy and agent policy of TOPIC is about 3 times higher than other representative policies.

Keywords. Heterogeneous resource management, software-defined network, task migration, imitation learning

1. Introduction

1.1. Background

Considering the deterioration of the ecological environment and the shortage of energy, Electric Vehicles (EVs), utilizing renewable energy [1], are gradually occupying the vehicular market. However, the contradiction between users' requirements and EVs' limited resources in the latency-sensitive scenario with energy constraints is still one of the thorniest issues to be solved. Integrating heterogeneous edge resources distributedly to overcome overloaded vehicular tasks is a promising approach to reduce the workload of devices and execution cost [2, 3].

1.2. Literature Review

However, the edge resources, consisting of Communication, Computation and Caching (3C), provided by Mobile Edge Computing (MEC) servers deployed in real-world scenarios are rather limited [4]. Homogeneous resource management has been studied adequately, while the study of heterogeneous resource integration is still immature. Zhou *et*

¹Corresponding Author: Xiaojie Wang, E-mail: xiaojie.kara.wang@ieee.org.

al. [5] integrated caching and computational resources through designing Service Function Chain (SFC), which is a common method to integrate heterogeneous resources. The 2C models in [6, 7] jointly optimized communication and computational resources allocation. In [6], the downloading communication phase was accelerated with computation replication in MEC networks to decrease the latency of communication. Ding *et al.* [7] minimized the weighted energy consumption for ground devices by joint ground user equipment association, multi-user transmit precoding, computation task assignment and communication resource allocation. The joint optimization problem was solved based on traditional optimization methods.

For the integration of 3C resources, how to comprehensively integrate these three heterogeneous resources is rather challenging. Luo *et al.* [8] proposed an efficient Fog-enabled 3C resource sharing framework (F3C) and solve resource allocation issues based on the auxiliary graph and proposed F3C algorithm to optimize energy consumption. Tang *et al.* [9] proposed a 3C resource integration framework to motivate device cooperation and schedule resources more flexibly. However, the investigated scenarios [8, 10] are relatively static, and the dynamics of mobile devices are not taken into account in the proposed algorithms.

However, the infrastructure resources are always regarded as sufficient in traditional studies [11], which is not realistic in the real world, especially in the dynamic scenario with fierce competition. The complexity of integrating heterogeneous resources promotes studies with leveraging Software-Defined Network (SDN), improving the resource pooling level via decoupling traditional network architecture.

1.3. Research Challenges and Contributions

Integrating heterogeneous edge resources with SDN is a promising approach for flexible resource scheduling and network service automation, motivating our work. In addition to the complexity of managing heterogeneous resources, scheduling resources in the dynamic ultra-dense scenario has fiercer competition among EVs with limited batteries than traditional vehicles with internal combustion engines, especially for the selfish and rational users. To solve the abovementioned issues, we construct a task migration framework integrating heterogeneous resources and propose a distributed Task cOoperative migration Policy Integrating 3C resources (TOPIC) based on Imitation Learning (IL) [2] to maximize the obtained Quality of Service (QoS). The main contributions are summarized as follows: (1) We construct an SDN-enhanced intelligent, cooperative task migration framework and formulate the optimization problem; (2) We design a distributed IL-based cooperative task migration policy, TOPIC, to make a trade-off between rent and the obtained service rate; (3) The efficiencies of our proposed policy has been validated with the datasets of Hangzhou (China), compared with representative policies.

The rest of this paper is organized as follows: the system model and problem formulation are illustrated in Section 2. Section 3 specifies the TOPIC policy. Experimental results and performance evaluation are analyzed in Section 4, followed by the conclusion Section 5.

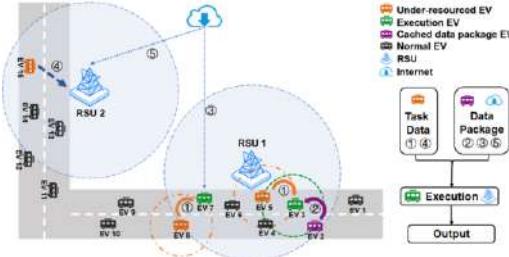


Figure 1. An SDN-enhanced architecture for 3C resource integration.

2. System Model and Problem Formulation

2.1. System Model

A dynamic vehicular network is considered as illustrated in Fig. 1. This paper investigates task migration problem via integrating 3C resources in one cluster-based vehicular network consisting of v EVs and an Road Side Unit RSU R deployed with an SDN controller. The vehicular task generated by EV V_i is denoted as S_i with service type S_i^{st} , $S_i^{st} \in \mathcal{K}$ where \mathcal{K} represents the set of service types. As shown in Fig. 1, vehicular task execution is divided into three parts, i.e., input, execution and output.

Herein, the input of vehicular task consists of task data and the data package required by the service type corresponding to the task. The task data is transmitted from the under-resourced EV, illustrated as process ① and ④ by Vehicle-to-Vehicle (V2V) communication and Vehicle-to-RSU (V2R) communication, respectively. Since the mobility of vehicles, Orthogonal Frequency-Division Multiple Access (OFDMA) is leveraged to guarantee steady communication. The signal noise ratio of channel ij is calculated by $p_i g_{i,j} / \sigma^2$ without interference, and the signal to interference plus noise ratio between V_i and R is defined as $p_i g_{i,r}^2 / \sigma^2 + I_{i,r}$. Herein, p_i , $g_{i,*}$ (including $g_{i,j}$ and $g_{i,r}$), σ^2 and $I_{i,r}$ respectively indicate the transmission power, channel gain, additive white Gaussian noise and interference. The inference is generated by other users occupied V2R sub-channels simultaneously, $I_{i,r} = \sum_{i=1}^v x_{i,r} p_{i,r} (g_{i,r})^2$ and $x_{i,r}$ indicates the decision variable to indicate the channel occupation. The communication rate among devices (including EVs and RSU) can be obtained by Shannon equation.

The required data package can be obtained via downloading from the Internet (process ③ and ⑤) or sharing by other vehicles caching this data (②). The idle caching resource can be utilized to cache data rather than repeatedly downloading for reducing the redundant cost [12]. Herein, data package obtained rate can be denoted by $\sum_{j=1}^v y_{j,h} R_{j,h}^{comm} + (1 - \sum_{j=1}^v y_{j,h}) R_j^{down}$, where binary variable $y_{j,h}$ indicate the decision of transmitting data package from V_h to V_j . In this paper, the output of task execution is viewed as small enough to be negligible [13]. After obtaining input data, task can be executed in device. The computing rate R_i^{comp} (Mbits per second) of S_i can be calculated as $\sum_{j=0}^v x_{i,j} s_i / t_i^{comp}$. Herein, $j = 0$ indicate the RSU.

Since the rationality and selfishness of vehicular users, we set battery power and available computational capacity as the main aspects to define the unit price of renting computation resources of EVs and RSU, which are calculated by $\rho^{comp}(f_j, \varepsilon_j) = (\alpha e^{\beta/\varepsilon_j}) / f_j$ and $\rho_r^{comp}(f_r, 1) = (\alpha e^{\beta}) / f_r$, respectively. Herein, parameters α and β rep-

resents price coefficients to adjust the impact of computational resources and battery power on unit rent at a time slot.

2.2. Problem Formulation

In this investigation, we define QoS as the ratio of payment and the sum service rate, i.e., $U_i = R_i^{total} / P_i^{total}$. The optimization objective is set as maximizing the obtained QoS, i.e.,

$$\max_{x_{i,j}, y_{j,h}} U_i, \quad (1)$$

and is subject to task migration delay constraint, i.e., $t_i^{mig} \leq \min\{d_i, d_{i,j}^{comm}\}$, indicating the migration delay cannot exceed the delay limitations of S_i and communicable distance; power consumption constraint, i.e., $\varepsilon_i^{cost} < \varepsilon_i, V_i \in V$, indicating the cost of each device cannot exceed the minimum power threshold; and communication capacity constraint, i.e., $\sum_{i=1}^v r_{i,r}^{comm} \leq Z_r, V_i \in V$, representing the occupied communication resources cannot exceed the communication capacity.

3. Task Cooperative Migration Policy Integrating 3C Resources

To solve the optimization problem formulated in the previous section, we design a distributed IL-based TOPIC policy. The policy consists of an expert policy and an agent policy. Compared to vehicles, RSU deployed with an SDN controller can observe the global states, which can be viewed as the expert node to execute expert policy with global observation. We propose a Deep Reinforcement Learning (DRL)-based expert policy and the Markov decision process is modeled as a quaternion $\langle \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P} \rangle$, where \mathcal{S} , \mathcal{A} , \mathcal{R} and \mathcal{P} represent system state, system action, reward function and state transition probability of TOPIC, respectively. The definitions of state, action and reward are specified as follows: The environment state $\mathcal{S}(t)$ is constructed with 3C states. The action $\mathcal{A}(t)$ consists of task execution and required data package obtainment, which is defined as: $\mathcal{A}(t) = \{x(t), y(t)\}$, where $x(t)$ and $y(t)$ are two vectors of binary variables. After executing the $\mathcal{A}(t)$, agent obtains $\mathcal{R}(t) = \sum_i^v U_i(t)$, if the constraints are satisfied; $\mathcal{R}(t) = 0$, otherwise. To obtain the optimal policy, the accumulate reward of the optimization problem is formulated as follows: $\mathcal{R}_i = \max_{x_{i,j}, y_{j,h}} \mathbb{E}[\sum_{t=0}^{T-1} \gamma^t \mathcal{R}_i(t)]$, where γ represents a discount factor indicating the impact of the current feedback gradually decreases over time, $\gamma \in (0, 1]$.

When expert node (RSU) obtains a task migration request of S_i , the SDN controller constructs adjacent matrix $\mathbb{M}_{v \times v}$ with communication rate between devices, i.e., $m_{i,j} = r_{i,j}^{comm}$, and observed the 3C states $\mathcal{S}(t)$. The SDN controller makes task migration decisions with a Double Deep Q-Network (DDQN) algorithm to maximize the obtained QoS by making a satisfying trade-off between service rate and payment with the proposed pricing mechanism. The core of the DRL-based expert policy is to obtain the optimal action with the maximum accumulate reward. With the global observation and the two-hop communicable search range of the requesting device, the caching data hit rate can be improved to reduce the redundant cost.

RSU generates the trajectories of expert policy through offline training, and the trajectories are transmitted to covered vehicles for training Deep Neural Network (DNN)-

Table 1. Simulation parameter settings

Parameter	Value
Bandwidth of RSU	50MHz
Transmit power of RSU	0.5W
Gaussian channel noise	5×10^{-13}
Available computational capacity of RSU	12Gcycles/s
Unit rent of RSU computational resource	0.7token/Gcycles
Unit rent of RSU communication resource	0.07token/Mbps
Unit rent of RSU download resource	0.2token/Mbps
Download rate of RSU	4Mbps
Available computational capacity of EV	[4, 8, 12, 16, 20]Gcycles/s
Download rate of EV	2Mbps
Unit rent of EV for resource downloading	0.1token/Mbps
Service type of tasks	{1, 2, 3, 4, 5}
Data size of task with service type K	$K \times 1Mbit$
Computational resource required by task with service type K	$K \times 2Gcycles$
Data size of data package required with service type K	$K \times 50Mbits$
Delay limitation of service type K	$K \times 30s$

based online agent policy. The observations obtained by requesting vehicles and decisions are viewed as the features and labels, respectively. Vehicles construct a DNN-based classifier to label the observation for obtaining task migration decisions.

4. Performance Evaluation

In this section, the simulation setup is introduced first. Then, the proposed TOPIC policy (includes an expert policy and an agent policy) is evaluated based on the real-world traffic dataset in Hangzhou (China).

4.1. Simulation Setup

We utilize PyCharm 2019.03 version based on Python 3.7.4 to perform experiments in a 64bit Windows 10 operating system computer with 128 GB RAM and an Intel(R) Xeno(R) CPU E5-2640 v4 with 2.4GHz frequency. TOPIC is realized on TensorFlow 1.14.0. Based on the characteristic of V2R communication, the channel gain of RSU and EVs are calculated by $127 + 30 \times \log_2 L$ [14], where L , indicating the distance between devices (includes EVs and RSU), can be obtained by GPS. The central coordinate [30.2547, 120.2741] of the dataset is set as the location of RSU. The channel of V2R is divided into 10 sub-channels. To simplify, the number of service types is reduced to 5, i.e., $K = 5$. Other simulation parameters are stated in Table 1.

To demonstrate the superiorities of TOPIC, we define six indicators, detailed as follows: (1) Obtained QoS: the ratio of the sum service rate to total payment; (2) Average rent: the average rent of migrating a task for execution; (3) Sum service rate: the sum service rate of executing a task; (4) Average power consumption: the average power consumption of under-resourced EVs; (5) Average increased Time-To-Live (TTL): the average increased TTL of the under-resourced EVs; (6) Average delay: the average execution latency of migrated tasks.

To demonstrate the effectiveness of the proposed TOPIC, the selected representative schemes are detailed as follows: (1) 3C DRL-based Policy [15]: It is a task offloading policy integrating 3C resources based on duelling DDQN to allocate 3C resources within one-hop communication range; (2) Computation+cache: computing can only be executed locally without communication resources; (3) Communication+cache: computing tasks can be migrated to RSU or executing locally by making a trade-off between the communication and caching resources; (4) Communication+computation: this policy allocates communication and computational resources based our proposed TOPIC policy with caching state locally.

4.2. Experimental Results

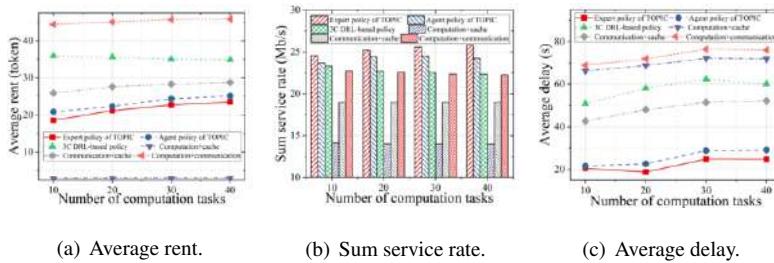


Figure 2. Performance of different computational workload.

Fig. 2 evaluates the performance with the distinct computational workload of the cluster. According to the definition of QoS, the performance of average rent cost and sum service rate of executing tasks under distinct computation workload are shown in Figs. 2(a) and 2(b), respectively. It can be observed that although the rent based on computation+cache policy is the lowest, computation+cache policy cannot utilize communication resources to reduce the local workload and increase the QoS, resulting in the lowest sum service rate for task execution. When the number of computing tasks is 40, the obtained sum service rate of TOPIC is 80% higher than computation+cache policy. The expert policy of TOPIC obtains the highest service rate with average rent second to computation+cache policy. Following the trajectories generated by the expert policy, the proposed agent policy can be better approximate to the performance of the expert policy. The rent of agent policy is only 7% higher than that of expert policy, and the obtained service rate is only 4% lower than that of expert policy. Figs. 2(a) and 2(b) indicate the proposed expert policy TOPIC makes a satisfying trade-off between the rent and the obtained service rate to maximize the total QoS under a large amount of computation workload. Fig. 2(c) shows the average delay cost of executing one task with distinct computation workloads, respectively. The average delay of the expert policy of TOPIC is much lower than the 3C DRL-based policy and other 2C policies when computation tasks are 40. That is because TOPIC policies within the two-hop search range decrease the competition for available resources increases as the computation workload becomes large. Agent policy of TOPIC imitates and follows the trajectories of expert policy for a satisfying trade-off between the obtained service rate and rent and further decreases the average delay for task execution.

The performance with distinct cached data package sizes is illustrated in Fig. 3. Fig. 3(a) evaluates the obtained QoS of different policies with distinct unit sizes of the data

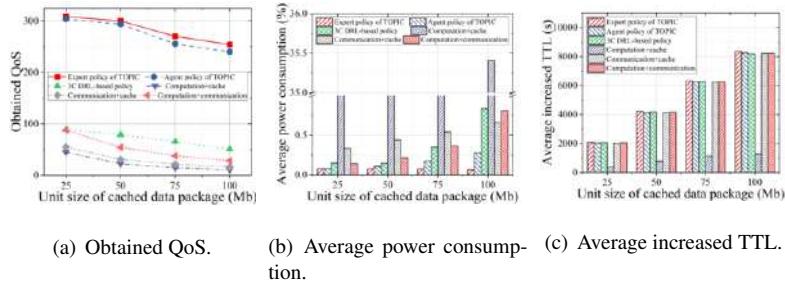


Figure 3. Performance of different cached data package size.

package. Even though the QoS of TOPIC (including expert policy and agent policy) decreases with the increasing unit data package size, its advantages over the 3C DRL-based policy increase. That is because the 3C states are constructed by the expert policy of TOPIC within a larger search range, improving the cache hit rate, and the proposed DNN-based agent policy imitates the trajectories of expert policy to label the states for making decisions accurately. In Fig. 3(b), the increasing rate of the TOPIC is much slower compared with other policies. For computation+cache policy, the power consumption of executing tasks that require large data packages is huge. The power consumption of the expert policy and agent policy is 62.09% and 43.65% lower than that of 3C DRL-based policy, respectively. That is because although the power consumption of 3C DRL-based policy can be reduced by integrating 3C resources, while the power consumption of waiting for completing tasks increases with the increasing data package. Fig. 3(c) evaluates the increased TTL of under-resourced EVs compared with local execution. Policies greatly increase TTL except for the computation+cache policy. Experimental results show that the computation+caching policy has to execute tasks locally without communication resources, resulting in large energy consumption. However, policies integrating communication resources migrate tasks to other devices for execution, greatly reducing energy consumption and extending the TTL of EVs. Herein, TOPIC (including expert policy and agent policy) with the higher cache hit rate, reducing the most redundant cost to save the battery power and increase the life cycle of vehicles.

5. Conclusion

This investigation focused on integrating 3C heterogeneous edge resources in Internet of EVs based on the SDN technique in the scenario with limited infrastructure resources. We formulated the optimization objective as maximizing the obtained QoS per token to motivate device cooperations. To make a satisfying trade-off between devices available resources and obtained service rate, we considered the constraints, including mobility, delay limitation, power consumption and available resources. A distributed task cooperation migration method based on IL was proposed to maximize the obtained QoS, namely TOPIC. Performance evaluations based on the real-world traffic dataset in Hangzhou (China) illustrated the effectiveness of our proposed policy from the perspectives of computation workload and unit cached data package size. Experimental results demonstrated that the QoS obtained by the expert policy and agent policy of TOPIC proposed in this work is about 3 times higher than those of other policies.

Although distributed IL is safer than centralized policy, designing a security mechanism to monitor malicious nodes to protect privacy in the ultra-dense dynamic scenario with overloaded communication is our further research topic.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61971084 and 62001073, and in part by the Dalian Young Science and Technology Star under Grant 2020RQ002.

References

- [1] H. Chen *et al.*, “Green Internet of vehicles: Architecture, enabling technologies, and applications,” *IEEE Access*, vol. 7, pp. 179 185–179 198, 2019.
- [2] X. Wang *et al.*, “Imitation learning enabled task scheduling for online vehicular edge computing,” *IEEE Transactions on Mobile Computing*, pp. 1–1, 2020.
- [3] Z. Ning *et al.*, “Blockchain-enabled intelligent transportation systems: A distributed crowdsensing framework,” *IEEE Transactions on Mobile Computing*, pp. 1–1, 2021.
- [4] Z. Ning *et al.*, “5G-enabled UAV-to-community offloading: Joint trajectory design and task scheduling,” *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2021.
- [5] Z. Zhou *et al.*, “Online orchestration of cross-edge service function chaining for cost-efficient edge computing,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 8, pp. 1866–1880, 2019.
- [6] K. Li *et al.*, “Exploiting computation replication for mobile edge computing: A fundamental computation-communication trade-off study,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4563–4578, 2020.
- [7] C. Ding *et al.*, “Joint optimization of transmission and computation resources for satellite and high altitude platform assisted edge computing,” *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.
- [8] S. Luo *et al.*, “Fog-enabled joint computation, communication and caching resource sharing for energy-efficient IoT data stream processing,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3715–3730, 2021.
- [9] M. Tang *et al.*, “Enabling edge cooperation in tactile Internet via 3C resource sharing,” *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 11, pp. 2444–2454, 2018.
- [10] L. Xu *et al.*, “Socially driven joint optimization of communication, caching, and computing resources in vehicular networks,” *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.
- [11] Z. Ning *et al.*, “Intelligent edge computing in Internet of vehicles: A joint computation offloading and caching solution,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2020.
- [12] Z. Ning *et al.*, “Joint computing and caching in 5G-envisioned Internet of vehicles: A deep reinforcement learning-based traffic control system,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2020.
- [13] Z. Ning *et al.*, “Intelligent resource allocation in mobile blockchain for privacy and security transactions: a deep reinforcement learning based approach,” *Science China Information Sciences*, vol. 64, no. 6, pp. 1–16, 2021.
- [14] Y. Sun *et al.*, “EMM: Energy-aware mobility management for mobile edge computing in ultra dense networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2637–2646, 2017.
- [15] Y. He *et al.*, “Integrated networking, caching, and computing for connected vehicles: A deep reinforcement learning approach,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 44–55, 2018.

Video Pre-Caching Joint Hand-Off and Content Delivery in Multi-Access Edge Computing Based EONs

Yutong CHAI¹, Shan YIN, Lihao LIU, Liyou JIANG, Shanguo HUANG
State Key Lab of Information Photonics and Optical Communications, China

Abstract. Multi-access Edge Computing (MEC) performs as a feasible solution when it comes to content delivery, for it can bring contents much closer to users. However, the hand-off (HO) and latency that occur in user movement reduce the users' quality of service. In this work, we consider the problem of high mobility handoff and content delivery of video streaming in the MEC based EONs. We propose a video pre-caching algorithm considering handoff and content delivery. The algorithm firstly selects the content delivery method and chunks the video accordingly using a preset threshold. Secondly, the algorithm chooses the shortest transmission path and calculates the latency time using Dijkstra method. Simulation results show that our algorithm significantly reduces the latency time and balances the server load compared to the other two baselines.

Keywords. MEC, handoff, content delivery, metro optical network

1. Introduction

The user case of video streaming required by vehicular users in the smart city falls in the category of ultra-reliable and low latency communication (URLLC) of 5G service type, making existing networks face new challenges. As a prospective computing paradigm, Multi-access Edge Computing (MEC) greatly decreases the waiting time between request and response through the way of pushing computational resources and storage resources to the network edges (e.g., Cellular Base Stations (BSs)). Furthermore, the 5G Automotive Association (5GAA) defined the concept of Cooperative Intelligent Transportation Systems (C-ITS) in 2017, stating that MEC would be the enabling technology for V2X (Vehicle to Everything) communications [1]. To make MEC construction more efficient, a heterogeneous network called Macro-Small cell network has become an emerging network deployment method.

However, handoffs (HOs) of vehicular users between BSs and even MECs happen frequently due to users' high mobility, decreasing the constant QoE of content [2]. Concurrently, how to minimize the latency time to make the content delivery process more efficient is another important issue. A few related solutions are aiming at solving the HOs problems in literature, such as edge caching based on Markov prediction [3], mobility prediction based RAN cache [4]. Also, several studies are exploring the

¹ Corresponding author: Yutong Chai; Email: chaiyutong@bupt.edu.cn

content delivery strategy. In literature [5], a distributed algorithm based on matching theory to minimize latency in the MEC-caching service network was provided. In literature [6], an MEC-based architecture was proposed to improve the performance of adaptive content delivery. This research led us to hypothesize that considering the handoff problem in advance can reduce the latency time of content delivery, which has not been studied in the previous research.

In this paper, we especially use video streaming as the required content of vehicular users in the smart city with MECs, which is a computation-intensive and caching-intensive service. Since the cost of storage resources is lower than the cost of bandwidth resources, distributing content in Multi-access Edge Computing is used as an effective method, allowing users to access the chunk of large-size video streaming from multiple MEC servers in advance. In addition, the main features of Elastic Optical Networks (EONs) like large bandwidth, low latency and flexible scheduling capabilities, make it a powerful technique to transmit information between BSs, thereby reducing the transmission cost [7]. Considering the above factors, we combine macro base stations (MBSs) and small-cell base stations (SBSs) which are all equipped with an MEC server to bring storage resources closer to users. Based on the structure, the video streaming can be pre-cached in the MEC server according to the traveling route of vehicles. Therefore, instead of migrating the whole content from one server to another, this paper proposes an algorithm which pre-caches the divided chunks in the BS MEC server covering the area where the vehicle is located. In this way, the content transfer time and HO time supporting service continuity in the dynamic and heterogeneous MEC network can be reduced effectively. The details of the design and simulation results are presented in this paper.

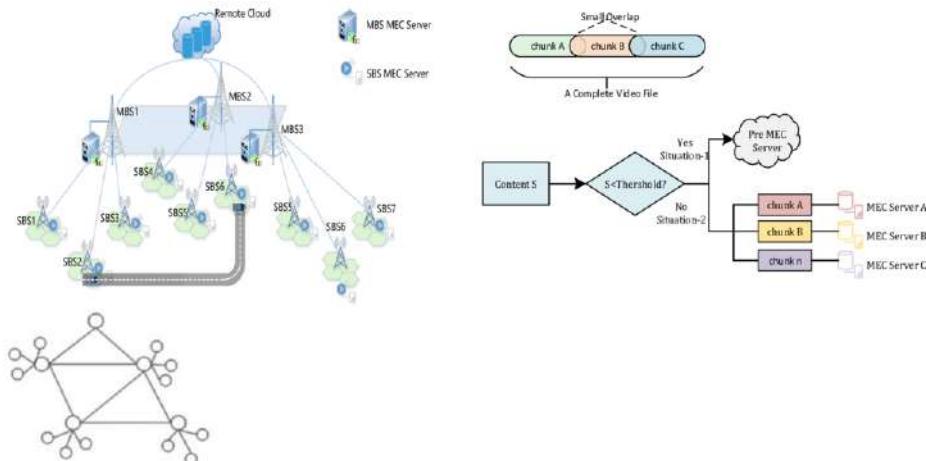


Figure 1. Illustration of (a) MEC based EON, (b) Content fetch decision and (c) Network Topology

2. The MEC Based Macro Small Cell Network Architecture

We consider a communication structure including multiple MBSs and SBSs which are all deployed with an MEC server as shown in Figure1. (a), so that they can monitor content requests sent by all users within the coverage area and make corresponding content delivery decisions. Also, there is a remote cloud. The service range of the MBS

covers all SBSs, which are deployed without overlapping with each other. From previous research [8], we know that users can only get the content from SBSs with the nearest geographic location and the video streaming is either delivered fully from the original storage place or stored separately in different MEC servers. To illustrate the problem, we can divide it into 2 situations which are shown in Figure 1. (b). Situation-1 indicates that the video is fetched from the MEC server of the original MBS and directly sent to the user. Situation-2 indicates that the video is divided into multiple chunks and pre-cached in the MEC server corresponding to the position of the vehicular user. During each user HO, the content must be migrated from the source MBS node to the new MBS node. Accordingly, the chunked video streaming should be pre-cached in the new MEC(NMEC) server before the vehicular user arrives at the new SBS cell. It should be noted that each chunked video fragment should have a small overlap to guarantee the integrity of the video streaming. There have been lots of research previously on predicting the traveling route of vehicles. For example, a convolution neural network can be applied to predict the road traffic situation in literature [9]. Therefore, we can assume that the route information of the car is predicted already so that we can mainly deal with the content delivery problem which determines the source of the content and the appropriate request scheduling process.

3. Pre Caching Algorithm Joint Handoff and Content Delivery

We have a set of MBS node $M=\{1,2,\dots,m\}$ with the maximum storage capacity S_m , and a set of SBS node $N=\{1,2,\dots,n\}$ with the maximum storage capacity S_n . The network topology $G(m,n,l)$ consists of m MBS nodes, n SBS nodes, and optical link l , as shown in Figure.1(c). The j_{th} video request is denoted as $R_j(P(t), C_j, T_{max,j})$, $P(t)$ is the vehicular users' location changing with time, while C_j is the required video's size (in bits) and $T_{max,j}$ is the maximum latency time for the request. It should be noted that the transmission latency from user to SBS node can be ignored as a relatively small constant cause the communication is wireless. The modulation level l_j is assigned to R_j according to the distance of delivery. f_j is the frequency slot to which the request needs to be allocated. B is the capacity of a frequency slot. Then the latency time LT_j for request R_j is given in Eq. (1), which is the time of transferring required content from the server storing it to the vehicular user through different servers. To determine whether a video streaming should be chunked, a threshold S_c is defined in advance according to the video size in practical.

$$LT_j = C_j / (l_j \times B \times f_j) \quad (1)$$

We propose a pre-caching algorithm joint handoff and content delivery in Table 1. Before the algorithm starts, we randomly initialize the load of each server and access the vehicular users' request to the SBS nodes at $t1$. In Line 4-8, we can get the position of each user and the MBS they access to before and after their movement. Thus, we can decide which MEC server the requiring content would be delivered to after the user moves. In Line9-11, the algorithm decides whether the content is delivered from the remote cloud or nearby MEC server. Once the content delivery method is scheduled, we need to calculate the latency time, which is shown in Procedure1 using Dijkstra

method. The distance between nodes A and B is defined as $d(A, B)$. Through our proposed algorithm, the amount of data to be transferred during the HO stage can be reduced since we deliver the chunked video segment instead of the whole one. It also should be noticed that the chunked video should have a small overlap with the ones before and after it to assure the integrity of the video. In this way, the resulting interruption time can be significantly reduced.

Table 1. Algorithm 1

Algorithm1: Pre-caching algorithm joint handoff and content delivery	Procedure1: Calculate the latency time
<p>1: Input: Content Request R_j and network topology G</p> <p>2: Output: The source of the requested content, path, and latency time of the request</p> <p>3: for each request of the vehicular users in R_j do</p> <p>4: 5: search its position in $t1$ $P(t1)$ and its connected MBS node $M1$</p> <p>6: 7: search its position in $t2$ $P(t2)$ and its connected MBS node $M2$</p> <p>8: 9: if $C_j < S_c$ then: deliver video from the remote cloud;</p> <p>10: calculate the latency time in Procedure1</p> <p>11: else:</p> <p>12: divide the video into chunks according to $P(t)$</p> <p>13: Pre-cache the chunk in MBS node M</p> <p>14: calculate the latency time in Procedure1; end if</p> <p>15: for</p>	<p>Select the closet node to $P(t1)$ as start node N</p> <p>for all node S in graph G do:</p> <p> for all neighbor nodes of S do:</p> <p> if $d(neighbour_node, N) < d(N, S) + d(S, neighbour_node)$ then:</p> <p> Update $d(N, neighbour) = d(N, S) + d(S, neighbour_node)$</p> <p> find next neighbor node; end for</p> <p> end for;</p> <p> Determine the modulation level l</p> <p> Calculate the latency time by Eq.1</p>

4. Simulation Results and Discussions

The network topology graph shown in Figure.1 (c) with 17 nodes, 4 MBS MEC servers and 12 SBS MEC servers are used for simulations. The width of the FS is set to 12.5GHz, the size of a video file is 10MB. For comparison, we deployed two baseline algorithms. One is migrating the required video from the MEC server which cached the content in time $t1$ and delivering it to the user, the other is delivering the required video straightly from the remote cloud. Since vehicular users are randomly accessed into our network, we test 10 samples to produce an overall performance of our pre-caching algorithm. The simulation result in Figure2. (a) shows that the performance of our algorithm is significantly better than the other two strategies by achieving a reduction in latency cost. In the meantime, the HO interruption time can be reduced since the required video chunks have been already cached in the MEC server corresponding to users' driving route which has been predicted before. What's more, we can observe that delivering contents from the remote cloud is most time costing for most cases. The reason is that the remote cloud is the furthest away from users, which causes more transmission time than the other two methods. Figure2. (b) indicates the server load of 4 MBS nodes, from the result we can see the load is more balanced by applying our algorithm than using the other two strategies, such that the servers' storage space utilization is higher. Consider a situation when one MEC server is almost full, it is

more feasible to accommodate a chunked video streaming than a complete one. In the meantime, the risk of content cannot be delivered successfully due to the crash of servers is reduced by the proposed algorithm.

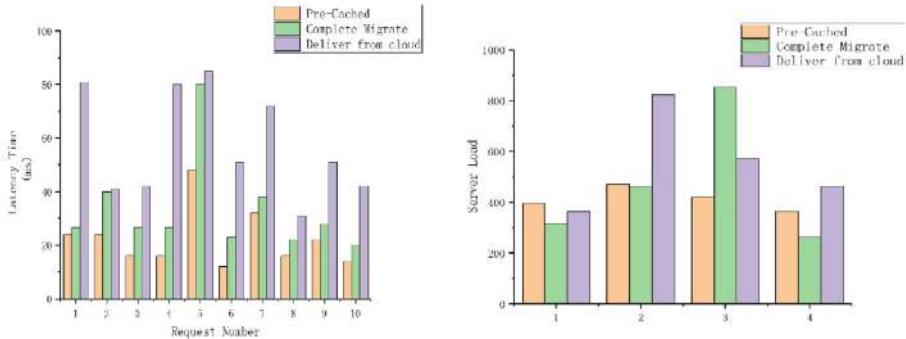


Figure 2. Calculation results of (a) Latency Time and (b) Server Load

5. Conclusion

In this paper, we propose a pre-caching algorithm joint handoff and content delivery, including pre-caching the video streaming in the corresponding server node, deciding the source of the video streaming, choosing the shortest transmission path and calculate latency time. Simulation results show the algorithm could reduce HO time and latency time at the scenario of dealing with high mobility vehicular users' requests. In the future, some more works following this paper can be considered, such as how to automatically decide the threshold S_c for each content to be delivered, and how to decide the size of the small overlap to make the resource utilization more efficient.

Acknowledge

This work is supported in part by the National Natural Science Foundation of China (Nos.61821001,61771074, 61601054).

References

- [1] Francesco Spinelli, et al, "Toward Enabled Industrial Verticals in 5G: A Survey on MEC-Based Approaches to Provisioning and Flexibility", IEEE,2020
- [2] Nandish P., et al, "Mobility Awareness in Cellular Networks to Support Service Continuity in Vehicular Users ", ICOIACT,2020
- [3] K. Poularakis and L. Tassiulas, Exploiting user mobility for wireless content delivery, International Symposium on Information Theory, Istanbul, Turkey, 2013.
- [4] H. Li and D. et, al, "Hu, Mobility prediction based seamless RAN-cache handover in HetNet", IEEE WCNC,2016.
- [5] T. Liu, J. Li, B. Kim, C.-W. Lin, S. Shiraishi, J. Xie, and Z. Han, "Distributed file allocation using matching game in mobile fog-caching service network," in Proc. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, Apr. 2018, pp.499–504.

- [6] Y. Li, P. A. Frangoudis, Y. Hadjadj-Aoul, and P. Bertin, "A mobile edge computing-based architecture for improved adaptive http video delivery," in Proc. IEEE CSCN'16, Oct. 2016, pp. 1–6
- [7] Z. Liu, et al., "Joint jobs scheduling and lightpath provisioning in fog computing micro data center networks," JOCN (2018):10(7): B152.
- [8] Zizheng Guo,et al, " Distributed Caching and Lightpath Provisioning in Multi access Edge Computing based Elastic Optical Networks" OSA,2020
- [9] Jinglin Li, et,al,"An End-to-End Load Balancer Based on Deep Learning for Vehicular Network Traffic Control", IEEE INTERNET OF THINGS JOURNAL,2019

Research and Implementation of High-Speed Data Streams Symbol Synchronization Algorithm Using Training Sequence in IMDD-OOFDM System

Dongsheng ZHANG^a, Gang ZHANG^b, Jiawei WU^c, Yunjie XIAO^c, Liang LIANG^d,
Chao LEI^a and Huibin ZHANG^{a,1}

^a*Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China*

^b*Global Energy Internet Research Institute Co. LTD, Nanjing, China*

^c*State Grid Shanghai Electric Power Company, Shanghai, China*

^d*State Grid Zhejiang Electric Power Co. LTD, Hangzhou, China*

Abstract. We propose a symbol synchronization algorithm for high-speed data streams in IMDD-OOFDM system using a training sequence. Sampling point phase offset approximately sustains within $\pm\pi/32$ and symbol synchronization deviation stabilizes within ± 0.5 sampling point in a real-time system of 1.5Gsa/s.

Keyword. Symbol synchronization algorithm, high-speed data streams, IMDD-OOFDM, training sequence

1. Introduction

With the increasing bandwidth of optical communication networks, the increasingly supported sampling rate of the physical layer and more application of high-speed data streams systems will be needed. At present, the intensity-modulation and direct-detection optical orthogonal frequency division multiplexing (IMDD-OOFDM) systems based on FPGA can be run the bit rate of 20.37 Gb/s [1]. The increasing of sampling rate poses a challenge to the symbol synchronization of system.

Symbol synchronization is the first step of prime subsequent receiving processes in IMDD-OOFDM systems. In the system, OFDM symbols are continuously transmitted by using a fixed-length frame. Therefore, the purpose of symbol synchronization is to catch the starting position of the frame's header.

In fact, Sampling Clock Offset (SCO) inevitably affects the accuracy of symbol synchronization in high-speed data streams. SCO occurs due to the frequency offset between the transmitter and the receiver [2]. Therefore, SCO leads to the sampling

¹ Corresponding author: Huibin ZHANG, the State Key Laboratory of Information Photonics and Optical Communication (IPOC), Beijing University of Posts and Telecommunications (BUPT), No.10 Xitucheng Road Haidian District, 100876 Beijing China; E-mail: zhanghuibin@bupt.edu.cn

point phase offset (SPPO) of the receiver, which accumulates into the integer sampling points deviation of the receiver over time.

In order to solve the problem of sampling point deviation caused by SCO, symbol synchronization of the low-speed data streams of IMDD-OOFDM system is performed on each frame and the length of each frame is set to be relatively short. In the above method, SPPO does not need to be considered, because SCO don't lead to integer sampling points deviation in the low-speed data streams system.

However, the frame length and sampling rate of high-speed data stream are several or even ten times higher than those of low-speed data stream, so that SCO lead to integer sampling points deviation in the high-speed data streams system. Moreover, SPPO becomes a non-negligible factor for symbol synchronization compared with low-speed data streams system. Therefore, symbol synchronization algorithm of low-speed data stream is not suitable for high-speed data stream. In summary, it is significant and meaningful to research symbol synchronization algorithm for high-speed data streams in the IMDD-OOFDM system.

In this paper, we propose a new symbol synchronization algorithm for high-speed data streams using a training sequence. The effect of SCO was eliminated by combining the proposed symbol synchronization algorithm and Frequency Offset Adjustment (FOA) in [3]. Moreover, based on FPGA, we implement the 1.5Gsa/s real-time IMDD-OOFDM system over 120km standard single mode fiber. The experimental result verifies that SPPO approximately keeps range from $-\pi/32$ to $+\pi/32$ and the deviation of symbol synchronization stabilizes within ± 0.5 sampling point. These results show that each OFDM symbol is precisely located without any integer sampling point deviation.

2. System setup and the principle of the symbol synchronization algorithm

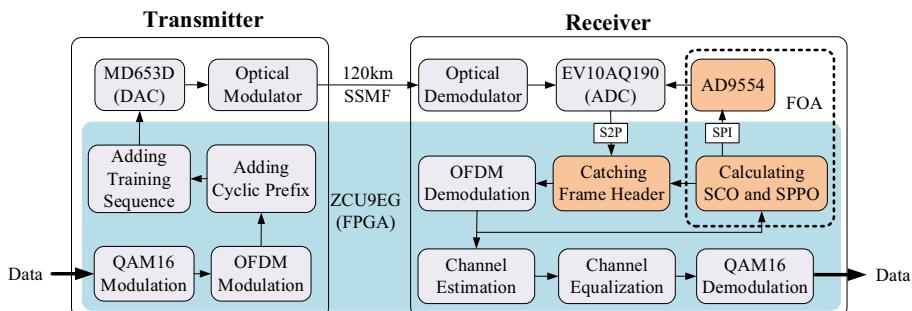


Figure 1. Block diagram of the IMDD-OOFDM system

Figure 1 shows the block diagram of the IMDD-OOFDM system over 120km standard single mode fiber to realize symbol synchronization algorithm for high-speed data streams. The part of the blue block diagram is realized in FPGA, which includes the DSP processing of the transmitter and the receiver. DAC and ADC are set to 1.5Gsa/s real-time sampling. FPGA receives serial sampling data from ADC and turns them into 40-point parallel sampling data. The length of the frame has 256 OFDM symbols. The frame is composed of 253 OFDM symbols and a frame header which is a training

sequence. Training sequence uses 3 OFDM symbols length which has 120-point data to implement symbol synchronization. OFDM modulation and demodulation are implemented with 32-point Inverse Fast Fourier Transform (IFFT) and 32-point Fast Fourier Transform (FFT). Cyclic Prefix (CP) is set to 8-point data. The length of the OFDM symbol has 40-point data after adding cyclic prefix.

Symbol Synchronization is implemented by three orange block diagrams in Figure 1. AD9554 is the clock chip that can compensate for the SCO of ADC in real-time. According to the principle of FOA in [3], the data on subcarriers No. 1, 5, and 9 of 256 OFDM symbols in each frame are used to calculate SCO and SPPO. The value of SCO is sent to AD9554 through the SPI interface. Finally, AD9554 uses this value to compensate for the SCO of ADC. It is worth noting that SPPO can reflect symbol synchronization performance. An OFDM symbol has a phase of 2π . Because the system uses 32-point IFFT and FFT, each point of the OFDM symbol represents a phase of $2\pi/32$. Therefore, if the absolute value of SPPO is greater than $\pi/16$, the symbol synchronization deviation is greater than 1 sampling point and is judged to have failed. In our proposed symbol synchronization algorithm, FOA is used to control the value of SPPO between $-\pi/32$ and $\pi/32$ nearby, which means that our symbol synchronization deviation is controlled within ± 0.5 sampling point.

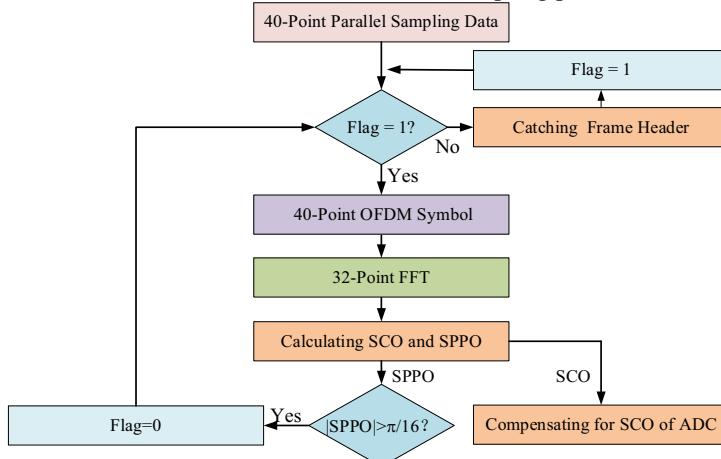


Figure 2. Flow diagram of the proposed symbol synchronization algorithm

Figure 2 shows a flow diagram of the proposed symbol synchronization algorithm in the IMDD-OOFDM system. If catching frame header succeeds, the flag is set to 1. Otherwise, the flag is set to 0. When the system starts to run, the flag is equal to 0. At this time, SPPO does not accumulate into the deviation of a sampling point. Therefore, the module of catching frame header receives 40-point parallel sampling data and utilizes the character of time-domain data of training sequence to catch frame header. When catching frame header is successful, the flag is set to 1. At this time, the system can know which sampling point in the 40-point parallel sampling data is starting position of the OFDM symbol and get complete 40-point OFDM symbols. System changes the time-domain data of OFDM symbols into the frequency-domain data of OFDM symbols through performing 32-point FFT. The module of calculating SCO and SPPO uses the frequency-domain data on subcarriers No. 1, 5, and 9 of OFDM symbols to calculate SCO and SPPO in real-time [3]. Finally, the value of SCO is sent to AD9554 through the SPI interface and AD9554 uses this value to compensate for the

SCO of ADC. At the same time, the value of SPPO can be used to determine whether the result of catching the frame header is valid. If the absolute value of SPPO is greater than $\pi/16$, which means the deviation of catching frame header is greater than 1 sampling point, the result of catching frame header is not valid. Therefore, the flag is set to 0 and the system starts catching the frame header again. If the absolute value of SPPO is less than $\pi/16$, the result of catching the frame header is always valid and the system can directly obtain complete 40-point OFDM symbols from 40-point parallel sampling data. Unlike the symbol synchronization of low-speed data streams which requires catching frame header for each frame, we proposed symbol synchronization algorithm does not require catching frame header for every frame. Therefore, the execution complexity of our catching frame header algorithm is much lower than that of low-speed data streams.

Using the correlation of time-domain data of training sequence catches frame header. We assume that the time-domain data of the training sequence consists of A, B, and C in Eq. (1).

$$A = [a_{L-1}, a_{L-2}, \dots, a_0], B = [b_{L-1}, b_{L-2}, \dots, b_0], C = [c_{L-1}, c_{L-2}, \dots, c_0], a_i = b_i = c_i, i \in [0, L-1] \quad (1)$$

We define R as the value of the correlation of A, B, and C in Eq. (2).

$$R = \sum_{i=0}^{L-1} |c_i - b_i| + |b_i - a_i| \quad (2)$$

Ideally, R is 0 in Eq. (2). Due to the existence of noise, the actual value of R is not equal to 0 but is very small. Because other OFDM symbols of the frame are not related to each other, the correlation values of these symbols are very large. Therefore, we can use R of the training sequence as judgment criteria for symbol synchronization. According to this characteristic, we calculate R between any $3 \times L$ -point sampling data in a frame and find the minimum value of R. Therefore, the sample point corresponding to this minimum value is the starting point of the training sequence. Namely, the header of each frame is found.

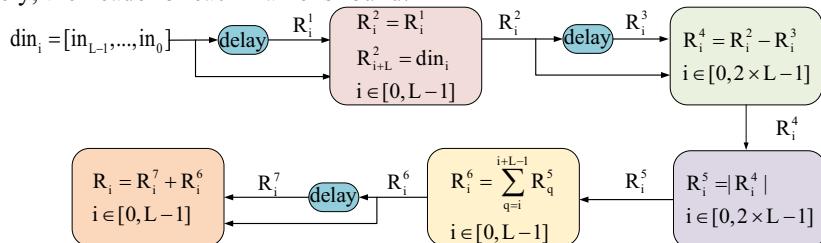


Figure 3. Eight-level pipeline architecture to calculate the R of each point in L-point parallel sampling data.

We design the eight-level pipeline architecture to calculate the R of each point in L-point parallel sampling data, as illustrated in Figure 3. The length of the OFDM symbol is L. The serial sampling data of ADC becomes L-point parallel sampling data after $1/L$ serial-to-parallel transformation. din_i is L-point parallel sampling

data. R_i is the R of each point of d_{in_i} . First-level, third-level, and seventh-level are delaying one clock cycle. Second-level is to transform two L -point parallel data into one $2 \times L$ -point parallel data. Four-level is subtraction operation. Fifth-level is the absolute value operation. Sixth-level is the sum of L -point data. Eighth-level is an addition operation. $R_i^1, R_i^2, R_i^3, R_i^4, R_i^5, R_i^6$, and R_i^7 are the results of each level.

According to the eight-level pipeline architecture, we can calculate the R of Eq. 2 for each point in L -point parallel sampling data. We suppose the length of the frame has M OFDM symbols. Each frame has $M \times L$ R according to this pipeline architecture. Therefore, the starting position of each frame can be determined by finding the minimum of $M \times L$ R . In this paper, the value of L is 40, and the value of M is 256. Therefore, each frame has 10240 R and R is the value of the correlation between each sampling point and the following 119 sampling points.

3. System equipment and results analysis

We have researched and implemented an equipment of the 1.5Gsa/s real-time IMDD-OFDM system over 120km standard single mode fiber on the PCB integrating FPGA, ADC, DAC, and AD9554, as illustrated in Figure 4. We utilize the VIVADO software to get the value of sampling point phase offset and R of Eq. (2) and analyze them.

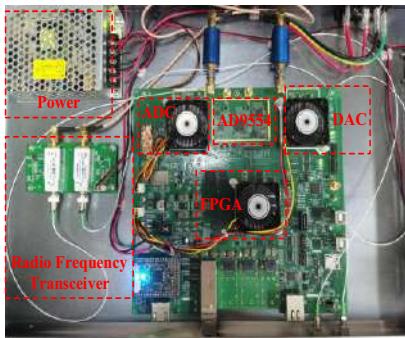


Figure 4. System equipment.

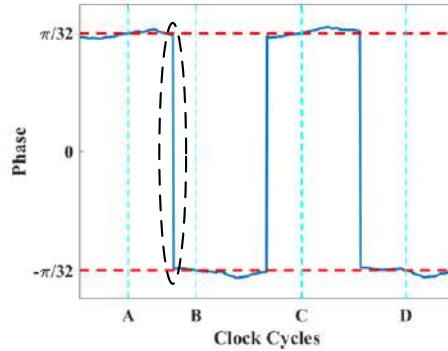


Figure 5. The curve of SPPO

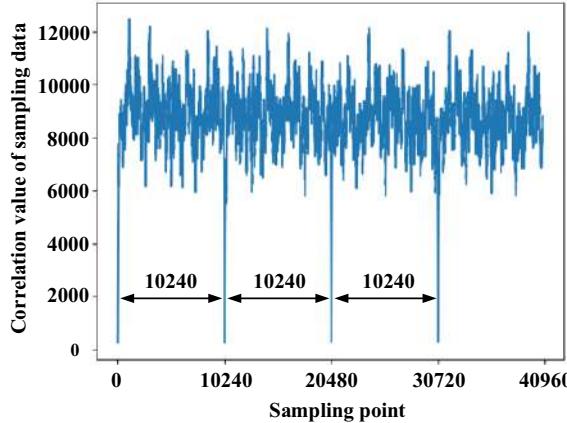


Figure 6. Frame header timing metric chart.

Figure 5 shows the curve of SPPO. When the absolute value of SPPO is equal to $\pi/32$ at A, B, C, and D, the value of SCO is sent to AD9554 through the SPI interface. After several clock cycles, the absolute value of SPPO begins to decrease. The value of SPPO is kept between positive and negative $\pi/32$ nearby. Because an OFDM symbol has a phase of 2π , 0.5 point of OFDM symbol represents a phase of $\pi/32$. Therefore, symbol synchronization deviation is controlled within ± 0.5 sampling point. It is worth noting that the curve of SPPO is a similar sine curve with an ultra-long period. Due to VIVADO software's limited computing power, a complete period of SPPO's curve can't be measured. Therefore, we discretely measure the values of peak and trough of SPPO. As shown in Figure 5, the curve jump at the black dotted line ellipse is not a true phase jump but a measurement time gap between approximately positive and negative $\pi/32$. The other two curves jump for the same reason.

Figure 6 shows the frame header timing metric chart about symbol synchronization algorithm, whose abscissa is the number of sampling points and the ordinate is the value of R of Eq. (2) between each sampling point and the following 119 sampling points. As we can see from the chart, the sampling points corresponding to the four smallest correlation values are the starting points of the four frames header. The number of sampling points per frame is 10240($256 \times 40 = 10240$) and the intervals between the sampling points corresponding to the four smallest correlation values are also 10240 points. Therefore, the starting point of each frame header is located without any integer point deviation.

4. Conclusions

We have studied and implemented a new symbol synchronization algorithm for a high-speed data streams in a 1.5Gsa/s real-time FPGA-based IMDD-OOFDM system over 120km standard single mode fiber. The system experimentally demonstrated that the sampling point phase offset was approximately controlled within $\pm \pi/32$ and symbol synchronization deviation was handled between -0.5 and +0.5 sampling points. It's worth noting that this symbol synchronization algorithm does not require symbol

synchronization for every frame and can be applied to symbol synchronization of wireless channel.

Acknowledgment

This research is supported by The Headquarters Management Science and Technology Project of State Grid Corporation of China(5700-202117190A-0-0-00).

References

- [1] Bruno, J.S., Almenar, V., Valls, J., & Corral, J.L. (2018). "Real-time 20.37 Gb/s optical OFDM receiver for PON IM/DD systems." *Optics express*, 26 15, 18817-18831.
- [2] X.Q. Jin, R.P. Giddings, E. Hugues-Salas and J.M. Tang, "Real-time experimental demonstration of optical OFDM symbol synchronization in directly modulated DFB laser-based 25km SMF IMDD systems" *Opt. Express*, vol. 18, no. 20, pp. 21100-21110, Sep. 2010.
- [3] Song, Xinyu , et al. "Research and Implementation of Frequency Offset Adjustment in Real-Time IMDD OOFDM System." *Asia Communications and Photonics Conference 2020*.

Spectrum Availability Aware Routing and Resource Allocation for Point-to-Multipoint Services in Mixed-Grid Optical Networks

Feng WANG^a, Qingcheng ZHU^b, Xiaolong LI^a, Jiaming GU^b, Zhenhua YAN^a,
Xiaosong YU^b, Yongli ZHAO^{b,1}

^aState Grid Electric Power Technical Research Institute, Yinchuan 750001, China

^bBeijing University of Posts and Telecommunications, Beijing, 100876, China

Abstract. Mixed-grid optical networks are in a migration state where fixed-grid and flex-grid optical networks coexist. To carry point-to-multipoint (P2MP) services in mixed-grid optical networks, routing and resource allocation (RRA) problems need to be solved. Once the RRA fails, services will be blocked and then influence quality of service. The minimized spectrum for satisfying the bandwidth request of services is called as a frequency block (FB). For a service, the total number of available FBs embodies the spectrum availability on a link. Because the fixed-grid and flex-grid links have different channel spacing, spectrum availability on fixed-grid and flex-grid links needs different evaluation method. We propose a RRA algorithm in mixed-grid optical networks for P2MP services by being aware of spectrum availability. The spectrum availability is evaluated according to fixed-grid and flex-grid constraints. Our proposed algorithm achieves the lower blocking probability (BP) than that of benchmark RRA algorithms according to simulation results.

Keywords. Mixed-grid optical network, point-to-multipoint, resource allocation, routing, spectrum availability

1. Introduction

With the popularity of point-to-multipoint (P2MP) services such as online conferencing and IPTV, the fixed-grid Wavelength Division Multiplexing optical network, which has a low spectrum resource utilization and lacks flexibility, is hard to carry the increasing network traffic. It is necessary for the optical network to migrate from fixed grid to flexible grid. To prevent a large amount of interference to the existing fixed-grid optical network, the fixed grid and the flexible grid are still in a state of coexistence [1]. Such optical network is called as mixed-grid optical networks. For P2MP services, the same data is propagated from one source to multiple destinations. It is important to select a routing which occupies less spectrum resources for P2MP services in mixed-grid optical networks.

¹ Corresponding Author, Yongli Zhao, Beijing University of Posts and Telecommunications, Beijing, 100876, China; E-mail: yonglizhao@bupt.edu.cn.

Much research about technologies and algorithms to carry P2MP services in optical networks has been conducted. A P2MP architecture based on digital subcarrier multiplexing has been presented, which saves remarkable cost and makes the network it enables simpler [2]. The multicast performance over spectrum elastic optical networks has been analyzed and results show the flexible spectrum allocation in elastic optical networks provides lower blocking probability for multicast compared to the ITU-T grid-based optical networks [3]. In elastic optical network, the all-optical P2MP problem has been investigated by using subtree schemes. And the proposed subtree algorithms could effectively reduce the request blocking probability by using slightly more bandwidth-variable transponders (BVTs) and better utilization of the higher-level modulation formats [4]. Flexible features of flex-grid networks are exploited for the dynamic placement of controllers enabled by virtualization of controllers [5]. A new algorithm, called mixed-grid-aware dynamic resource allocation, is proposed to solve the route, spectrum, and modulation-format allocation problem in a mixed-grid network while considering interoperability constraints [6]. A layered graph based routing and spectrum assignment algorithm for multicast is proposed by considering distance-adaptive modulation in mixed-grid optical networks [7].

In this paper, the problem of routing and resource allocation (RRA) for P2MP services is solved in mixed-grid optical networks. The proposed algorithm is aware of spectrum availability during the process of RRA with distance-adaptive modulation formats. According to the simulation results, it achieves the lowest blocking probability compared with benchmark RRA algorithms.

2. Problem Statement and Network Model

In mixed-grid optical networks, given the network topology, the RRA problem for P2MP services can be divided into two independent problems, including the routing problem and the resource allocation problem. 1) The routing problem means that when the P2MP service arrives, the routing trees need to be calculated from the source node to all the destination nodes. If the routing tree passes through the fixed-grid node and the bandwidth requirement exceeds the wavelength interval, the P2MP service needs to be supplied by multiple routing trees. 2) Resource allocation problem means that sufficient spectrum resources need to be allocated for it according to the bandwidth requirements of the P2MP service. During the allocation process, certain constraints need to be followed, including spectrum consistency, spectrum adjacency, and spectrum non-overlapping. This paper aims to solve the RRA problem by being aware of available spectrum resources for P2MP services in mixed-grid optical networks by designing appropriate algorithms, thereby reducing the blocking ratio of networks and improving network performances.

Topology $G(V, E)$ for the physical network is given, where $V = \{v_1, v_2, \dots, v_{|V|}\}$ and $E = \{e_1, e_2, \dots, e_{|E|}\}$ are the sets of fiber links and optical nodes. $e_i = (v_{si}, v_{di}, w)$, which is a three-tuple, starts from node v_{si} to node v_{di} , and its weight equals w . $s_e = \{f_1, f_2, \dots, f_F\}$ is the bit-mask set to show status of frequency slots (FSs) on link e . $f_i = 0$ if the i^{th} FS is free, otherwise $f_i = 1$; F is the total number of FSs on link e . Two types of optical nodes including fixed-grid nodes and flex-grid nodes exist in the network. Link $e_i = (v_{si}, v_{di}, w)$ is a fixed-grid link when v_{si}, v_{di} contains fixed-grid nodes; otherwise, it is a flex-grid link. A P2MP service is represented as $MR(s, D, b)$,

where s and $\mathbf{D} = \{d_1, d_2, \dots, d_{|\mathbf{D}|}\}$ are the source node and the set of destinations respectively, b is the bandwidth requirement. The minimized spectrum for satisfying the bandwidth request of services is called as a frequency block (FB). A FB can consist of several continuous FSs.

3. Spectrum Availability Aware Routing and Resource Allocation (*SAA-RRA*) Algorithm in Mixed-Grid Optical Networks

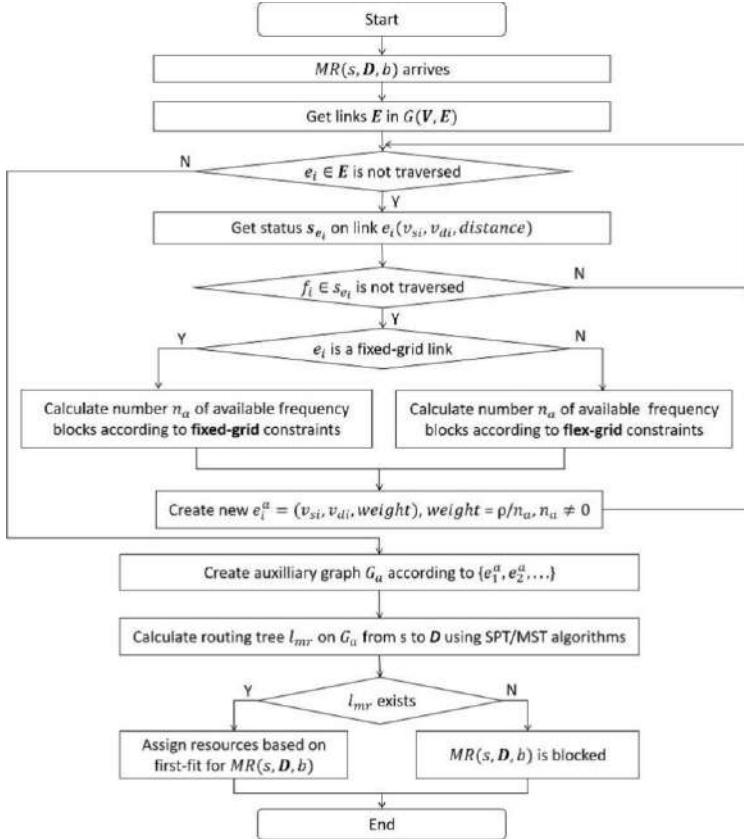


Figure 1. The flowchart of *SAA-RRA* algorithm

For satisfying the requirement of P2MP services efficiently, we propose a spectrum availability aware routing and resource allocation (*SAA-RRA*) algorithm in mixed-grid optical networks. The detailed flowchart of the proposed *SAA-RRA* algorithm is shown in Figure 1. When a P2MP service $MR(s, \mathbf{D}, b)$ arrives, firstly the link set E is got. The proposed algorithm traverses E in network topology $G(V, E)$. For the link $e_i = (v_s, v_d, distance) \in E$, the bit-mask set s_e is got which shows status of FSs on link e_i . Then, it traverses f_i on e_i to check whether the FS f_i is free. During the FS status traversing process, the important step is to calculate the number of available FBs over different grid of links for the service $MR(s, \mathbf{D}, b)$. If the e_i is from flex-grid node to flex-grid node, the e_i is a flex-grid link; otherwise, it is a fixed-grid link. If the e_i is a fixed-grid link, the number of available FBs will be calculated according to fixed-grid

constraints. And if the e_i is a flex-grid link, the number of available FBs will be calculated according to flex-grid constraints.

Based on the calculated number n_a of the available FBs over e_i , if $n_a \neq 0$, that is, available spectrums exist that meet the P2MP service's requirement, the new auxiliary link $e_i^a = (v_{si}, v_{di}, weight)$ is created. Here, the source and destination of e_i^a are the same with that of e_i but the $weight = \rho/n_a$ is different, ρ is a constant. The weight shows the number of available spectrum resources of the physical link, and the more the spectrum resources are available, the smaller the weight is. After traversing E in network topology $G(V, E)$, the auxiliary graph $G_a(V_a, E_a)$ is created according to the established auxiliary links e_i^a . Then, routing tree is calculated on G_a for $MR(s, D, b)$ from source node s to destination nodes D using *minimum spanning tree (MST)* and *shortest path tree (SPT)* algorithms. *MST* aims to reduce the total resource consumption. *SPT* aims to reduce the time delay of data in the network. At last, if the calculated routing tree from source node to destination nodes exists, enough spectrum resources are assigned for MR with First-fit and it means that the MR is carried successfully; otherwise, the MR is blocked.

4. Simulation Results and Analysis

The proposed *SAA-RRA* algorithm is simulated with different routing tree calculation method *MSP* and *SPT* algorithms respectively, and are represented by *SAA-RRA-MST* and *SAA-RRA-SPT*. The compared RRA algorithms for P2MP are based on distance-aware RRA (*DA-RRA*) algorithm. The benchmark algorithm firstly finds whether each link contains available FBs, and then compose the auxiliary topology of links with available FBs, and calculate the shortest path for the service according to the distance. The weight of link in the auxiliary topology is set to the distance of the corresponding physical link. *DA-RRA-MST* and *DA-RRA-SPT* represent the compared algorithms that use *MST* and *SPT* respectively. COST239 network topology is used in the simulation and it contains 11 nodes and 25 bi-directional links. The simulated traffic load is from

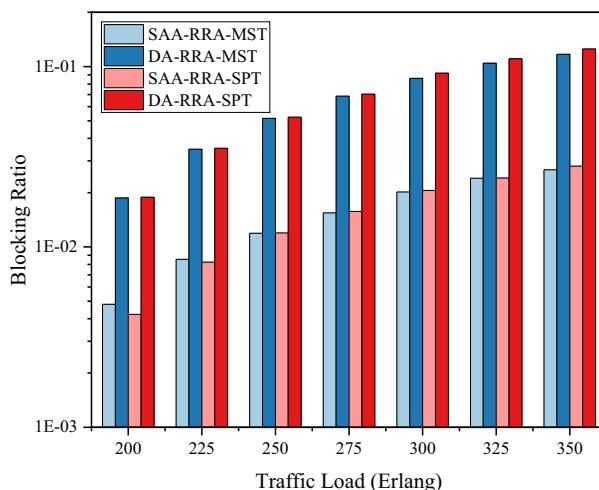


Figure 2. Simulation results in terms of blocking probability under 3 flex-grid nodes

200 Erlang to 350 Erlang, and the spacing is 25 Erlang. We do the simulations under three different scenarios including different scenarios considering different number of flex-grid nodes in the network. In detail, simulations with flex-grid node set {2,4,6}, {2,4,6,8,10}, {2,3,4,6,7,8,10,11} are conducted respectively.

Figure 2 shows the BP of the simulated algorithms under 3 flex-grid nodes in mixed-grid optical network. We can find that *SAA-RRA-MST* and *SAA-RRA-SPT* algorithm have lower blocking probability (BP) in this case, because our proposed algorithm can be aware of available spectrum that meets the P2MP service's requirement by setting the weight of links in the constructed auxiliary graph based on the number of available FBs for services, while the benchmark algorithms cannot.

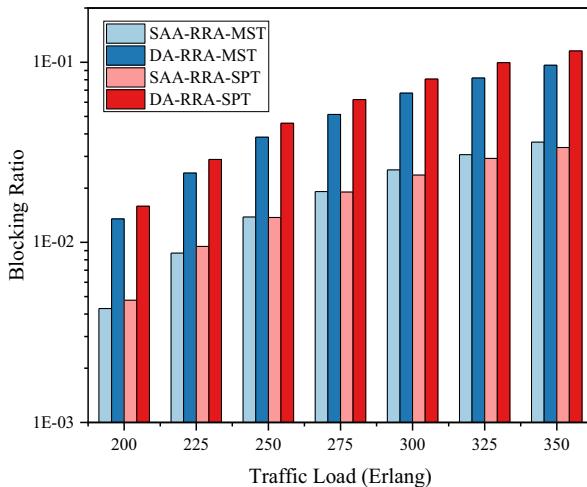


Figure 3. Simulation results in terms of blocking probability versus under 5 flex-grid nodes

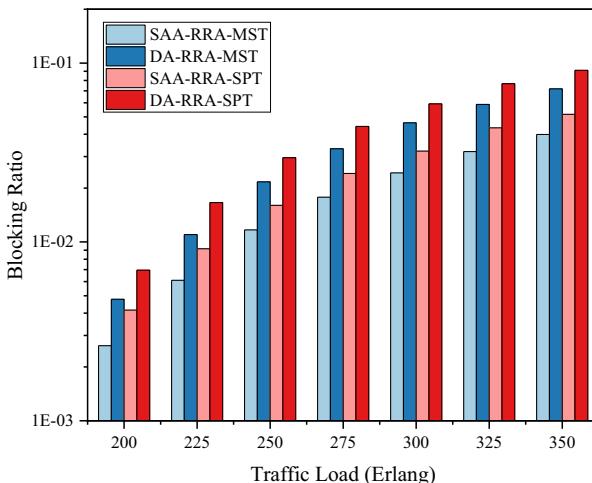


Figure 4. Simulation results in terms of blocking probability versus under 8 flex-grid nodes

Figure 3 shows the BP of some algorithms under 5 flex-grid nodes. The BP of our proposed algorithms is also the lowest in this case. The phenomenon shows that

different numbers of flex-grid nodes in the network influence the gap between the *SAA-RRA* and *DA-RRA* algorithms. The reason is that the spectrum granularity of flex-grid link is smaller than that of flex-grid link, the same spectrum can carry more services. Hence, when flex-grid nodes increase, spectrum utilization of network is improved and the BP drops, and the benchmark algorithm is less likely to block the service due to the path selection.

Figure 4 shows the BP of some algorithms under 8 flex-grid nodes. The BP of our proposed algorithms is also the lowest. In this case, compared with Figure 2 and Figure 3, the blocking probability of the *SAA-RRA*(*DA-RAA*) algorithms by using *SPT* and *MST* has a bigger gap. The reason is that the routing tree calculated by *MST* has the lower total weights compared with that calculated by *SPT* algorithms.

5. Conclusion

To solve the RRA in mixed-grid optical networks, we propose an algorithm called *SAA-RRA* for P2MP services by being aware of spectrum availability, which is simulated with both *MST* and *SPT* algorithm. Compared with the benchmark algorithm, the *SAA-RRA* has lower BP. Simulation results show that the *SAA-RRA* algorithms achieve the lowest blocking probability compared with benchmark algorithms. In cases where flex-grid nodes in physical topologies are different, the more flex-grid nodes, the more obvious the performance advantage of the *SAA-RAA* algorithm.

Acknowledgements

This work was supported by fund of science and technology research and development project of Yinchuan (2021BDE13001).

References

- [1] X. Yu et al. Migration from fixed grid to flexible grid in optical networks. *IEEE Communications Magazine*. 2015 Feb;53(2):34-43
- [2] D. Welch et al. Point-to-Multipoint Optical Networks Using Coherent Digital Subcarriers. *Journal of Lightwave Technology*. 2021 Aug;39(16):5232-5247
- [3] Q. Wang and L. Chen. Performance analysis of multicast traffic over spectrum elastic optical networks. *OFC/NFOEC*; 2012; pp. 1-3.
- [4] Z. Fan, Y. Li, G. Shen and C. C. Chan. Distance-Adaptive Spectrum Resource Allocation Using Subtree Scheme for All-Optical Multicasting in Elastic Optical Networks. *Journal of Lightwave Technology*. 2017 May;35(9):1460-1468
- [5] S. Rahman et al. Dynamic Controller Deployment for Mixed-Grid Optical Networks. 2018 Asia Communications and Photonics Conference (ACP); 2018; pp. 1-2.
- [6] T. Ahmed et al. Dynamic routing, spectrum, and modulation-format allocation in mixed-grid optical networks. *IEEE/OSA Journal of Optical Communications and Networking*. 2020 May; 12(5):79-88.
- [7] Q. Zhu, X. Yu, Y. Zhao and J. Zhang. Layered Graph based Routing and Spectrum Assignment for Multicast in Fixed/Flex-grid Optical Networks. 2020 Asia Communications and Photonics Conference (ACP) and International Conference on Information Photonics and Optical Communications (IPOC); 2020; pp. 1-3.

Low-Cost Deployment Scheme of VNF and PNF in Optical Datacenter Networks

Jianghua WEI, Xin LI, Jingjie XIN, Ying TANG, Lu ZHANG and Shanguo HUANG¹

*School of Electronic Engineering, Beijing University of Posts and Telecommunications,
Beijing 100876, China*

Abstract. In this paper, we study the problem of network function deployment while user demands are served by both virtualized network functions (VNFs) and physical network functions (PNFs). A hybrid deployment scheme of VNFs and PNFs (HDVP) is proposed. A heuristic algorithm is developed for the HDVP scheme with the aim of minimizing the total cost of VNFs and PNFs. The results show the HDVP scheme achieves lower cost than the conventional only PNFs serving user demands scheme.

Keyword. Virtualized network functions, physical network functions, hybrid deployment

1. Introduction

Virtualized network functions (VNFs) play an important role because of their flexibility and scalability in today's optical datacenter networks (ODNs) [1]. Network function virtualization (NFV) technology decouples network functions (NFs) from proprietary hardware and realizes NFs implemented by software, namely VNFs [2]. VNFs run on standard commercial-off-the-shelf (COTS), while physical network functions (PNFs) run on proprietary hardware. Proprietary hardware has high processing capacity, high operation expense (CAPEX) and high capital expenditures (OPEX), while COTS has low processing capacity, low CAPEX and low OPEX [3, 4]. VNFs are installed into virtual machines (VMs) that provide virtual running environment, increasing additional time consumption, i.e. the booting time and shutdown time of VMs, and the installation time of VNFs. VNFs have low cost but poor time performance, while PNFs have high cost but good time performance [5]. In this paper, we think the cost of VNFs and PNFs depends on the utilization time and prices. Only using VNFs to serve user demands (VNF-only) cannot meet **stringent** delay requirements, and only using PNFs (PNF-only) is too expensive. The hybrid deployment of VNFs and PNFs is a promising scheme to achieve low cost while meeting **stringent** delay requirements. So far, many VNF deployment schemes have been proposed to meet delay requirements. A cost-efficient VNF placement and scheduling (VPS) scheme in public cloud networks was proposed which incorporates the impact of several realistic factors, e.g., VNF threading attributes, VM booting time, and VNF installation time [6]. Moreover, machine learning-based

¹ Corresponding Author: Shanguo Huang, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; E-mail: shghuang@bupt.edu.cn

approaches for predicting VPS decisions were proposed to shorten processing time of applying the deployment decision to networks [7, 8]. These VNF-only deployment schemes are inapplicable to demands with **stringent** delay requirements because of the poor time performance of VNF. In [9,10], authors proposed a placement scheme in a hybrid scenario, where VNFs and PNFs jointly serve user demands. Nevertheless, they deployed PNFs as priority, increasing cost.

To guarantee low cost when serving user demands, the NFs deployment is not confined to PNF, and any NFs can be mapped into VNFs or PNFs. It is possible to deploy VNFs and PNFs to serve one user demand. In this paper, we propose a novel hybrid deployment of VNF and PNF (HDVP) scheme which uses VNFs and PNFs to jointly serve one user demand to minimize the cost while meeting stringent delay requirements. Heuristic approach is developed for the HDVP scheme. Simulation results show that the HDVP scheme can reduce the cost in ODNs.

2. Hybrid Network Function Placement and Scheduling Scheme

A ODN topology is represented as $G(V, P, E)$, where V denotes the set of VNF-capable nodes, P denotes the set of PNF-capable nodes, and E denotes the set of fiber links. A user demand r is represented as $r(a_r, s, b, F, d, t_r)$, where a_r denotes the arriving time, s and b represent the source node and destination node of demand r , respectively, d is the data size, t_r stands for the delay requirement. $F = \langle f_1, f_2, \dots, f_k \rangle$ indicates NFs of demand r , which run in order, $f_k \in F$. A network function (NF) is represented as $f_k(c, p_{f_k}^v, p_{f_k}^p)$, where c denotes the number of required CPU cores by the VNF, $p_{f_k}^v$ and $p_{f_k}^p$ are the processing capacities of the VNF and PNF respectively.

VM additional delay t_{add} is calculated in Eq. (1), where t_b is the booting time, t_i is the VNF installation time and t_s is the shutdown time.

$$t_{add} = t_b + t_i + t_s \quad (1)$$

The total delay of a user demand r , i.e., t , is calculated in Eq. (2), where t_{tran} is the transmission delay, t_{prop} is the propagation delay, and t_{proc} is the processing delay.

$$t = t_{tran} + t_{prop} + t_{add} + t_{proc} \quad (2)$$

In the proposed HDVP scheme, each NF will be mapped into a VNF or a PNF according to the NF mapping principle, i.e., a user demand r with higher delay requirement will have a larger ratio coefficient of PNFs. The basic time represents total processing time while all NFs of a user demand are mapped to VNFs, and it is calculated using Eq. (3), where $p_{f_k}^v$ is the processing capacity of the VNF. The ratio coefficient λ is defined as Eq. (4) and it determines the number of PNFs and VNFs serving user demands.

$$t_{basic} = \sum_{f_k \in F} \frac{d}{p_{f_k}^v} \quad (3)$$

$$\lambda = \frac{t_{basic}}{t_r} \quad (4)$$

Processing capacity gap ΔC represents the difference of PNF processing capacity and VNF processing capacity, which is calculated by Eq. (5). The NF with big ΔC is more likely to be mapped into the PNF. The shortest path is selected to locate VNFs and PNFs to reduce the propagation delay. The VM holding VNFs is reused to reduce the transmission delay. P^{CPU} and P^{PNF} are the prices of the CPU cores and the price of the physical hardware, respectively.

$$\Delta C = p_{f_k}^p - p_{f_k}^v \quad (5)$$

For each user demand, the cost is calculated by Eq. (6), where t_{proc}^v and t_{proc}^p are the total processing time of VNFs and PNFs serving the user demand, respectively.

$$\varphi = P^{CPU} \cdot (t_{add} + t_{proc}^v) \cdot c + P^{PNF} \cdot t_{proc}^p \quad (6)$$

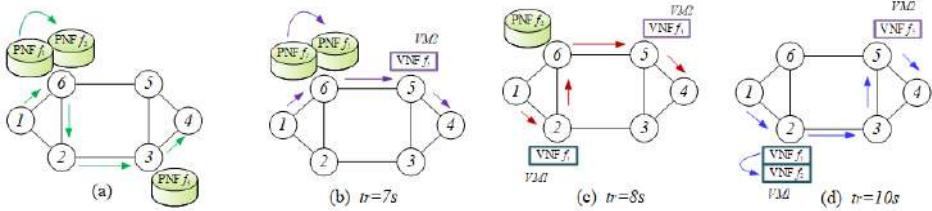


Figure 1. different deployment schemes: (a) PNF-only scheme, (b)-(d) HDVP schemes

Figure 1 illustrates the advantages of the proposed HDVP scheme (Fig. 1(b)-(d)) over PNF-only scheme (Fig. 1(a)). Supposing VNF-capable node 2 holds VM1 allocated one CPU core, VNF-capable node 3 holds VM2 allocated two CPU cores, PNF-capable node 3 holds PNFs f_3 and PNF f_1 , and node 6 holds PNF f_2 , respectively. Supposing the bandwidth between nodes is 10Gbps, and the bandwidth allocated to links of each VM and PNF is 5Gbps. The price P^{CPU} is 30 times as big as P^{PNF} . We consider one user demand r (0s, 1, 4, $\langle f_1, f_2, f_3 \rangle$, 2GB, t_r), where f_1 (1, 1 Gbps, 5 Gbps), f_2 (1, 1 Gbps, 10 Gbps), and f_3 (2, 2 Gbps, 5 Gbps). As shown in Fig. 1(a), when r arrives in the network, it incurs the transmission delay $1G/10Gbps=0.1s$. After PNF f_1 processes the data of r (processing delay is $1G/5Gbps=0.2s$), the data will be transmitted to PNF f_2 and the processing delay is $1G/2Gbps=0.5s$. The delay calculations for f_2 and f_3 are similar. The total time consumption is 1.8s, and the total cost is 24.

In the proposed HDVP scheme as shown in Fig. 1(b), $t_r = 7s$. The basic time consumption $t_{basic} = 2/1 + 1/2 + 2/2 = 5s$. The gaps of f_1 , f_2 , and f_3 are $\Delta C_{f_1} = 5 - 1 = 4Gbps$, $\Delta C_{f_2} = 10 - 1 = 9Gbps$ and $\Delta C_{f_3} = 5 - 2 = 3Gbps$. So the NF order of PNF mapping precedence is f_2 , f_1 , and f_3 . The ratio coefficient $\lambda = 71.4\%$, according to the NF mapping principle I in Table 1, the number of PNF is 2. We deploy f_1 and f_2 into PNFs and deploy f_3 into a VNF. The delay calculations for f_1 and f_2 are similar to that of PNF-only scheme. For f_3 , both PNF processing delay and VM additional delay are calculated. The time consumption is 4.7s, and total cost is 22.2. We change t_r

to 8s and 10s, and the deployment strategies are shown in Fig. 1(c) and (d), respectively. Table 2 records delay and cost between the HDVP scheme and the PNF-only scheme.

Figure 1(d) shows the results of the HDVP scheme are the same as that of VNF-only deployment scheme when $t_r \geq 8.2s$. If $t_r < 8.2s$, VNF-only cannot meet the delay requirement. This example shows that the HDVP scheme can reduce more cost than that of the PNF-only deployment scheme while meeting stringent delay requirements.

Table 1. NF mapping principles

The proportion of PNF	λ -I	λ -II	λ -III
0	[0, 55%]	[0, 70%]	[0, 40%]
1/3	[55%, 70%]	[70%, 80%]	[40%, 60%]
2/3	[70%, 85%]	[80%, 90%]	[60%, 80%]
1	[85%, 100%]	[90%, 100%]	[80%, 100%]

Table 2. Result comparisons between two schemes

Scheme	delay requirement t_r	PNF number	cost	delay
PNF-only	5	3	24	1.8
HDVP	7	2	22.2	4.7
	8	1	12.3	7.2
	10	0	9.3	8.2

3. Heuristic Algorithm for HDVP

We propose a heuristic approach for the HDVP scheme to minimize the cost of VNFs and PNFs. The HDVP algorithm is detailed as follows. Q indicates the NF set of demand r , which is sorted in increasing order of ΔC , $q \in Q$, Q_v and Q_p indicate the sets of VNFs and PNFs for demand r respectively, β indicates the number of PNFs, $\beta = |Q_p|$, $p_{f_k}^v$ and $p_{f_k}^p$ indicate the VNF processing capacity and PNF processing capacity respectively, and k indicates the index of NF f_k in F , $1 \leq k \leq |F|$.

Specifically, we first initiate basic time t_{basic} and set Q , Q_v , and Q_p . Next, we determine the NFs of demand r . From line 3 to 6, VNF processing capacity and PNF processing capacity of each NF are obtained, basic processing time t_{basic} is updated, and processing capacity gap is calculated. In line 8, parameter λ is calculated. Then, according to the NF mapping principle, checking λ which level it belongs to, and the number of PNFs β is obtained. In lines 9-10, NFs are added into set Q in the descending order of ΔC . Set Q is divided into two parts. The former is set Q_p , whose size is equal to β and the latter is set Q_v . From line 12 to 19, each NF f_k of set Q_v is mapped into a VNF. If adjacent NFs in the set F are both in set Q_v and require the same number of CPU cores, they are located in a same VM. Else they are located in a same node. If $k = 1$, or NF f_{k-1} is in set Q_p , a new VM is used for f_k . PNFs are employed to serve set Q_p in line 21. Total data size processed at a certain moment by a PNF should be smaller than its maximum throughput to avoid congestion. In line 23-24, light-paths are calculated between the locations where two consecutive VNF instances are hosted and the total delay is calculated. In line 25-28, we check whether the delay requirement of the user demand is satisfied or not.

Algorithm: Heuristic algorithm for HDVP**Input:** demand r (a_r, s, b, F, d, t_r) and NF mapping principle.**output:** the cost of VNFs and PNFs φ , and the PNF numbers β .

1. Initialize basic time $t_{basic} = 0$, $\mathcal{Q} = \emptyset$, $\mathcal{Q}_v = \emptyset$ and $\mathcal{Q}_p = \emptyset$.
2. Find the set F of demand r , i.e. $F = (f_1, f_2, \dots, f_{|F|})$
3. **for** each NF $f_k \in F$, **do**
4. Get the VNF processing capacity $p_{f_k}^v$ and PNF processing capacity $p_{f_k}^p$ respectively.
5. $t_{basic} += d / p_{f_k}^v$
6. $\Delta C_{f_k} = p_{f_k}^p - p_{f_k}^v$
7. **end for**
8. Calculate the parameter $\lambda = t_{basic} / t_r$, check the level of λ in NF mapping principle and calculate β (i.e., the number of PNFs).
9. Sort f in ΔC descending order, and put the results into the set $\mathcal{Q} = (q_1, q_2, \dots, q_{|F|})$.
10. $\mathcal{Q}_p = (q_1, \dots, q_\beta)$, $\beta = |\mathcal{Q}_p|$; $\mathcal{Q}_v = \mathcal{Q} - \mathcal{Q}_p$.
11. **for** each NF f of demand r **do**
12. **if** NF $f_k \in \mathcal{Q}_v$ **then**
13. **if** the index of NF $k > 1$ and $f_{k-1} \in \mathcal{Q}_v$ **then**
14. **if** the required CPU cores are the same **then**
15. place f_k to the VM holding f_{k-1} .
16. **else**
17. place f_k to the same node of f_{k-1} .
18. **else**
19. place f_k to the VM with c CPU cores
20. **else**
21. place the NF to a PNF, whose throughput is enough to process all data size at the time.
22. **end for**
23. Calculate the light-path between the datacenters hosting f_k .
24. Calculate the delay $t_{total} = t_{tran} + t_{prop} + t_{add} + t_{proc}^v + t_{proc}^p$ and check whether the latency requirement is satisfied.
25. **if** $t_{total} \leq t_r$ **then**
26. calculate the cost $\varphi = P^{cpu} \cdot (t_{add} + t_{proc}^v) \cdot c + P^{PNF} \cdot t_{proc}^p$.
27. **else**
28. mark it blocked.

4. Results and Analysis

We evaluate the HDVP scheme in the US Backbone network topology presented in Fig. 2. According to NF mapping principles (I), (II) and (III) in Table 1, the corresponding experimental results of the HDVP scheme are represented as HDVP (I), (II) and (III). The price ratio of VNFs and PNFs is set to 1/10. Demands with different user nodes are randomly generated. We adjust delay requirement scales (ddl_s) to change demands. In this experiment, the blocking rate of VNF-only scheme is greater than 25%, so VNF-only scheme is not considered. PNF-only scheme is deployed as comparison. We compare the cost with different NF mapping principles and different delay requirement scales. Then we focus on the PNF usage.

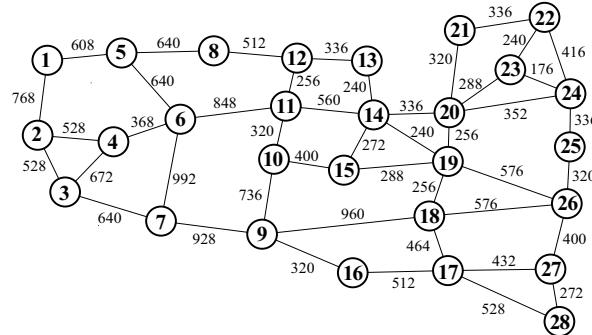


Figure 2. Network topology

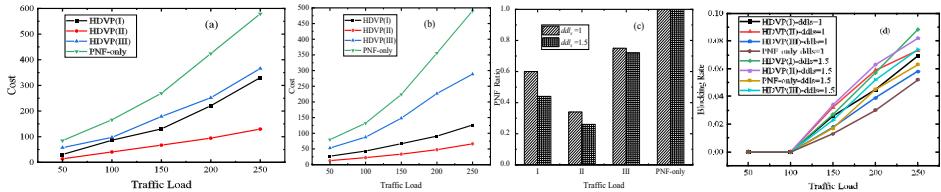


Figure 3. Numerical results (a) cost and PNF ratio ($ddl_s = 1$) (b) cost and PNF ratio ($ddl_s = 1.5$) (c) PNF usage (d) blocking rate

Figure 3(a) and (b) show the cost increases with traffic load for each NF mapping principle at $ddl_s = 1$ and at $ddl_s = 1.5$ respectively. Figure 3(c) shows the PNF usage ratio of different schemes at $ddl_s = 1$ and at $ddl_s = 1.5$ respectively. Figure 3(d) show the blocking rate increases with traffic load at $ddl_s = 1$ and at $ddl_s = 1.5$ respectively. Each HDVP scheme achieves slightly higher blocking rate compared with the PNF-only scheme at the same ddl_s .

At the same traffic load, the cost of each HDVP scheme is smaller than that of PNF-only scheme and the cost of the HDVP scheme (II) is the smallest. For different delay requirements, the cost of each scheme is smaller at $ddl_s = 1.5$ than the cost at $ddl_s = 1$, and the PNF ratio of each HDVP scheme is smaller at $ddl_s = 1.5$ than the PNF ratio at $ddl_s = 1$. For the same delay requirement, the HDVP scheme (II) outperforms the HDVP scheme (I), the HDVP scheme (III) and the PNF-only scheme. The value of ddl_s is bigger in Fig. 3(b) than that in Fig. 3(a), i.e., the delay requirement of the demand is lower in Fig. 3(b) than that in Fig. 3(a). In the same scheme with the same load, the bigger ddl_s , the smaller PNF proportion, and the smaller cost. Compared with the PNF-only scheme

at $ddl_s = 1$, the HDVP scheme (I) (II) and (III) reduce PNF ratio above 40%, 66%, and 25%, and reduce cost above 43.0%, 75.1%, and 32.4%, respectively. Compared with the HDVP scheme (I) and (III) at $ddl_s = 1$, the HDVP scheme (II) reduces PNF ratio above 43.3%, and reduces cost above 47.4%.

5. Conclusion

We propose a hybrid deployment of VNFs and PNFs to minimize cost consumption while meeting stringent delay requirements. In this study, we focus on the problem of dynamic deployment of VNFs and PNFs in ODNs where the number of PNFs and VNFs serving a user demand is determined by the ratio of its basic time to its delay requirement. Numerical results indicate that the proposed HDVP scheme can reduce the cost of VNFs and PNFs significantly compared with the PNF-only scheme.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Nos. 61701039 and 61821001) and the project was supported by Fund of State Key Laboratory of IPOC (BUPT) (No. IPOC2021ZT15), P. R. China.

References

- [1] W. Xia, et al., “A Survey on Data Center Networking (DCN): Infrastructure and Operations,” *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 640-656, 2017.
- [2] B. Yi, et al., “A comprehensive survey of Network Function Virtualization”, *Computer Netw.*, vol. 133, pp. 212-262, 2018.
- [3] H. Ko, et al., “Optimal middlebox function placement in virtualized evolved packet core systems”, in *17th APNOMS 2015*, Busan, Korea(South), pp. 511–514.
- [4] D. Bhamare, et al., “Optimal virtual network function placement in multi-cloud service function chaining architecture”, *Computer Commun.*, vol. 102, pp. 1–16, 2017.
- [5] ETSI, “Network Functions Virtualisation-Introductory White Paper ,” (Issue 1),2012
- [6] T. Gao, et al., “Cost-Efficient VNF Placement and Scheduling in Public Cloud Networks,” *TCOM*, vol. 68, no. 8, 4946-4959, Aug. 2020.
- [7] S. Lange, et al., “Machine Learning-based Prediction of VNF Deployment Decisions in Dynamic Networks,” in *20th APNOMS 2019*, Matsue, Japan, pp. 1-6.
- [8] S. Park, et al., “Machine Learning-based Optimal VNF Deployment,” in *21st APNOMS 2020* , Daegu, Korea (South) , pp. 67-72.
- [9] J. Zhang, X. Lu and D. K. Panda, “Performance Characterization of Hypervisor-and Container-Based Virtualization for HPC on SR-IOV Enabled InfiniBand Clusters,” 2016 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Chicago, IL, 2016, pp. 1777-1784.
- [10] J. Cao, et al., “VNF Placement in Hybrid NFV Environment: Modeling and Genetic Algorithms,” in IEEE 22nd ICPADS 2016 , Wuhan, China, pp. 769-777.

Complex-Coefficient Microwave Photonic Filter Based on Orthogonally Polarized Optical Single-Sideband Modulation

Jinwang QIAN ^a, Junling SUN ^a, Pengge MA ^{a,*}, Xinlu GAO ^b and Shanguo HUANG ^b

^a Zhengzhou University of Aeronautics, Zhengzhou 450046, China

^b Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. A complex-coefficient microwave photonic filter with continuous tunability is proposed and experimentally demonstrated. The filter taps are based on a 360° tunable microwave photonic phase shifter, which is realized by orthogonally polarized optical single-sideband (OSSB) modulation. The experimental results are shown and regarded as good performance for the proposed filter. The phase shift for the two taps covers a full 360° range from 8 GHz to 26 GHz. Frequency responses with different center frequency are measured within 20-21 GHz with the full free spectral ranges (FSRs) of 185 MHz and 285 MHz, respectively.

Keywords. Microwave photonic filter, complex coefficient, microwave photonic phase shifter

1. Introduction

Microwave photonic filter (MPF) has important applications in the processing of microwave signals due to its inherent advantages, including broad bandwidth, high tunability, and immunity to electromagnet interference [1]. Complex-coefficient filters are able to keep the FSR and shape unchanged while tuning the center frequency, so they have the most potential for applications among the MPFs with different coefficients. Many MPFs with complex coefficients have been reported [2-6] including those based on SSB modulation [2], stimulated Brillouin scattering (SBS) [3], fiber Bragg grating (FBG) [4], programmable photonic processor [5]. However, the problem of bias drift is existed in the scheme based on SSB modulator. The structure of MPF based on SBS is complex since it needs two or more modulators. The key limitation associated with FBG-based method is the limited tuning range. Furthermore, these schemes utilize two or more laser sources, making the system costly and complicated. In [6], a complex-coefficient MPF using one single optical source is proposed. However, the phase difference between the orthogonally polarized signals should be maintained at $\pi/2$ by adjusting bias voltage of the polarization modulator, and this condition is not easy to implement in an experiment.

In this paper, we propose a two-tap complex-coefficient MPF using a polarization-dependent dual-parallel Mach-Zehnder modulator (PD-DPMZM). Compared with the

* Corresponding author: Pengge Ma; E-mail: mapenge@163.com.

existing schemes, the proposed system requires only one optical source and one modulator. And the system structure is not complicated and extendable to more taps without adding additional laser source. By simply controlling a polarization controller (PC), the MPF is continually tunable in the full FSR without changing the filter shape and FSR.

2. Principle

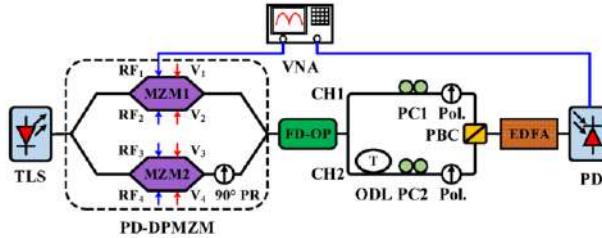


Figure 1. The schematic diagram of the proposed tunable MPF.

Figure 1 shows the proposed MPF. An optical carrier from a tunable laser source (TLS) is fed into a PD-DPMZM. A radio frequency (RF) signal generated by a vector network analyzer (VNA) is applied to one arm of the child MZM1 to accomplish a carrier-suppressed dual-sideband (CS-DSB) modulation. The Child MZM2 has no RF signal and is biased at its maximum transmission point. Then the modulated signal is input to a Fourier-domain optical processor (FD-OP) to separate the upper sideband into the upper path and the lower sideband into the lower path. By doing this, an orthogonally polarized OSSB signal can be obtained at each path. Considering small signal modulation only, the output optical fields at the upper path and lower path are respectively given by

$$E_{upper}(t) = \begin{bmatrix} E_{x1} \\ E_{y1} \end{bmatrix} \propto \frac{1}{4} E_0 \begin{bmatrix} J_1(\beta) e^{j(\omega_0 + \omega_m)t + j\frac{\pi}{2}} \\ e^{j\omega_0 t} \end{bmatrix} \quad (1.1)$$

$$E_{lower}(t) = \begin{bmatrix} E_{x2} \\ E_{y2} \end{bmatrix} \propto \frac{1}{4} E_0 \begin{bmatrix} J_1(\beta) e^{j(\omega_0 - \omega_m)t + j\frac{\pi}{2}} \\ e^{j\omega_0 t} \end{bmatrix} \quad (1.2)$$

Where E_0 is the amplitude of the optical carrier; ω_0 and ω_m are the angular frequency of the optical carrier and RF signal, respectively; $J_1(\beta)$ is the first-order Bessel function of the first kind.

Two PCs are used to tune the phase difference between E_x and E_y in the two paths, respectively. The two OSSB signals are then polarization-combined by the polarizers (Pol.), respectively. If the combined signals are injected into a photodetector (PD), two electrical currents can be obtained

$$i_{upper}(t) \propto \sin(2\alpha_1) \cos(\omega_m t + \phi_1 + \frac{\pi}{2}) \quad (2.1)$$

$$i_{lower}(t) \propto \sin(2\alpha_2) \cos(\omega_m t - \phi_2 - \frac{\pi}{2}) \quad (2.2)$$

Where ϕ_1 and ϕ_2 are the phase differences between E_x and E_y in the two path, respectively; α_1 and α_2 are the polarization angles between the principal axis of the PD-DPMZM and the upper and lower polarizer, respectively. From Eq. (2), the two recovered RF signals can realize full 360° phase shift by adjusting ϕ_1 and ϕ_2 .

The signals after FD-OP propagate through different time delays by introducing an optical delay line (ODL) between the two optical paths. The recovered RF signal after PD is a summation of $i_{upper}(t)$ and $i_{lower}(t)$, as

$$i(t) = i_{upper}(t) + i_{lower}(t) = \sum_{n=1}^2 a_n \cos[\omega_m t + \phi_n + \omega_m(n-1)T] \quad (3)$$

Where a_n is the RF power in each tap; ϕ_n is the RF phase shift in each tap, which can be independently adjusted from 0 to 2π by tuning the PC at each path; T is the relative time delay between the two taps.

The transfer function of the proposed MPF is given by

$$H(\omega_m) = a_1 + a_2 e^{j\omega_m T} e^{j\Delta\phi} \quad (4)$$

Where $e^{j\Delta\phi}$ is the complex coefficient, and $\Delta\phi$ is the phase difference between the two taps. The frequency response of the proposed MPF can be rewritten as

$$H'(f) = H(f + \Delta\phi/2\pi T) \quad (5)$$

Where f is the frequency of the RF signal.

In Eq. (5), the center frequency change of $\Delta\phi/2\pi T$ is realized by adjusting the phase difference between the two taps. The proposed MPF is continually tunable over the full FSR, since the phase shift covers full 360° range.

3. Experimental Results and Discussion

A proof-of-concept experiment is implemented based on the setup shown in Figure 1. A 10.0 dBm optical carrier at ~ 1550.0 nm emitting from a TLS (Coherent Solutions MTP1000) is fed to the PD-DPMZM (Fujitsu FTM7980EDA). A -25 dBm RF signal generated by the VNA (Agilent 8722ES) is amplified by a broadband electrical amplifier (OA4MVM3), which has a wide frequency range from 30 kHz to 45 GHz with a gain of 27 dB. The amplified RF signal is applied to the upper arm of child

MZM1 of PD-DPMZM, and MZM1 is biased at its MITP to realize CS-DSB modulation. Child MZM2 is biased at its MATP without adding a RF signal. The modulated signal is sent to a FD-OP (Finisar Waveshaper 4000s). An orthogonally polarized OSSB signal is obtained at each path.

We measure the phase shift property of the recovered RF signal in the upper path while the lower path is off. The orthogonally polarized optical carrier and sideband are combined by the PBC and then amplified by an Erbium Doped Fiber Amplifier (EDFA, Conquer KG-EDFA-P) before sending to a PD (Optilab PD30). The output power of the EDFA is 1.3 dBm. The power-frequency response and phase-frequency response of the recovered RF signal within 8-26 GHz are measured by the VNA. The RF phase shifts are recorded from -180° to 180° with a step of 45° by tuning PC1, as shown in Figure 2(a). It indicates that a continuously tunable phase shifter with stable performance over a wide bandwidth is obtained. Figure 2(b) shows that the relative power variations are within ± 2 dB for different phase shifts.

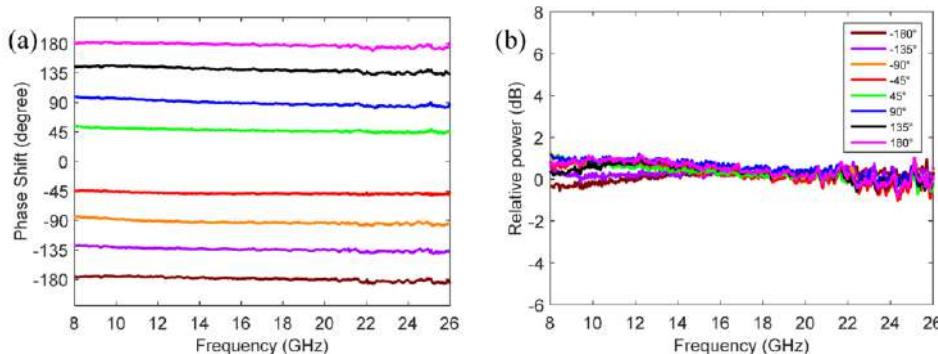


Figure 2. (a) Phase shifts of the RF signal within 8-26 GHz, (b) Relative power variations for different phase shifts.

In order to implement a two-tap MPF, an ODL is inserted into the lower path. The two orthogonally polarized OSSB signals in the two paths go through different time delays and then mutually interfere in the PD after the EDFA. Figure 3 shows the frequency responses of proposed filter with different center frequency within 20-21 GHz. As shown in figure 3(a), the FSR is maintained at 185 MHz which is close to the theoretical value of 187 MHz when the relative time delay T is about 5.35 ns.

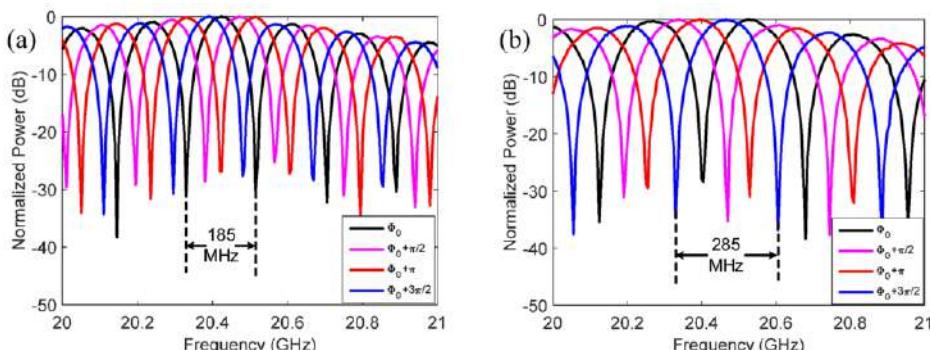


Figure 3. Measured frequency responses of the proposed MPF with FSRs of (a) 185 MHz and (b) 285 MHz.

To demonstrate the tunability of the FSR, different ODLs are used in the lower path to introduce different time delays between the two paths. When the relative time delay is changed to about 3.54 ns, the measured FSR of the filter is nearly 285 MHz, as shown in Figure 3(b), which is approach to the theoretical value of 282 MHz. From Figure 3, the proposed MPF is tunable over the full FSR without affecting the FSR. Meanwhile, by inserting different ODL between the two paths, the FSR can be changed.

4. Conclusions

We propose and demonstrate a complex-coefficient MPF based on orthogonally polarized optical single-sideband modulation. The proposed MPF is continuously tunable in the full FSR without affecting the filter shape and FSR. The proposed MPF is simple, easily tunable, and extendable to more taps.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (Nos. U1833203); and in part by the Key Scientific and Technological Projects in Henan Province (Nos. 212102210519).

References

- [1] Minasian R A, Chan E, Yi X. Microwave photonic signal processing. *Optics Express*, 2013, 21(19):22918-22936.
- [2] Sagues M, Loayssa A, Capmany J, et al. Tunable complex-coefficient incoherent microwave photonic filters based on optical single-sideband modulation and narrow-band optical filtering. *Optical Fiber Communication Conference*. 2007: OWU5.
- [3] Wang W T, Liu J G, Mei H K, et al. Microwave photonic filter with complex coefficient based on optical carrier phase shift utilizing two stimulated Brillouin scattering pumps. *IEEE Photonics Journal*, 2015, 7(1): 1-8.
- [4] Xu E, Yao J. Frequency- and notch-depth-tunable single-notch microwave photonic filter. *IEEE Photonics Technology Letters*, 2015, 27(19): 2063-2066.
- [5] Xu D, Cao Y, Tong Z, et al. A continuously tunable microwave photonic notch filter with complex coefficient based on phase modulation. *Optoelectronics Letters*, 2017, 13(1): 13-15.
- [6] Zhang C, Yan L S, Pan W, et al. A tunable microwave photonic filter with a complex coefficient based on polarization modulation. *IEEE Photonics Journal*, 2013, 5(5):5501606-5501606.

Node Importance Based Protection in Power-Grid Optical Backbone Communication Networks

Xiaobo LI^a, Guoli FENG^a, Run MA^a, Lu LU^b, Kaili ZHANG^b, Xinnan HA^a, Wenbin WEI^a, Ning WANG^a, Xiaosong YU^b and Yongli ZHAO^{b,1}

^aState Grid Ningxia Electric Power Co. LTD, Information communication corporation, Yinchuan 750001, China

^bBeijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. Power-grid optical backbone communication network is a special communication network serving for power system. With the development of new internet technology, there are more and more services carried by power-grid optical backbone communication networks. It plays an important role in the protection of nodes, especially important nodes which often carry important information of the network, when the network is under heavy traffic load. Hence, to solve this problem, we propose the concept of node importance and design a node importance-based protection algorithm under heavy traffic load scenario in power-grid optical backbone communication networks. Simulation results show that the proposed node importance based protection algorithm can obviously reduce blocking probability of the important nodes and improve the performance of the entire network in terms of blocking probability.

Keywords. Node importance, resource shortages, power grid, optical networks

1. Introduction

The security and stability of power-grid optical backbone communication networks play an important role in the whole power system. And the current power grid construction is constantly improving. The power communication system is becoming more and more complex, the scale of the power communication network continues to expand [1]. However, as the key service of power system increasingly depends on communication and information system, power system is more vulnerable to network attack [2]. In addition, in the scenario of heavy traffic, once a fault occurs in the network, it will bring more serious losses. They all pose challenge to the survivability of the power-grid optical backbone communication network. At the same time, with the development of big data, cloud services, artificial intelligence and other emerging technologies, networks can provide more diverse types of services [3]. And the establishment of core nodes, such as data center and control center, makes some nodes more important than others, and the traffic originating from or destined to these nodes needs to be guaranteed first.

¹ Corresponding Author, Yongli Zhao, Beijing University of Posts and Telecommunications, Beijing, 100876, China; E-mail: yonglizhao@bupt.edu.cn.

For the evaluation of node importance, several methods have been proposed. For example, in Ref [4], the authors proposed a weighted K-shell decomposition method, which combines the nodes in the network, the degree index and the number of iterations to rank the importance of nodes. Ref [5] analyzes the importance of nodes based on four indexes: feature vector centrality, proximity centrality, intermediary centrality. In this paper, we propose a new definition of node importance, which takes into account the nodes carrying important services. In addition, there are also many research on the survivability of network. In Ref [6], a flexible differentiated protection mechanism is proposed, which can switch between different path protection schemes (such as no protection, shared path protection and 1 + 1 protection) to adapt to different service availability. In Ref [7], the authors present ILP formulations to solve the multi-layer survivability problem. However, the protection of some nodes under heavy traffic load has not been studied yet.

To solve this problem of the protection of some nodes under heavy traffic load, in this paper, we propose the concept of node importance (NI) in *Section 2*, which gives different importance levels for network nodes. The corresponding routing and wavelength allocation strategy is designed in *Section 3* to ensure that nodes with different importance can be protected with higher requirements. *Section 4* gives the simulation settings and results which indicate that proposed node importance based protection (NIP) scheme can obviously improve the performance of important nodes in terms of blocking probability.

2. Problem Statement

Survivability mechanism includes protection mechanism and recovery mechanism. The basic idea of the protection mechanism is to reserve a part of redundant resources as the standby resources in advance. When the links or nodes fail, the main system affected by the failure will be quickly switched to the standby system. The protection mechanism can be divided into link protection, path protection and multi segment protection. But node-based protection schemes have not been widely studied. The algorithm proposed in this paper is a node-based protection strategy. Firstly, the concept of node importance is defined. For node importance, some studies define the node with the highest degree as important nodes, and some studies define the node with the largest traffic as important nodes. However, with the establishment of core nodes such as data center and control center, these nodes may not only carry the most traffic, but also often carry important service information in the network. Therefore, this paper defines those nodes responsible for core service as important nodes and gives priority to their protection.

In addition, the current routing and wavelength allocation schemes in power-grid optical backbone communication networks generally use Dijkstra algorithm or KSP algorithm to select the shortest route, and then use FF algorithm to select the wavelength that meets the requirements. This scheme is relatively simple to implement, but it ignores the traffic load distribution in the network and the load balance of the link. It is not flexible enough to deal with some problems in heavy traffic load scenarios. The traditional schemes treat all nodes equally, but in practice, the importance of network nodes is often not the same. In this paper, the nodes responsible for core services in the network are defined as important nodes. Since these nodes carry important services information in the network, the services originating from or sending to these nodes need to be guaranteed in priority, so we set higher node importance (NI) for these nodes. When

the network resources are abundant, the KSP algorithm is used to find the path and the FF algorithm is used to allocate the wavelength. When the adjacent links and secondary adjacent links of important nodes are short of resources, certain spectrum resources need to be reserved for the services of important nodes. At this time, only the services from / to important nodes are carried.

3. Node Importance based Protection (NIP) scheme

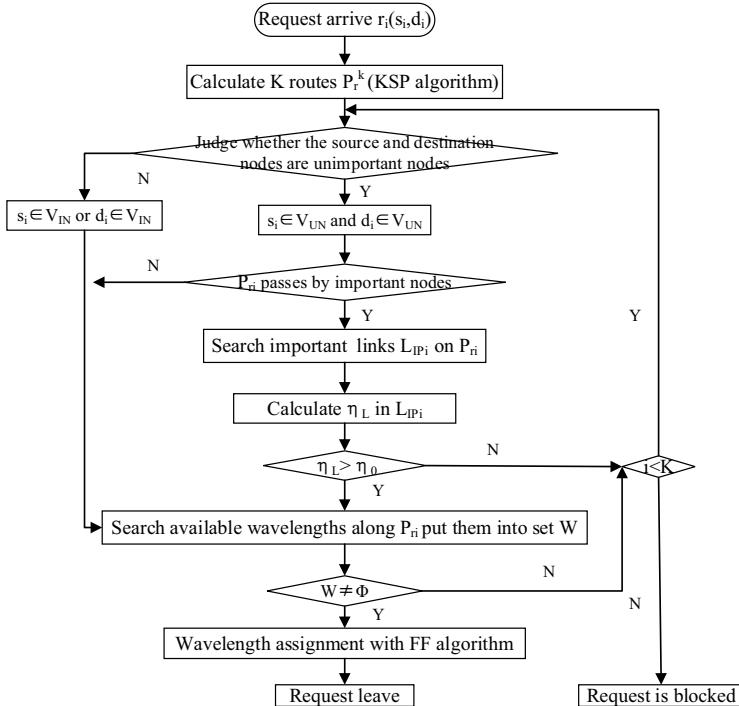


Figure 1. The flow chart of node importance based protection (NIP) scheme.

The flow chart of node importance based protection scheme is shown in Figure 1. In our algorithm, node importance has two levels: high and low. The nodes of different importance are called important nodes and unimportant nodes, respectively. When the network resources are sufficient, or the service only passes through low-importance nodes, we only need to allocate according to the KSP and First-fit algorithm. If the service passes through high-importance nodes and there are few remaining wavelengths in the adjacent links of these important nodes, we need to perform some additional processing to ensure that these resources are reserved for services from/ to important nodes. In the flow chart, the topology of the network is represented as $G(V, E)$, where V and E are the node set and link set respectively. The important node set in the network is $V_{IN} = \{v_{IN1}, v_{IN2}, \dots, v_{INn}\}$, the important link set is $E_{IN} = \{e_{IN1}, e_{IN2}, \dots, e_{IN2}\}$. The unimportant node set is $V_{UN} = \{v_{UN1}, v_{UN2}, \dots, v_{UNn}\}$, the unimportant link set is $E_{UN} = \{e_{UN1}, e_{UN2}, \dots, e_{UNn}\}$. The service request is denoted as $r_i(s_i, d_i)$, where s_i is the source node and d_i is the destination node.

After the service request arrives, firstly use KSP algorithm to calculate K shortest routes $P_r^k = \{P_{r1}, P_{r2}, \dots, P_{rn}\}$, and get the i_{th} ($i < K$) path P_{ri} from the path set. Then judge the type of source and destination nodes. If the source node or destination node is important node, search available wavelengths along P_{ri} and allocate wavelengths with FF algorithm directly. If the source node and destination node are unimportant node, then judge whether the path passes by important node, if not, try to allocate the wavelength. Otherwise, search important links L_{IPI} and calculate the percentage η_L of free wavelengths of each link in the important link set that the path passes through. If η_L is greater than the percentage of wavelengths reserved for important links η_0 , which indicates that there are sufficient wavelength resources on the important link, so this path is available and FF algorithm is used to allocate the wavelength. Until all K paths are executed, if all K paths are allocated unsuccessfully, the request is blocked.

4. Simulation Analysis

To evaluate the performance of the proposed *NIP algorithm*, we use the *14-node NSFNET topology* in our simulation. There are 80 wavelengths in each optical link, and each request demands for a single wavelength. *Node 5* and *Node 9* are set to be important nodes so that the traffic from/to *Node 5* or *Node 9* will be protected. *Figure 2-4* shows the simulation results under 3 different wavelength resources reserved percentages. *Figure 2, 3, 4* are under 4%, 3%, 2% respectively. Blocking probability of important nodes are calculated by the number of requests and blocked requests that source node or/and destination node are important nodes. Also, blocking probability of unimportant nodes as well as the blocking probability of important and unimportant node, that is, entire network, are listed together. Besides, to evaluate the impact of adding important nodes on the entire network, a *Non-Node Importance based Protection (NNIP) algorithm* without considering important node is also simulated for comparison.

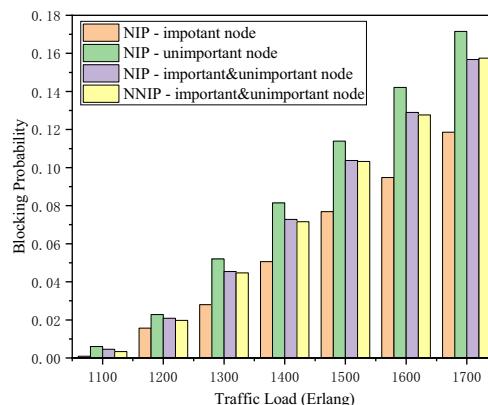


Figure 2. Blocking Probability of important nodes compared with unimportant nodes and the entire network when 4% wavelengths reserved for important links.

In *Figure 2*, we can see that the *NIP algorithm*, which considers reserving wavelengths can largely reduce the blocking probability of important nodes. As traffic load increases, blocking probability reduce more compared with the network blocking

probability. But for exchange, the blocking probability of unimportant nodes increases. It is inevitable that when the network under the high payload, important nodes will no more as the relay nodes. Absence of candidate paths will must lead to this result. As for the entire network, it shows that *NIP algorithm* leads to a slight rise of the network blocking probability than *NNIP* when traffic load is in the range of 1100-1600 Erlang. When traffic load went to 1700 Erlang, *NIP algorithm* achieves lower network blocking probability than *NNIP*. To further investigate the influence of the percentage of reserved resources in *NIP algorithm* on network blocking probability, we conduct the simulation when 3% and 2% wavelengths reserved for important links.

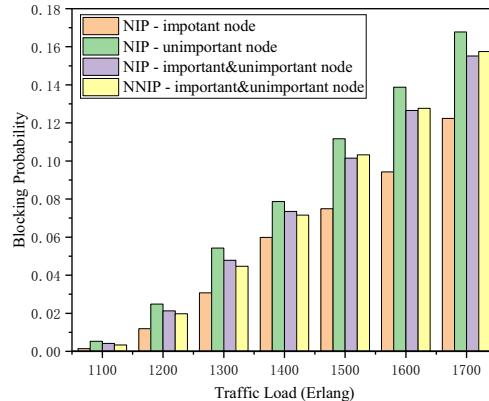


Figure 3. Blocking Probability of important nodes compared with unimportant nodes and the entire network when 3% wavelengths reserved for important links.

In *Figure 3*, it shows a similar trends as *Figure 2* that blocking probability of *NIP* with important node reduce more and with unimportant node increases more compared with total network blocking probability as traffic load increases. However, both of them is less than that in *Figure 2* when 4% wavelengths reserved, which means the protection effect to important nodes is weaken. When it comes to the network blocking probability, *NIP algorithm* achieves lower network blocking probability than *NNIP* under heavy payload situation, i.e., traffic load values from 1500 to 1700 Erlang.

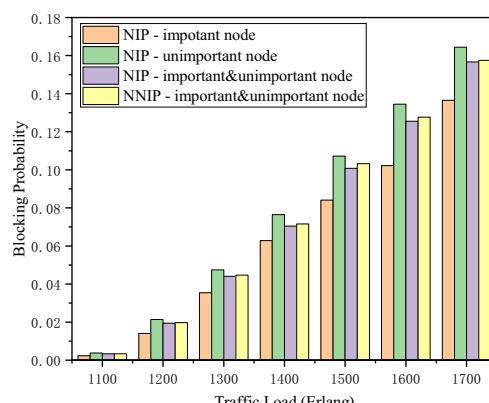


Figure 4. Blocking Probability of important nodes compared with unimportant nodes and the entire network when 2% wavelengths reserved for important links.

We have known that the blocking probability of entire network using *NIP algorithm* could be either higher or lower than using *NNIP algorithm* without considering important nodes. Comparing *Figure 2, 3 and 4*, it manifests that network blocking probability also related to the percentage of wavelengths reserved for important links. For example, in *Figure 2*, the blocking probability of network using *NIP algorithm* is mostly higher than that using *NNIP algorithm*. While in *Figure 4*, the situation is opposite to *Figure 2* that the blocking probability of network using *NIP algorithm* performs better in most cases (blocking probability of network using *NNIP algorithm* are same in *Figure 2, 3 and 4*). Over all, the performance of important nodes in all situations can be largely improved using the proposed algorithm, and only a few wavelengths reserved can improve the performance not only for important nodes, but for the entire network. If better protection to important nodes is needed, a little cost of the entire network should be paid.

5. Conclusions

This paper focus on the problem of protecting several core nodes in power-grid optical backbone communication networks under heavy traffic scenario. The concept of node importance is designed to distinguish important nodes with others. A node importance-based protection algorithm is proposed and evaluated under different resources reserved. It is found that the proposed algorithm could protected important nodes effectively, and can even improve the performance of the entire network in terms of blocking probability under heavy payload situation.

Acknowledgements: This work was supported by the Mass Science and Technology Innovation Project of Ningxia Electric Power Co., LTD. Information Communication Company of State Grid (No. 5229XT21000J).

References

- [1] H. Ding, Y. Wang, A. Liu, D. Yang, D. Chen and R. Xiu. Review on Intelligent Optimization Technology for Survivability of Power Grid Backbone Communication Networks. 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA); 2021 Jun 28-30; Dalian, China. p. 499-502.
- [2] Y. Wu, H. Xu and M. Ni. Defensive Resource Allocation Method for Improving Survivability of Communication and Information System in CPPS Against Cyber-attacks. in Journal of Modern Power Systems and Clean Energy. 2020 Jul; 8(4): 750-759.
- [3] X. Liu, D. Niu, H. Zhang and H. Dong. Analysis of Key Technologies in the Construction of Ubiquitous Power Internet of Things. Proceedings of 2019 2nd International Conference on Intelligent Systems Research and Mechatronics Engineering (ISRME 2019); 2019; Taiyuan, Shanxi. p.557-560.
- [4] J. Xing, J. Chen, X. Sun, X. Zhang and R. Zhang. A K-shell Improved Method for the Importance of Complex Network Nodes. 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS); 2018; Enshi, China. p. 639-643.
- [5] X. Li, Y. Han, X. Wu and D. A. Zhang. Evaluating node importance in complex networks based on TOPSIS and gray correlation. 2018 Chinese Control and Decision Conference (CCDC); 2018; Shenyang, China. p. 750-754.
- [6] X. Chen et al. Flexible Availability-Aware Differentiated Protection in Software-Defined Elastic Optical Networks. in Journal of Lightwave Technology. 2015 Sep; 33(18): 3872-3882.
- [7] P. Papanikolaou, K. Christodoulopoulos, E. Varvarigos. Joint multi-layer survivability techniques for IP-over-elastic-optical- networks. in IEEE/OSA Journal of Optical Communications and Networking. 2017 Jan; 9(1):A85-A98.

A High-Performance Infrared Imaging System with Adaptive Contrast Enhancement

Feng DENG^a, Zhong SU^a, Rui WANG^b, Jun LIU^{c,d,1} and Yanzhi WANG^c

^a*Beijing Key Laboratory of High Dynamic Navigation Technology, Beijing Information Science & Technology University, Beijing, China*

^b*State Key Laboratory of Nickel and Cobalt Resources Comprehensive Utilization, Jinchang, China*

^c*Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA*

^d*Department of Electrical & Computer Engineering, College of Engineering, Northeastern University, Boston, MA, USA*

Abstract. Most of the existing infrared imaging systems employ the scheme of FPGA/FPGA+DSP with numerous peripheral circuits, which leads to complex hardware architecture, limited system versatility, and low computing performance. It has become an intriguing technical problem worldwide to simplify the system structure while improving the imaging performance. In this paper, we present a novel real-time infrared imaging system based on the Rockchip's RV1108 visual processing SoC (system on chip). Moreover, to address the problem of low contrast and dim details in infrared images with a high dynamic range, an adaptive contrast enhancement method based on bilateral filter is proposed and implemented on the system. First, the infrared image is divided into a base layer and a detail layer through bilateral filter, then the base layer is compressed by an adaptive bi-plateau histogram equalization algorithm, and finally a linear-weighted method is used to integrate the detail layer to obtain the image with enhanced details. The experimental results indicate that compared with traditional algorithms, our method can effectively improve the overall contrast of the image, while effectively retaining the image details without noise magnification. For an image of 320*240 pixels, the real-time processing rate of the system is 68 frames/s. The system has the characteristics of simplified structure, perceptive image details, and high computing performance.

Keywords. Infrared imaging system, RV1108, real-time, contrast enhancement, bilateral filter, bi-plateau histogram equalization, high dynamic range

1. Introduction

The infrared (IR) imaging system converts the temperature distribution of the target object into a human visual distinguishable image by non-invasive detection of the object infrared radiation using infrared thermography technology. With innovative

¹ Corresponding Author, Jun Liu, Robotics Institute, School of Computer Science, Carnegie Mellon University, 4227 Newell-Simon Hall, 5000 Forbes Avenue Pittsburgh, PA 15213, USA; E-mail: junliu2@andrew.cmu.edu.

breakthroughs in infrared technology, the application of IR imaging systems has expanded from the military to various civilian industries, which is playing an increasingly important role nowadays. Research on high-performance IR imaging systems is particularly urgent.

Since the first forward-looking infrared imaging system was invented in the United States in the 1960s [1], the infrared thermal imaging system has evolved from the first generation to the third generation [2, 3]. However, the FPGA/FPGA+DSP system structure with plenty of peripheral circuits is widely used for the image processing in most of the existing IR systems, which leads to complex hardware architecture, limited system versatility, and low computing performance. On the other hand, IR images are prone to loss of local detail information when adapt data from the detector to the display. The two main reasons for this problem are: (1) Limited by the constraints of the IR detector manufacturing process, the original IR image has the characteristics of low contrast, blurred details and poor signal-to-noise ratio; (2) The IR detector's response output is 14bits or 16bits for high-performance IR imaging system, which cannot be managed by the general 8bits display devices. Therefore, in the process of the high dynamic range compression, detailed textures with small background temperature differences in the image are vulnerable to lost. It has become an intriguing technical problem worldwide to optimize the system structure while enhancing computation performance and improving details of potential or weak targets in high dynamic range IR images in new generation IR imaging system.

Traditional infrared image enhancement algorithms usually use methods such as automatic gain control (AGC) [4] and histogram equalization (HE) [5, 6], which can better adjust the contrast information of the image, but it is prone to "over bright" or "over dark" in local areas, and lose small details in raw images. The image layer based and transform domain based approaches have better contrast enhancement compared to AGC and HE. F. Branchitta [7] presented a detail enhancement algorithm BF&DRP for high dynamic range compression based on bilateral filtering. C. Zuo et al. [8] proposed an improved method BF&DDE, which using adaptive Gaussian filtering to avoid gradient reversal artifacts caused by the over enhancement of the detailed layer. N. Liu et al. [9] presented a method based on guided filtering to get better contrast enhancement. However, there are many parameters needed to be carefully tuned in the image layer based algorithms for the specific scenario so as to take on the best visualization effect. Transform-domain based approach [10, 11] and AI (artificial intelligence) [12, 13] based method can effectively enhance the small details in the target scene of the IR image, but the algorithms have high time consuming and space complexity, which is not conducive to implementation on embedded systems with limited system resources and high real-time requirements.

In this paper, we present a novel real-time infrared imaging system based on the Rockchip's RV1108 visual processing SoC (system on chip). And moreover, to address the problem of low contrast and dim details in infrared images with a high dynamic range, an adaptive contrast enhancement method based on bilateral filter is proposed and implemented on the system. Our system surpasses traditional IR imaging systems in terms of system structure, imaging quality and real-time performance.

The rest of this paper is organized as follows. Section 2 describes the system architecture in brief. Section 3 elaborates on adaptive contrast enhancement method. Section 4 presents our experimental results and evaluation. Section 5 concludes the paper with remarks on our future work.

2. System architecture design

Infrared imaging system is generally composed of optical system, infrared detector, AD conversion circuit and signal processor. In our proposed system, a continuous zoom infrared germanium lens is used as optical lens, an uncooled focal plane array with far infrared band of 8–14 microns is employed as infrared detector, and Rockchip's RV1108 is adopted as main image processor of the system.

RV1108 is a professional visual processing SoC, which embeds a new-generation DSP for digital processing and an ARM Cortex-A7 single-core processor for system scheduling and application. The ARM microprocessor is designed for system configuration, framework management, DSP driver, and I/O communication interaction. DSP in RV1108 is in charge of executing various image processing algorithms with high efficiency by utilizing specific DSP optimization techniques. Critical system and algorithm codes are stored in DSP internal memories to speed up the program execution with low latency. In addition, RV1108 is capable of supporting H.264 video encoder/decoder up to 1440p@30fps, simultaneously inputting up to 4 cameras, and displaying on one screen [14]. It also supports various camera interfaces such as MIPI (mobile industry processor interface)-CSI (camera serial interface), CVBS (composite video broadcast signal) in, and lots of display interfaces such as MIPI-DSI (display serial interface), HDMI (high definition multimedia interface) 1.4, CVBS out, and serial/parallel RGB. Since RV1108 is a highly integrated and cost-effective SoC, it can fully meet the requirements of a high-performance IR imaging system in terms of efficiency, power consumption, and cost-effectiveness. The block diagram of the system hardware architecture is shown in Fig.1.

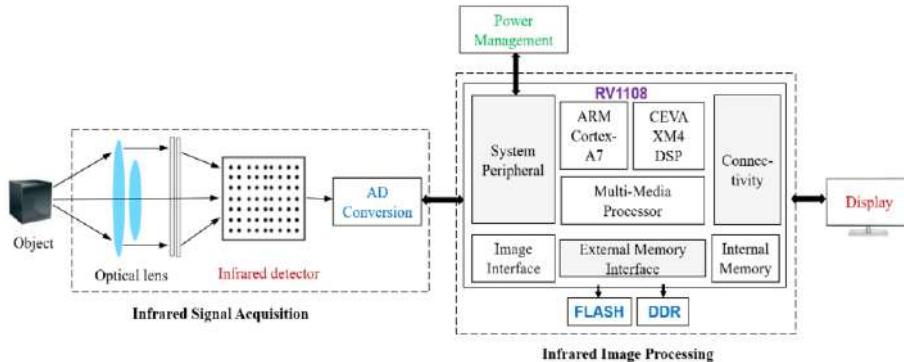


Fig. 1 Block diagram of the system architecture

3. Proposed method

The original input IR images with a wide dynamic range of 14-bit (or more) results in the invisible or faint details visualization during the high dynamic compressing process. Inspired by the visualized effect of BF&DRP method, we propose an adaptive contrast enhancement method, which can adaptively adjust the algorithm parameters so as to achieve the best visualization performance in different scenarios. Fig.2 shows the flowchart of our algorithm. First, bilateral filter is adopted to decompose the image into

the base component and the detail component, which corresponds to the base layer and the detail layer, respectively. Then, an adaptive bi-plateau histogram equalization method is performed on the base layer to dynamically enhance the image contrast according to the image grayscale in different scenarios. At last, we integrate the two parts by linear-weighted method and quantize the result to 8-bit display range.

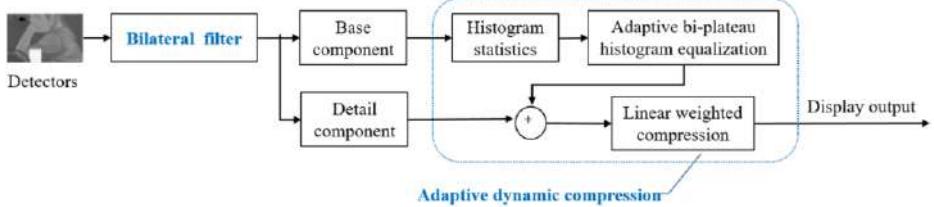


Fig. 2 Flowchart of adaptive contrast enhancement algorithm

3.1 Bilateral filter

Bilateral filter can be acknowledged as a successful filter unlike traditional filters for edge preservation and noise elimination. Using a Gaussian G_σ as a decreasing function, and considering a gray-level image I , the result I^b of the bilateral filter is defined by [15]:

$$I_p^b = \frac{1}{W_p^b} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(I_p - I_q) I_q \quad (1)$$

with $W_p^b = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(I_p - I_q)$

where S denotes the spatial domain, I_p is the gray value of image I at position p , the parameter σ_s defines the size of the spatial neighborhood used to filter a pixel, and σ_r controls how much an adjacent pixel is down weighted because of the intensity difference. W^b normalizes the sum of the weights.

Fig.3 shows the detail layer images obtained using bilateral filter with different filtering windows r . It can be seen, as the filtering window decreases, the detail edges after bilateral filtering become clearer under a certain σ_s and σ_r .

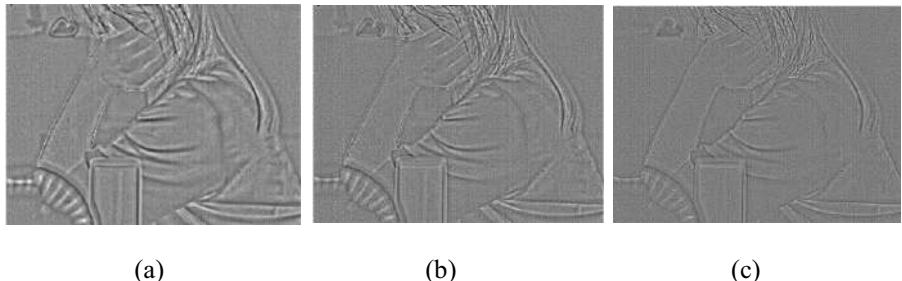


Fig. 3 Detail layer image using bilateral filter with different filtering window. (a) $r = 8$, $\sigma_s = 4$, $\sigma_r = 0.01$. (b) $r = 4$, $\sigma_s = 4$, $\sigma_r = 0.01$. (c) $r = 2$, $\sigma_s = 4$, $\sigma_r = 0.01$.

3.2 Adaptive dynamic compression

After bilateral filtering, we get base component of image as base layer. Then, the subtraction of the input image and the base component is then determined as the detail layer. Since the detail component of the image occupies a smaller dynamic range than the base component, if the dynamic range compression is performed on a uniform scale, the great effect of detail enhancement cannot be achieved. Therefore, we process each component separately. For base layer, the contrast is relatively low, and the image is dark with less details, we propose an adaptive bi-plateau histogram equalization (ABPHE) to enhance the contrast of background. Then integrate the base layer and detail layer by the linear-weighted method to display range. The proposed algorithm can constrain the over-bright background and noises by setting an upper plateau threshold, and enhance the image details and magnify the dim targets by setting a lower threshold value. The upper and lower threshold value can be adjusted adaptively according to different scenarios.

The method of calculating the upper plateau threshold T_{up} is as follows:

$$T_{up} = \text{mean}\{\text{POLAR}(1), \text{POLAR}(k), \dots, \text{POLAR}(P)\} \quad (2)$$

where POLAR is the set of local maximums of the histogram with zero statistics removed, and P is its size. POLAR(k) is the kth element in POLAR.

The calculation of local maximums depends on the value of the window length of the histogram, and enhancement performance will be affected by the calculation result of the upper plateau threshold. According to the experiment, the value of window length 20 is recommended.

The method of calculating the lower plateau threshold is described as pseudocode shown in Algorithm 1.

Algorithm 1 Method of calculating lower plateau threshold

Input: Image x; image histogram hist[];

Output: The lower plateau threshold

- 1: \triangleright Perform truncation operation on hist[] according to the given ratio, which in this algorithm is chosen 0.001
 - 2: $\text{trimage}[] \leftarrow \text{image_tail}(x, 0.001)$
 - 3: \triangleright Calculate the grayscale of the truncated image
 - 4: $L_{gray} = 2^{INT(\ln^{max.gray} / \ln^2)}$
 - 5: \triangleright Get the histogram of the truncated image
 - 6: $\text{trhist}[] \leftarrow \text{histogram}(\text{trimage})$
 - 7: \triangleright get the lower plateau threshold
 - 8: $T_{down} = \lfloor \text{trhist}(L_{gray}) / L_{gray} \rfloor$
-

Fig.4 shows the result comparison of base layer by HE and ABPHE. It is shown that some local areas turn out to be “over bright” in the base layer after processed by HE, such as building outline, pedestrian profile, and tree leaves. While the base layer has uniform brightness and good overall contrast by ABPHE.



(a) The base layer by HE

(b) The base layer by ABPHE

Fig.4 The result comparison of base layer by HE and ABPHE

3.3 Recombine of the base and detail layer

The base layer is projected into the proper range using ABPHE method. In order to display the enhanced image onto the general monitor devices with 8 bits, we first map the processed base layer and the detail layer into the gray scale (typically 0 to 255) by linear strength method. The processed base and detail components are represented by I_{ba} and I_{de} . Finally, the two components are recombined by linear-weighted algorithm to get the final output image I_{out} .

$$I_{out} = \alpha * I_{ba} + (1 - \alpha) * I_{de} \quad (3)$$

where α ($0 < \alpha < 1$) is weighted coefficient. The large the α , the less the detail information in the image. We empirically set $\{\alpha | 0.3 \leq \alpha \leq 0.5\}$.

4. Experimental results and performance evaluation

4.1 System prototype based on RV1108

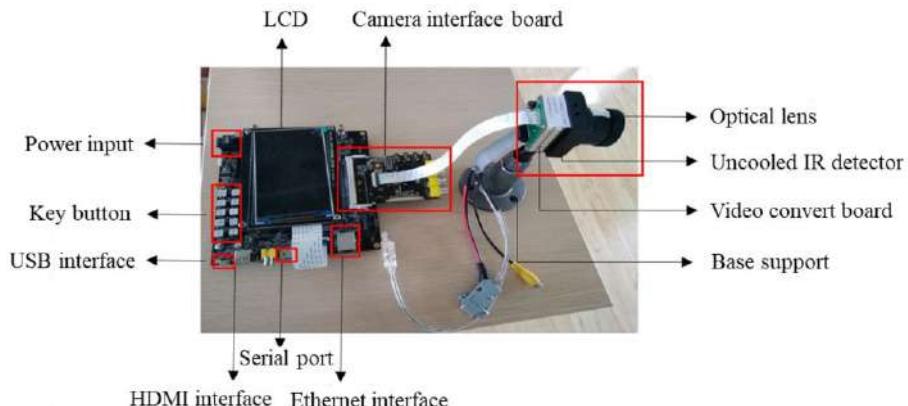


Fig.5 Infrared imaging system prototype based on RV1108

The IR imaging system prototype based on RV1108 developed in this paper is shown in Fig.5. The system uses 8-14 μ m uncooled IR focal plane detector, the resolution of the video image collected by the system is 320×240 pixels, and the accuracy of the raw image is 14-bit.

We ported and implemented the adaptive contrast enhancement method onto the prototype system. To evaluate the system performance, we choose two sets of image data in different scenarios, as shown in Fig.6 (a) and Fig.7 (a). Fig.6 (a) was collected by the proposed IR system. Fig.7 (a) was provided by FIIR with 16-bit data width and resolution 324*256 pixels. The proposed method was compared with the popular image enhancement algorithms, i.e., AGC, HE and BF&DRP.

4.2 Subjective tests

The comparison results of different algorithms and our algorithm is shown in Fig.6 and Fig.7. The experimental results illustrate that the proposed method provides the best visual effect and contrast enhancement over all two sets of images. Besides, it effectively retains the detail information and avoids amplification of the noise.

Fig.6 and Fig.7 provide a high dynamic range scenario with abundant details. It is noticeable that image by AGC has low contrast and less details, image by HE can map the contrast to a certain range, but has the “over bright” phenomenon. For example, the hot cup and girl fingers in Fig.6 (b)), the building in Fig.7 (b) are over-enhanced and halos are produced around the cup, building outline and tree leaves. The method of BF&DRP can improve the overall contrast while maintaining the details of the image, but it tends to magnify the noise in flat regions, for example, the noise in the background are highlighted shown in the Fig.6 (c) and Fig.7 (c). Our proposed method achieves the best performance in terms of contrast enhancement, details perception, and noise suppression. It can be seen evidently from Fig. 6 (d) and Fig.7 (d) that the overall contrast has enhanced, and some details obscure in the original images have visualized, such as the girl’s hair, fingers, clothes folds, edge of hot cup, fan wheel, pedestrian and tree shadow etc. Besides, our method is free of artifacts, producing a natural and homogeneous image without tuning parameters among different scenarios.



(a) AGC



(b) HE

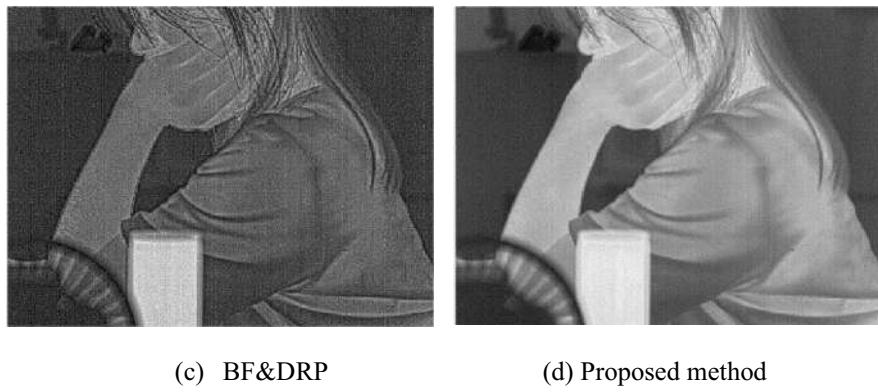


Fig.6 Comparison results of scene 1

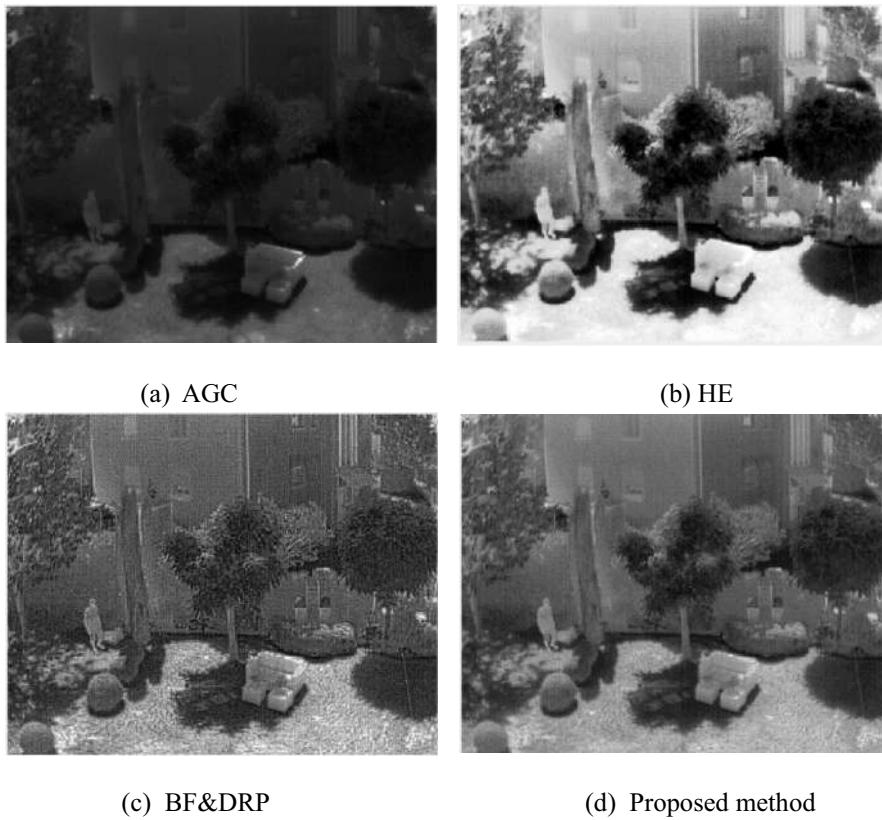


Fig.7 Comparison results of scene 2

4.3 Quantitative tests

In the quantitative tests, standard deviation (STD) and peak signal to noise ratio (PSNR) are used to objectively evaluate the image enhancement effect of different methods. The

test results are given in Table 1. It should point out that HE has higher STD value than BF&DRP and the proposed method, but its PSNR is too low to achieve satisfied visual effects. Compare with BF&DRP, the proposed method has higher STD and PSNR, which is proven to be more effective on image contrast enhancement.

Table 1 The objective evaluation indicators of different algorithms

Scenario	HE	BF&DRP	Proposed
STD	Fig.6	73.6742	38.0261
	Fig.7	73.7171	41.1081
PSNR	Fig.6	14.8819	16.6070
	Fig.7	8.1422	13.3014

4.4 Computation performance

To further demonstrate the system computation performance, we experiment the running time of our methods on the system prototype platform. Table 2 presents the performance evaluation of the core algorithms in terms of RV1108 cycles occupation, average execution time and quantity of code. For the image with size of 320*240 pixels, the average execution time of imaging enhancement algorithm is 14.50 ms/frame under the RV1108 DSP frequency of 600 MHz, namely the image processing speed of the system is 68 frames/s. The evaluation results illustrate that our algorithms have very low time/space complexity, and the real-time performance of the system is excellent and competitive.

Table 2 Computation performance evaluation

Core algorithms	RV1108 cycles (M/frame)	Average execution time (ms/frame)	Code quantity (KB)
Normalization	0.046	0.076	0.542
Bilateral filter	3.266	5.444	4.654
Histogram statistics	0.547	0.912	0.384
Bi-plateau histogram equalization	1.789	2.997	3.202
Adaptive enhancement	7.668	12.779	1.952

5. Conclusion

In this paper, we focused on designing a high-performance IR imaging system with state-of-art contrast enhancement method. Most of the previous works on IR imaging system are based on FPGA/FPGA+DSP architecture, which leads to complex system architecture and limited computing performance. Aiming at these issues, we designed a high-performance uncooled IR imaging system based on the Rockchip RV1108 visual processing SoC to increase the system integration and reliability. Moreover, to address

the problem of low contrast and dim details in infrared images with a high dynamic range, an adaptive contrast enhancement method based on bilateral filter is proposed and implemented on the system. The experimental results indicate that compared with traditional algorithms, our method can effectively improve the overall contrast of the image, while effectively retaining the image details without noise magnification. The system has the characteristics of simplified structure, perceptive image details, and high computing performance. In the future, we will further verify the adaptability of our algorithm in various scenarios, so as to enhance the robustness of the presented IR imaging system.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (41871348); Beijing Information Science & Technology University (2020KYNH224); and Beijing Key Laboratory of High Dynamic Navigation Technology (HND2019002).

REFERENCES

- [1] A. Rogalski. Infrared detectors: an overview [J]. *Infrared Physics & Technology*, 2002, 43 (3–5): 187–210. [http://dx.doi.org/10.1016/s1350-4495\(02\)00140-8](http://dx.doi.org/10.1016/s1350-4495(02)00140-8).
- [2] A. Rogalski. Progress in focal plane array technologies [J]. *Prog. Quantum Electron*, 2012, 36 (2–3): 342–473. <http://dx.doi.org/10.1016/j.pqantelec.2012.07.001>.
- [3] P. Martyniuk, J. Antoszewski, M. Martyniuk, L. Faraone, A. Rogalski. New concepts in infrared photodetector designs [J]. *Appl. Phys. Rev.* 1 (4) (2014) 041102. <http://dx.doi.org/10.1063/1.4896193>.
- [4] R. C. Gonzalez, R. E. Wood. *Digital Image Processing* [M]. 2nd ed. Upper Saddle River: Prentice-Hall, 2002.
- [5] H.O. Chen, N.S.P. Kong, H. Ibrahim. Bi-histogram equalization with a plateau limit for digital image enhancement [J]. *IEEE Transactions on Consumer Electronics*, 2010, 55(4): 2072-2080.
- [6] S. Agrawal, R. Panda, P. K. Mishra. A novel joint histogram equalization based image contrast enhancement [J]. *Journal of King Saud University- Computer and Information Sciences*, 2019: 1-11.
- [7] F. Branchitta, M. Diani, G. Corsini. New technique for the visualization of high dynamic range infrared images source [J]. *Optical Engineering*, 2009, 48(9): 1-9.
- [8] C. Zuo, Q. Chen, J. Ren. Display and detail enhancement for high-dynamic-range infrared images [J]. *Optical Engineering*, 2011, 50(12): 895-900.
- [9] N. Liu, D. Zhao. Detail enhancement for high-dynamic-range infrared images based on guided image filter [J]. *Infrared Physics & Technology*, 2014, 67: 138-147.
- [10] H. I. Ashiba, H. M. Mansour, et al. Enhancement of IR images using histogram processing and the undecimated additive wavelet transform [J]. *Multimedia Tools and Applications*, 2019, 78: 11277–11290.
- [11] Y. Chang, M. Chen, et al. Toward universal stripe removal via wavelet-based deep convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 58(4): 2880-2897
- [12] Z. He, Y. Cao, Y. Dong, et al. Single-image-based nonuniformity correction of uncooled long-wave infrared detectors: a deep-learning approach [J]. *Applied Optics*, 2018, 57(18): 155-164.
- [13] F. Lv, B. Liu, F. Lu. Fast Enhancement for Non-Uniform Illumination Images using Light-weight CNNs. *ACM Multimedia 2020*: 1450-1458
- [14] Fuzhou Rockchip Electronics Co., Ltd. Rockchip RV1108 Datasheet (Revision 1.4). June 2018.
- [15] Tseng Y C, Hsu P H, Chang T S. A 124 Mpixels/s VLSI Design for Histogram-Based Joint Bilateral Filtering [J]. *IEEE Transactions on Image Processing*, 2011, 20(11): 3231-3241.

ReDCN: A Dynamic Bandwidth Enabled Optical Reconfigurable Data Center Network

Xinwei ZHANG¹, Zuoqing ZHAO, Yisong ZHAO, Yuanzhi GUO, Xuwei XUE, Bingli GUO and Shanguo HUANG

*The State Key Laboratory of Information Photonics and Optical Communications,
Beijing University of Posts and Telecommunications, Beijing 100876, China*

Abstract. A reconfigurable optical data center network is proposed, in which the optical bandwidth can be automatically reconfigured by reallocating time slots based on the real time traffic. Numerical investigations validate that the network performance of packet loss after reconfiguration decreases by 58.5%, and the end-to-end latency decreases by 63.8% with comparison to the network with rigid link interconnections, and thereby increasing the 9.4% of throughput at load of 0.8.

Keywords. Reconfigurable bandwidth, optical data center network, time slots

1. Introduction

With the rapid development of traffic boosting applications, the traffic presents a rapid growth trend in data centers (DCs) [1-2]. This poses unprecedented challenges to the existing data center network (DCN) based on electric switching technology from the aspects of both switching technology and network architecture [3-4]. The development and application of optical switching technology in DCN has been extensively investigated to overcome the bandwidth bottleneck of electrical switches [5-6]. However, the high bandwidth between the top-of-racks (TORs) in optical DCNs is uniform after the network building, which cannot reallocate bandwidth according to the variety of traffic in real-time [7-8].

In the latest researches, a feasible solution is to configure specific transmission links or devices for specific applications using load distribution algorithms [7]. However, the deployment of load distribution mechanism will increase the complexity of network control and management, and then increase the cost, especially for large networks. Another method is to build a flexible reconfigurable network with the capability to dynamically adjust optical bandwidth [9-10]. However, the proposed schemes need complicated network interconnections and control mechanism. This is hard to afford the large-scale network with requirements of low management cost and power-consumption.

In this paper, we propose a reconfigurable optical DCN, ReDCN, which can allocate time slots according to the traffic proportion to then reconfigure the bandwidth. In the

¹ Corresponding Author, The State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China; E-mail: zhangxinwei@bupt.edu.cn.

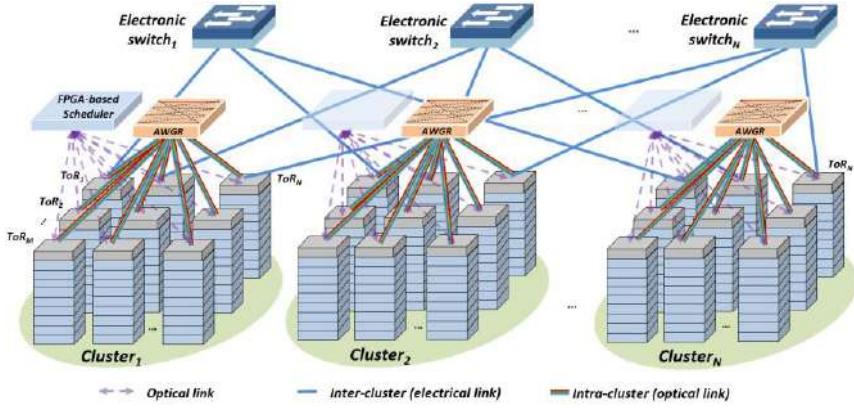


Figure 1. The structure of the proposed reconfigurable optical DCN.

proposed ReDCN network, we implement a field programmable gate array (FPGA)-based scheduler, which is used to deploy the reconfigurable instructions. The scheduler reconfigures the time slots and bandwidth according to the collected ToR traffic and topology information to provide adaptive optical bandwidth for the links in the cluster. The network can provide adaptable optical bandwidth to the hosted application with variety of traffic, and thereby reduce the packet loss rate, improves the latency and throughput performance.

2. Reconfigurable Architecture

Figure 1. shows the proposed network which is divided into n clusters. There are k servers interconnected through ToR in each rack, and each cluster contains n racks. The traffic (Ethernet frame) generated by the server is divided into three types (intra ToR, intra cluster and inter cluster) according to its destination. Ethernet frames are first processed at the Ethernet switch of each ToR. The intra-cluster links are interconnected by the AWGR, while the i -th electronic switch interconnects the i -th ToR of each cluster, where $i = 1, \dots, N$. Therefore, inter cluster communication only needs two hops. This interconnection improves the flexibility of the network.

2.1. Schematic of the ToR

The function block of ToR implemented by the FPGA is shown in **Figure 2.** Each rack contains k servers interconnected through Ethernet switches in ToR. The Ethernet switch processes Ethernet frames with different destinations generated by the server and accordingly forwards them to the AWGR and the electrical switch. In each ToR, n transceivers (TRXs) with corresponding electric buffers are deployed to store the traffic to the AWGR (p) and electric switch ($n-p$) respectively for intra cluster and inter cluster communication. By processing the MAC address of each frame destined for intra cluster and inter cluster links, ToR can calculate the traffic of each link and sends this traffic statistics (traffic distribution within and between clusters) to the FPGA based controller on the optical link.

2.2. Reconfigurable Scheme

In this work, the distributed scheduling system based on FPGA connects ToR through optical links to monitor and collect traffic and topology information. The controller redistributes the time slot according to the collected traffic information, and reconfigures the optical bandwidth of the connection in the cluster to adapt to the current traffic mode.

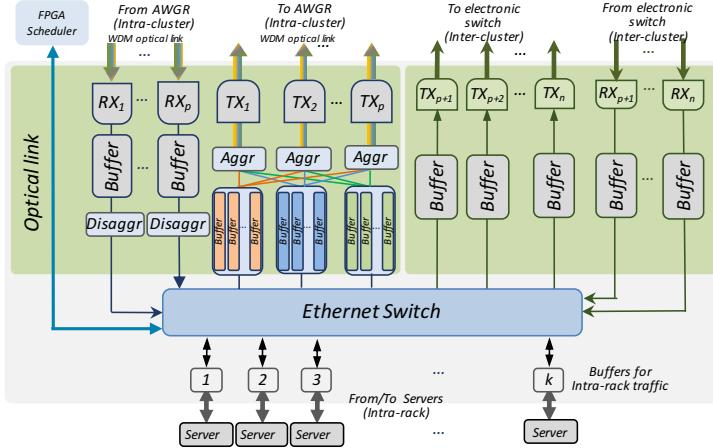


Figure 2. Schematic of the ToR

To highlight the characteristic of the reconfiguration scheme, an example assumes that 4 clusters containing 4 racks in a DCN. Each ToR has 4 TRXs, and they are connected through optical links with AWGR. The i -th buffer caches the traffic sent to the i -th ToR, $i=1, 2, 3, 4$. The traffic ratio distributing to different destinations is different. We design the reconfiguration scheme by reallocating the different scheduling time for each buffer so that they can forward traffic with corresponding bandwidth. As the

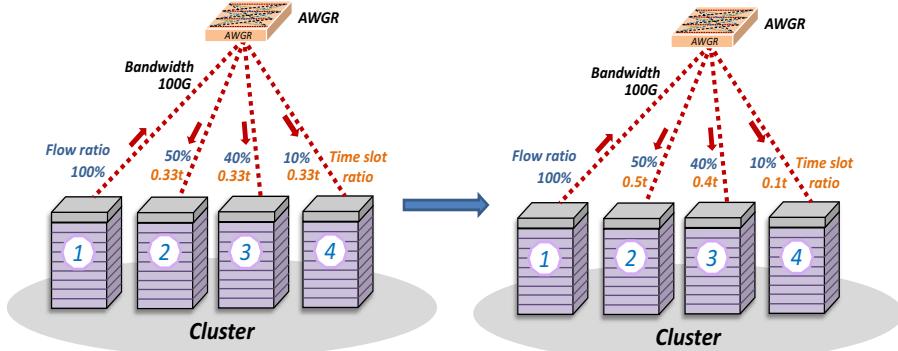


Figure 3. Reconfiguration scheme

reconfiguration scheme is shown in **Figure 3.**, ToR₁ sends 50%, 40%, 10% intra traffic to ToR₂, ToR₃, ToR₄, respectively. Before reconfiguration, the bandwidth is uniformly allocated for the links to the different destination. In our reconfiguration scheme, the controller will adjust the bandwidth allocation by reallocating time slots. So the cycle is accordingly divided into $0.5t$, $0.4t$ and $0.1t$ to schedule corresponding buffers, with t

representing one scheduling cycle. FPGA is a parallel processing logic element, the time slot allocation can be faster, so as to reconfigure the optical bandwidth and realize fast switching scheduling.

3. Simulation and Results

In our experiment, the DCN with 2560 servers is simulated on OMNeT++ platform, which is divided into 8 clusters, 8 racks in each cluster and 40 servers in each rack. ToRs have 2 ports, each port connect with 4 destination ToRs in the same cluster through an AWGR. The traffic ratio in the rack, intra-cluster and inter-cluster are shown in **Table 1.**. The Intra-cluster traffic proportion on each link connected to AWGR is the same and the division of different ToR's intra traffic is shown in **Table 2.**. After reconfiguration, the bandwidth is divided according to the traffic proportion. There are 4 time slots in each cycle, and all ToRs schedule a buffer in each time slot. The link rate between server and ToR is 10Gbps, and the link rate from each buffer to AWGR is 100Gbps. The buffer size of ToR is 80KB, and the buffer queue from the ToR to the server is 20KB.

Table 1. Studied traffic pattern.

Traffic	Intra-ToR		Intra-cluster		Inter-cluster	
	50%	37.5%	37.5%	12.5%	12.5%	12.5%

Table 2. The ratio of intra-cluster traffic

	ToR ₁	ToR ₂	ToR ₃	ToR ₄	ToR ₅	ToR ₆	ToR ₇	ToR ₈
ToR ₁	0	0.2	0.2	0.1	0.05	0.15	0.2	0.1
ToR ₂	0.2	0	0.1	0.2	0.15	0.05	0.1	0.2
ToR ₃	0.2	0.1	0	0.2	0.2	0.1	0.05	0.15
ToR ₄	0.1	0.2	0.2	0	0.1	0.2	0.15	0.05
ToR ₅	0.05	0.15	0.2	0.1	0	0.2	0.2	0.1
ToR ₆	0.15	0.05	0.1	0.2	0.2	0	0.1	0.2
ToR ₇	0.2	0.1	0.05	0.15	0.2	0.1	0	0.2
ToR ₈	0.1	0.2	0.15	0.05	0.1	0.2	0.2	0

Figure 4. shows the simulation results in terms of the packet loss, ToR-to-ToR latency, Server-to-Server latency and throughput before (2560 network) and after reconfiguration (2560 network-R). The packet loss begins to deteriorate when the load is 0.6 when the bandwidth is uniformly allocated, and it begins to deteriorate at the load of 0.8 after reallocating timeslots. Buffers in ToR will be filled faster so that the packet loss in the ToR buffers is more when the load is higher, verifying our reconfiguration scheme is more effective at the high load. Because the bandwidth is reallocated according to the traffic, the buffer can be scheduled in time, so the packet loss rate reduced by at most 58.5%. The ToR-to-ToR latency and Server-to-Server latency begins to deteriorate

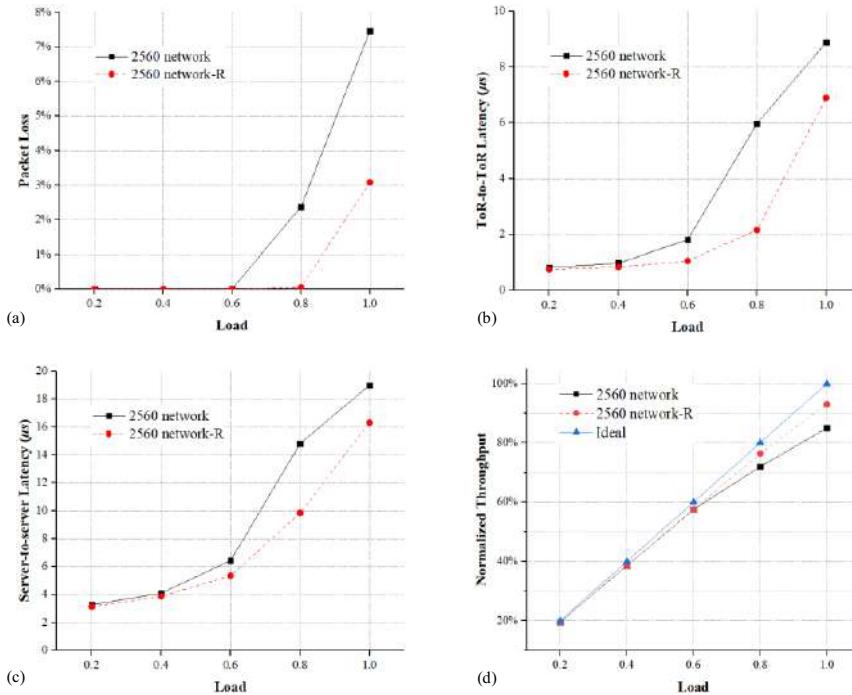


Figure 4. (a)Packet loss (b)ToR-to-ToR latency (c)Server-to-Server latency (d)Throughput

at the load of 0.6, and it gets improvement by 63.8% after reconfiguration since the bandwidth is adapted to real time traffic. When the load is higher, more traffic arrive at server's buffer queues, so the packet loss in the server queues becomes dominant. Because we only calculate the latency of successfully received traffic, so the latency after reconfiguration increases faster at load over 0.8. Due to less packet loss, the throughput is also improved by 9.4% when the link is fully occupied.

4. Conclusions

In this paper, we propose a reconfigurable optical DCN in which the optical bandwidth can be reallocated based on the monitored real time traffic. Numerical results prove the proposed ReDCN improves packet loss by 58.5%, end-to-end latency by 63.8% and throughput by 9.4% with compared to the network with rigid interconnections, validating the good flexibility of the proposed network.

Acknowledgements

The project was supported by Fund of State Key Laboratory of Information Photonics and Optical Communications (Beijing University of Posts and Telecommunications) (No. IPOC2021ZT08), P. R. China.

References

- [1] X. Xue et al., “ROTOs: A Reconfigurable and Cost-Effective Architecture for High-Performance Optical Data Center Networks,” *Journal of Lightwave Technology*, 38(13), pp.3485-3494, 2020.
- [2] M. Analytics, “The age of analytics: competing in a data-driven world,” ed: San Francisco: McKinsey & Company, 2016 (online). See <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>.
- [3] C. Guo, et al. “BeCube: a high performance, server-centric network architecture for modular data centers.” *ACM SIGCOMM*, 2009.
- [4] C. Guo, et al. “DCell: A scalable and fault-tolerant network structure for data centers,” *ACM SIGCOMM*, 38(4):75–86, 2008.
- [5] G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. E. Ng, M. Kozuch, and M. Ryan, “c-S-Through: part-time optics in data centers,” *ACM SIGCOMM Comput. Commun.* 40, 327–338 (2010).
- [6] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, “Helios: a hybrid electrical/optical switch architecture for modular data centers,” *ACM SIGCOMM Comput. Commun.* 41, 339–350 (2010).
- [7] X. Meng, V. Pappas, and L. Zhang, “Improving the scalability of data center networks with traffic-aware virtual machine placement,” *IEEE INFOCOM*, 2010, pp. 1–9.
- [8] X. Xue, F. Yan, B. Pan and N. Calabretta, “Flexibility Assessment of the Reconfigurable OPSquare for Virtualized Data Center Networks Under Realistic Traffics,” *2018 European Conference on Optical Communication (ECOC)*, 2018, pp. 1-3.
- [9] X. Xue, F. Wang, F. Agraz, A. Pagès, B. Pan, F. Yan, X. Guo, S. Spadaro and N. Calabretta. “SDN-controlled and orchestrated OPSquare DCN enabling automatic network slicing with differentiated QoS provisioning,” *Journal of Lightwave Technology*, 38(6), pp.1103-1112, 2020.
- [10] K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen, “OSA: An optical switching architecture for datacenter networks with unprecedented flexibility,” *IEEE/ACM Transactions on Networking*, vol. 22, no. 2, pp. 498-511, 2013.

Routing Optimization Based on OSPF in Multi-Layer Satellite Network

Mingjiang FU¹, Bingli GUO, Hai YANG, Chengguang PANG and Shanguo HUANG

State Key Laboratory of Information Photonics and Optical Communications,

Beijing University of Posts and Telecommunications, Beijing, 100876, China

Abstract. Multi-layer satellite network has become a hot spot for its wider coverage and higher bandwidth level. However, due to the frequent link changes and complexity of network, it is hard to find out a mechanism to handle well on long delay and high packet loss level. This paper proposes an optimized OSPF protocol called OOWLP to eliminate unnecessary routing convergence to optimize the packet loss level and delay ultimately. Link plan table, which records link contacting plan, will be used to update the link state database periodically so that we can eliminate the flooding procedure caused by scheduled link changes. On the other hand, Constrained Shortest Path First (CSPF) will be used to get business differentiated routes in multi-layer satellite network to optimized the throughput capacity in congestion scenario. We divide the sending packets into different businesses and get the routes for each business with longer duration limited by remaining bandwidth. Simulation results show that in normal scenario, average packet loss rate and delay performance are improved 17.42%, 51.44ms respectively, average packet loss rate and throughput capacity performance are optimized 79.05%, 9.81Mbps respectively in congestion scenario compared to standard OSPF. As a result, the proposed mechanism is able to shorten the average delay and lower the packet loss level in multi-layer satellite network.

Keywords. Optimized OSPF with Link Plan (OOWLP), Multi-layer Satellite Network (MLSN), Link Plan Table (LPT), Constrained Shortest Path First (CSPF)

1. Introduction

As an essential member in the 6G version, satellite network will be integrated with the terrestrial networks to provide users ubiquitous access capability [1]. With the development of satellite communication technology, single-layer satellite networks cannot meet the gradually increased needs of coverage area, Quality of Services (QoS) and robustness. Meanwhile multi-layer satellite networks (MLSN) strike into people's sight [2-4], which can provide a more efficient network with better performance than single-layer networks [5]. Link QoS will affect the data interchange in the inter-satellite link deeply. To meet the requirements, the satellite network needs effective QoS guarantee [6].

A distributed contact plan design scheme is proposed to divide the complete contact information into several partial parts to reduce its complexity [7], but frequent link changes still lead to severe packet loss. Korçak [8] proposed a multi-state virtual network

¹ Corresponding Author, Beijing University of Posts and Telecommunications, Beijing, 100876, China;
E-mail: 2019111636@bupt.edu.cn

topology to give us the new efficient handover mechanisms to reduce packet loss, but this scheme cannot handle packet loss well in link congestion scenario. Heuristic algorithms [9-11] lead us to optimize variety of QoS parameters, but it is difficult to apply to MLSN practically for the complexity of the mechanism and long duration of convergence. All in all, these mechanisms have their advantages in normal scenario, but never mentioned in congestion scenario. Enlightened from these mechanisms, we need a mechanism with fine QoS in both normal and congestion scenarios.

In this paper, Optimized OSPF with Link Plan (OOWLp) is proposed to reduce packet loss caused by frequent link change in satellite application scenarios. Link Plan Table (LPT) is introduced to store scheduled link information to update link state periodically. Furthermore, the data business is allocated to different routes according to the remaining link bandwidth and effective link duration to get the optimal transmission path in link congestion scenario, which will be mentioned in Constrained Shortest Path First (CSPF) algorithm herein below. Simulation results show that the enhanced routing strategy can eliminate unnecessary routing convergence to lower the packet loss level and select the optimal transmission path for MLSN routing selection process to shorten the average delay.

The remainder of this paper is organized as follows. Section 2 introduces the construction of multi-layer satellite constellation, including the generation of link plan table. Main strategies involving OOWLp and CSPF are amply presented in Section 3. Simulation results are shown in Section 4. In Section 5, conclusions are provided.

2. Multi-layer Satellite Constellation Construction

Generally, the multi-layer satellite network (MLSN) may consist of Inclined GeoSynchronous Orbit (IGSO) satellites, Medium Earth Orbit (MEO) satellites and Low Earth Orbit (LEO) satellites [12]. The contacting plan design procedure is also an essential approach for MLSN operation. AGI System Tool Kit (STK) [13] is a powerful tool to obtain the visibility data for a given MLSN constellation. With the visibility information, a time-varying topology could be generated through various algorithms, which records link adjacency relationship associated with their start and end time. Finally, link state information is extracted from the time-varying topology structure and composes the LPT, which contains the nodes' link connection state and effective time period to give us the basis of link update later. Our simulations will be set up in Omnet++ simulation environment, which can be freely used for academic research with an open simulation core and supporting alterable network topology [14].

3. Routing Optimization Strategy for MLSN

3.1. OOWLp

In the original rule of OSPF, neighbors are detected by the hello message interaction process, stored in the Interface Neighbor Information Structure (INIS) and gathered to constitute a Link State Advertising (LSA) which is stored in the Link State Database (LSDB). If we apply the standard OSPF directly to our MLSN, frequent link changes will cause the route convergence frequently, leading to long period of network instability, packet loss and longer hops.

To deal with the above-mentioned defects, we propose the OOWLP. As mentioned before, LPT is designed to preserve the link plan information of the whole network for a period of time. The ground station will send the new LPT periodically to the satellites who directly connect to it. Meanwhile the satellites will flood the compressed LPT to other satellites in MLSN.

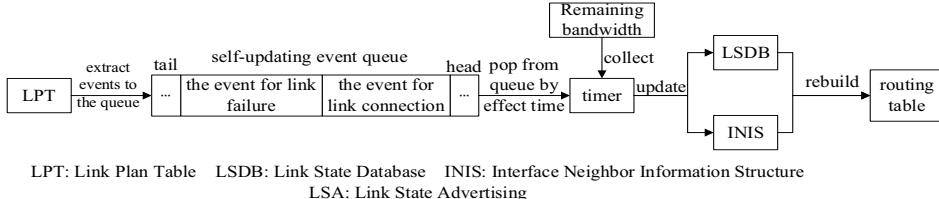


Fig 1. OOWLP updating process.

When any satellite gets the new LPT, the update procedures will come into operation as shown in Fig 1. Firstly, the events are extracted from LPT to update the LSDB and INIS regularly. These events are split into two types based on the LPT entry, including the event for link connection which relied on effective time start, and the event for link disconnected which is based on effective time end. The flag bit called “conn or intrp” is used to distinguish between the two types. Furthermore, a timer is set to handle the events who reach the effective time. For the link connection event, if the event matches an INIS, the neighbor IP and neighbor interface IP in the INIS should be replaced with the corresponding two parts in the event. In addition, the LSA, who matches the event, should update the link id from the neighbor IP of the event. As for the link failure event, it has the same processes to the previous, but we will clear the neighbor information in the corresponding INIS and LSA. Also, we will collect the link remaining bandwidth and add to the corresponding part in LSDB. Finally, we will rebuild the routing table.

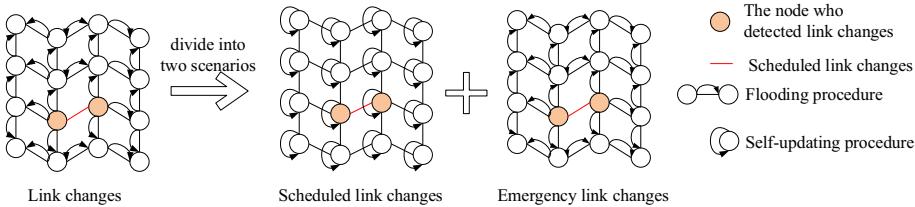


Fig 2. Flood process modification.

Meanwhile, we revise the judgement of flooding procedure as shown in Fig 2. In the original mechanics, if the on-off action appeared in one link, the two nodes of the link will flood the updating message to other nodes in the network. It will cost much of time on route convergence and occupying parts of the link resources. As a result, we divide the link changes to two kinds of situation, for most of time the link change can be predicted, and rare of time the link change is unpredictable. As mentioned before, we periodically update the LSDB and INIS by the LPT. Thus, we restrain the predictable flooding, and keep the link monitoring active for the emergency situation.

3.2. CSPF

Dijkstra algorithm is the core routing algorithm in the OSPF protocol. Based on the link information of the whole network in LSDB, the shortest path from the local node to other nodes in the network can be calculated by Dijkstra algorithm and added into the routing table. However, because of the resource constraints in space, link congestion problem will easily happen if the routes we choose are always in shortest hops. So, we put forward the CSPF on allocating data business to different routes which meet the business needs.

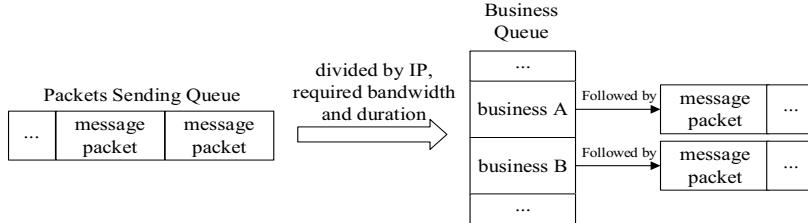
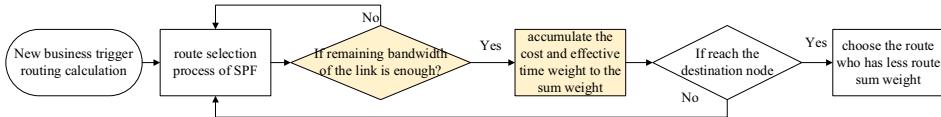


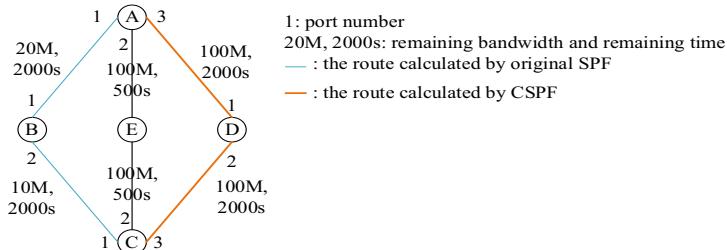
Fig 3. Divide packets from sending queue.

Then, we divide the packets from packet sending queue into different kinds of business, and put each packet into the corresponding business item of business queue. If the packet belongs to new business, record the IP information including source IP and destination IP, required bandwidth and duration of the packet to form a new business item into the business queue, as shown in Fig 3. Afterwards, we send the packet by business type.



(a) CSPF route calculation procedure.

Business: A to C Bandwidth: 80M Transmission time: 1500s



(b) CSPF route calculation compared with original SPF.

Fig 4. The principle of CSPF

Next, we set up the CSPF route calculation procedure based on the SPF algorithm as shown in Fig 4 (a). Route paths are calculated based on the remaining bandwidth and remaining effective time into picking up the links. The link we picked should meet the required bandwidth of the business, and the shortest remaining effective time of the route will be as long as possible. Therefore, we add the weight of effective time into the

accumulation of the link cost. The remaining effective time is divided by an appropriate value as the weight of the effective time to keep balance with the origin cost, here we choose the maximum remaining effective time as the value to divide each remaining effective time of the link. In Fig 4 (b), we provide an example to compare our CSPF with the original SPF. Here we set a data business from node A to node C, requiring bandwidth 80M and transmission time 1500s. Through the original SPF algorithm, we get the route path “A->B->C” with lack bandwidth, which may lead to packet loss. By CSPF, the route path will be replaced to “A->D->C”, which has enough bandwidth and longer duration.

4. Simulation and Results

In this simulation, we build out our MLSN in Omnet++ simulator as Chapter 2 mentioned. And we apply the two sets of routing algorithm to our scenario, one is the standard OSPF and the other is CSPF. Our OOWLP is used to eliminate unnecessary routing convergence and CSPF is to get the business differentiated routes with enough bandwidth. To verify the superiority of the OOWLP, we use the broadcast and multicast business to our scenario and compare the standard OSPF to our OOWLP through average delay and packet loss rate in different contract awarding frequency as shown in Fig 5 (a). As for CSPF, we use the partial congestion business and compare the standard OSPF to our CSPF through throughput capacity and packet loss rate in different sending rate, as shown in Fig 5 (b).

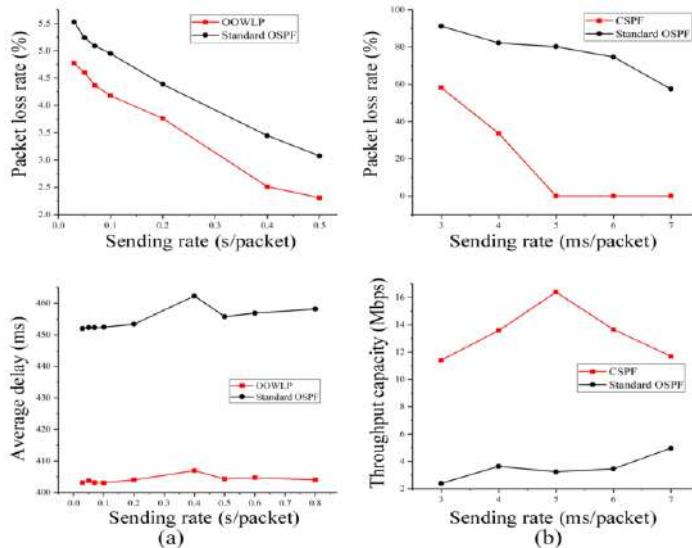


Fig 5 (a) OOWLP and Standard OSPF performance comparison in normal sending rate, (b) CSPF and Standard OSPF performance comparison in the congestion scenario

As shown in Fig 5 (a), the average packet loss rate in OOWLP is lower than the standard OSPF, which decrease 17.42%, and average delay decrease 51.44 ms. Our

OWLP eliminates unnecessary routing convergence to reduce the effect on packet loss and delay. In Fig 5 (b), the average packet loss rate of CSPF is improved with 79.05% and average throughput capacity is increased for 9.81 Mbps compared to the standard OSPF. Because data businesses are divided to the business differentiated routes to reduce the probability of link congestion and increase the throughput.

5. Conclusion

In this paper, we propose OOWLP strategy to eliminate the unnecessary routing convergence and enhance the network performance. Besides the CSPF is raised to get the business differentiated routes to obtain the optimized transmission paths under the link congestion situation in MLSN. Simulation results show that the OOWLP can reduce the packet loss rate and delay effectively. Furthermore the CSPF can decrease the packet loss rate and improve the throughput capacity in link congestion scenario.

Acknowledgements

This work was funded in part by BUPT innovation and entrepreneurship support program (2021-YC-A057), State Key Laboratory of Advanced Optical Communication Systems Networks, China.

Reference

- [1] NetWorld 2020's White Paper, "SatCom WG: The role of satellite in 5G", July 2014, <https://www.networld2020.eu/white-papers/>
- [2] Yu Zhang, Bo Wang, Bingli Guo, "A Research on Integrated Space-Ground Information Network Simulation Platform Based on SDN", published in Computer Networks 188 (2021) 107821.
- [3] Xiangming Zhu, Chunxiao Jiang, LinLing Kuang, "Capacity Analysis of Multi-layer Satellite Networks", published in 2020 International Wireless Communications and Mobile Computing
- [4] Zhizhong Yin, long zhang. "Qos-Aware Multicast Routing Protocol for Triple-layered LEO/HEO/GEO satellite IP networks". IEEE, October 2010.
- [5] ChaoYuan, Xiaoxiang Wang, "A Multicast Routing Algorithm for GEO/LEO Satellite IP Networks", in 2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing.
- [6] Tengyue Mao, Bin zhou , Zhengquan Xu, "A Multi-QoS Optimization Routing for LEO/MEO Satellite IP Networks", published in JOURNAL OF MULTIMEDIA, VOL. 9, NO. 4, APRIL 2014.
- [7] Wenfeng Shi, Deyun Gao, Huachun Zhou, "Distributed Contact Plan Design for Multi-layer Satellite-terrestrial Network", published in China Communications, volume 15, issue 1, Jan.2018.
- [8] Korçak, Ö. and Alagöz, F., "Virtual topology dynamics and handover mechanisms in Earth-fixed LEO satellite systems." Computer Networks, vol. 53, pp. 1497-1511, 2009.
- [9] Fei Long, Naixue Xiong, Athanasios V. Vasilakos, "A sustainable heuristic QoS routing algorithm for pervasive multi-layered satellite wireless networks", Wireless Networks 16, 1657-1673 (2010).
- [10] Y. H. Zhou; F. C. Sun; B. Zhang, "A Hierarchical and Distributed QoS Routing Protocol for Two-layered Satellite Networks", published in: The Proceedings of the Multiconference on "Computational Engineering in Systems Applications".
- [11] Fei Long, Fuchun Sun, Zhian Yang, "A Novel Routing Algorithm Based on MultiObjective Optimization for Satellite Networks", published in JOURNAL OF NETWORKS, VOL. 6, NO. 2, FEBRUARY 2011.
- [12] Wei Zhou, Yi-fan Zhu, Yao-yu Li, Qun Li, Qin-Zhang Yu, "Research on hierarchical architecture and routing of satellite constellation with IGSO-GEO-MEO network", published in "International Journal of Satellite Communications and Networking" volume 38, issue 2.
- [13] Junqing Qi, Zhuoming Li and Gongliang Liu, "Research on Coverage and Link of Multilayer Satellite Network based on STK", 2015 10th International Conference on Communications and Networking in China (China Com).
- [14] OMNeT++ home page: <http://www.omnetpp.org/>.

Data Augmentation Algorithm Based on Generative Antagonism Networks (GAN) Model for Optical Transmission Networks (OTN)

Liang CHEN^a, Kunpeng ZHENG^{b,1}, Yang LI^a, Xuelian YANG^a,
Han ZHANG^a, Yangyang LIANG^a, Junhua HUANG^a, Yuan
ZHANG^a, Pengfei XIE^a, Yongli ZHAO^b

^a*Information and communication branch of State Grid Corporation of China, China*

^b*Beijing University of Posts and Telecommunications, Beijing, China*

Abstract. OTN (Optical Transmission Networks) is one of the mainstream infrastructures over the ground-transmission networks, with the characteristics of large bandwidth, low delay, and high reliability. To ensure a stable working of OTN, it is necessary to perform high-level accurate functions of data traffic analysis, alarm prediction, and fault location. However, there is a serious problem for the implementation of these functions, caused by the shortage of available data but a rather-large amount of dirty data existed in OTN. In the view of current pretreatment, the extracted amount of effective data is very less, not enough to support machine learning. To solve this problem, this paper proposes a data augmentation algorithm based on deep learning. Specifically, Data Augmentation for Optical Transmission Networks under Multi-condition constraint (MVOTNDA) is designed based on GAN Mode with the demonstration of variable-length data augmentation method. Experimental results show that MVOTNDA has better performances than the traditional data augmentation algorithms.

Key words. Optical transmission networks, GAN model, variable-length data, data augmentation, multi Condition constraint

1. Introduction

With the rapid development of optical networks, OTN play a very important role of data transmission in the ground, having the characteristics of large bandwidth, low

¹ Corresponding Author, Kunpeng ZHENG; E-mail: kumpengzheng@bupt.edu.cn.

delay, and high reliability [1]. OTN technologies become more and more mature with expanding of bigger network scale, especially including rapidly attractive difficulties of failure diagnosis. Driven by the era of big data, machine learning nowadays becomes a possible application solution for big data in communication fields, deriving various artificial intelligence algorithms. One typical application of machine learning is exploited for data learning in communication networks, with the deeply-mined rules that are hidden in the data. It has profound-guiding significance for the stable operation and safe production of the existing networks, in terms of data-based traffic analysis, alarm prediction, failure location, and construction of communication network equipment health portraits [2, 3].

However, the amount of effective data is small, and it is not enough to support big data-based algorithms such as machine learning. To solve this problem, it is generally introduced data augmentation algorithms to preprocess the data. While most of the new data augmentation methods are complex, mainly because most of them are based on machine learning models, given by the more common Logistic Regression classifier (LR) [4- 6], K-nearest Neighbors (KNN) and other algorithms [7, 8]. Then to the current hot development of Deep Neural Network (DNN) algorithm [9, 10]. From this, the idea of learning through game antagonism and GAN [11] is introduced to systematically extract features from scarce pre-processed data. On this basis, The AI model for data generation is introduced. Some researchers have introduced Recurrent Neural Networks (RNN) into GAN model, mainly because RNN is suitable for learning features from data sequences. Therefore, Continuous Recurrent Neural Networks with Adversarial Training (C-RNN-GAN) [12], Time-series GAN (Time GAN) [13] based on C-RNN-GAN supplemented with the data static characteristics, Doppel GANger (DG) [14] that associates metadata and feature data have appeared successively. Reference [15] proposed an adaptive traffic data enhancement algorithm is based on the generative confrontation network, enhanced the original traffic data from three different real scenes, and verified the rationality and dynamics of the proposed algorithm through the three statistical characteristics of mean value, variance and Hurst index. Literature [16] also proposed a self-optimized data enhancement algorithm based on original GAN for the enhancement of large optical transmission network alarm data, and verified the effectiveness of the enhanced data through alarm prediction scenarios.

This paper focuses on the data enhancement technology over self-optimization communication networks, and completes the research on the multi-condition variable length data augmentation method. Through the above research content, the “conditional, variable length, and high accuracy” augmentation of the scarce data of the optical transport network has been realized, which provides data support for the performance improvement of artificial intelligence model in communication network scenarios.

2. GAN Module

The GAN model has the core idea of generating antagonism, consisting of Discriminator Module (DM) and Generator Module (GM) two important structures. As shown in Figure 1, it is the basic structure of GAN model. The objectives of DM and

GM are opposite, thus presenting the dual antagonism of training process of GAN. The overall task of GAN for the data generation is to acquire the knowledge of data characteristics through the learning. Then, the synthetic data is used to identify the distribution law of original data, and pass for truth according to the learned knowledge. The objective function of GAN is described in Equation (1).

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)}[\log(D(x))] + E_{x \sim P_g(x)}[\log(1 - D(G(z)))] \quad (1)$$

$P_{data}(x)$ represents the distribution law of the real data set. $G(x)$ represents the process of GM processing input data, such as the target multidimensional data transformed by the random Gaussian white noise vector z into. $D(x)$ represents the probability that data x complies with the distribution law P_{data} of the real data set rather than P_g of the composite data set.

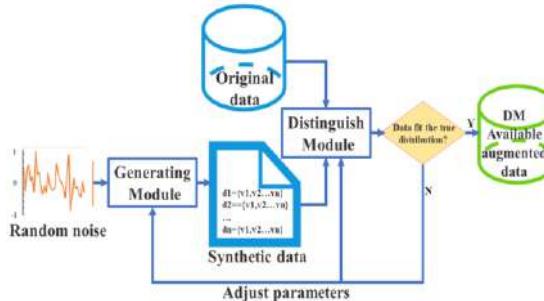


Figure 1. The basic structure of the GANs model.

3. VARIABLE-LENGTH DATA AUGMENTATION ALGORITHM FOR OTN UNDER MULTIPLE CONDITIONS (MVOTNDAA)

This section describes a timing data generator, i.e., the DoppelGANger (DG) [17]. As shown in Figure 2, it is one architecture diagram of DoppelGANger model which is adapted and applied in OTN data augmentation.

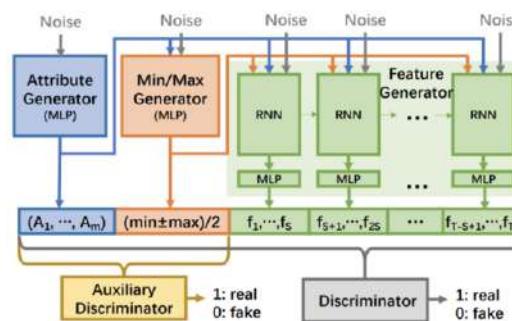


Figure 2. Doppel GANger model architecture diagram.

We conduct the DG model based on GANs and the Generator composed of RNN model, and the improved DG model effectively solves the limitations of timing data

Generator Module (GM). First, DG is an antagonistic generation network based on the GAN model. It uses Recurrent Neural Network (RNN) network model to improve the original GAN to synthesize more complex time series of data sets. For the application of DG, two innovations will be introduced to create the OTN data generator, i.e., a mature learning strategy and the mode collapse of GAN model, in DG to accelerate the convergence rate of GANs. Distinguishment Module (DM) adopts a well-designed noise mode processing strategy to improve the stability of GAN model.

As shown in Figure 3, one method is attribution separation generation method. The other method is a batch generation method, using to synthesize a small batch of long time series data sets. Then, a normalization method can be used to add normalization factor to GM to limit the range of data features. On the basis of DG, Convolutional Neural Networks (CNN) is introduced to analyze and recognize pictorial labels. It mainly aims at the primary optimization of topological data in OTN. MVOTNDA can augment the link data through the specified network links. The DG model is based on GAN's DG model and the generator composed of RNN model. Thus, the limitation of GM module of timing data can be solved effectively.

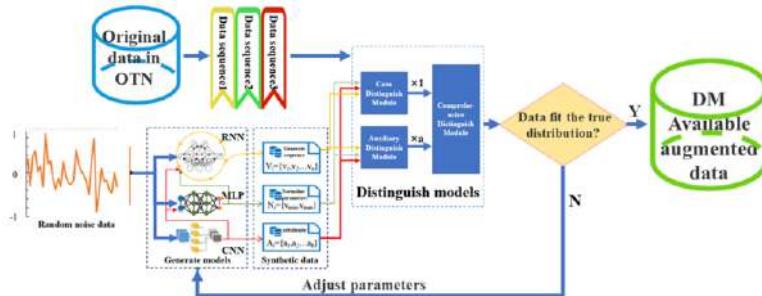


Figure 3. The block diagram of MVOTNDA algorithm based on DG model.

3.1 Objective optimization function of MVOTNDA

In this section, we designed the objective optimization function of MVOTNDA, Generation Module of MVOTNDA and Distinguishment Module of MVOTNDA. Reference [17] introduces Wasserstein distance to solve the model instability problem in the original GAN training process. When Wasserstein GAN (WGAN) model is trained, the adjustment of GM and DM training round ratio is omitted.

The introduction of Wasserstein distance is mainly reflected in the following two modifications of the objective function,

- (1) The logarithm function in the objective function is omitted;
 - (2) Each iteration update process of GM and DM, the model's relevant parameters were taken and deleted to keep them within a certain range.
- To make the objective function conform to the problem, we convert the equation into the following form:

$$WL(G, D) = \min_G \max_D E_{s \sim P_{DS^k}} [D(s)] + E_{z \sim P_z(z)} [-D(G(z))] \quad (2)$$

$G(z)$ is the new sample data generated by GM module. $D(s)$ is the value output by DM, represented the probability value that a generated sample passed by the GM to the DM will be identified as true.

P_{DS}^k represents the mathematical distribution of the original sample data set, P_Z represents the mathematical distribution of pseudo-data generated by the GM, $\min G$ and $\max D$ represent a process of simultaneous maximum and minimum optimization.

The detailed procedures of GAN model based on Wasserstein distance objective optimization function are shown in Table 1.

Table 1. Training process of GAN model based on Wasserstein distance objective optimization function.

Training process of GAN model based on Wasserstein distance objective optimization function	
Training parameter setting: Learning Rate, α ; Gradient Clipping, C ; Number of iterations of training, $\bar{\Pi}$; Number of iterations of the distinguish module, M .	
Initialize i and j to 0	
While $i < n$ do	
While $j < m$ do	
Take any k $\{z^1, \dots, z^k\}$ input vectors from the white noise distribution $P_z(z)$;	
K data columns are randomly selected from the mathematical distribution P of the original data set; The DM was updated by Gradient Descent method:	
If the data is measured data:	
$d_w := \nabla_{\theta_d} \frac{1}{k} \sum_l^k [-D(s^l) + D(G(z^l)) - \lambda(\ D_2(M(\hat{s}))\ _2 - 1)^2]$	
Else The data is label data	
$d_w := \nabla_{\theta_d} \frac{1}{k} \sum_l^k [-M(D_2(s^l)) + D_2(M(G_2(z^l))) - \lambda(\ D_2(M(\hat{s}))\ _2 - 1)^2]$	
$w := w + \alpha \cdot d_w$	
$w := \text{clip}(w, -c, c)$	
$j++;$	
End While	
Take any k $\{z^1, \dots, z^k\}$ input vectors from the white noise distribution $P_z(z)$;	
The GM is updated by Gradient Descent method:	
$g_w := \nabla_{\theta_g} \frac{1}{k} \sum_l^k [-D(G(z^l))]$	
$w := w + \alpha \cdot g_w$	
$i++;$	
End While	

As shown in Figure 2, it shows the importance to adjust the objective optimization function based on the Wasserstein distance. We introduce the Auxiliary Distinguishment Module (ADM) in the DM. The objective optimization function of MVOTNDA algorithm is designed as follows:

$$\min_G \max_{D_1, D_2} V(D, G) = WL(G, D_1) + \gamma WL(G, D_2) - \lambda(\|D_2(M(\hat{s}))\|_2 - 1)^2 \quad (3)$$

Due to the unknown influence of ADM, we introduce a coefficient item γ to adjust the parameters during the experiment.

3.2 Generating Module of MVOTNDA

As shown in Figure 2, the GM of MVOTNDA algorithm is mainly used to synthesize the false data that follows the mathematical distribution of the original sample set. It is composed of three neural networks, namely CNN, MLP and RNN. Among them, RNN is responsible for the generation of pseudo-data, MLP is responsible for the generation of normalized parameters, and CNN is responsible for the generation of label data.

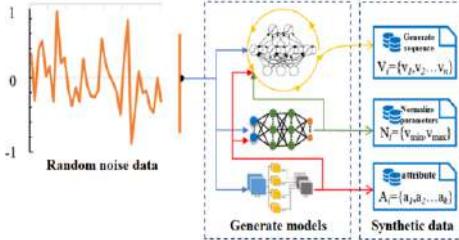


Figure 4. The GM structure.

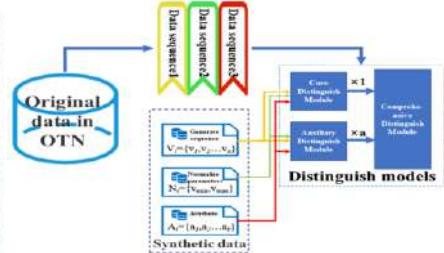


Figure 5. The DM structure.

CNN can identify the characteristics of each link in the topology and generate the corresponding data of each link, such as the flow data in the link. The output of CNN network acts on MLP normalized parameter generator to guide MLP to synthesize normalized parameters of data types identified by CNN. Finally, the output of CNN and MLP act on RNN together to guide RNN to generate sequence data under the corresponding label.

The GM structure of the MVOTNDA model is shown in Figure 4. By accepting the input of random white noise, the RNN model processes it to generate a sequence of side length:

$$\tilde{s}' = G(z) \quad (4)$$

Among them, z is a random white noise vector; \tilde{s}' represents a synthetic pseudo sequence whose length is variable. The objective optimization function of GM can be:

$$G_{loss} = E_{z \sim P_z(z)}[-D(G(z))] \quad (5)$$

$$G_{loss,2} = E_{z \sim P_z(z)}[-D_2(G_2(z))] \quad (6)$$

The loss value G_{loss} of the generated module needs to be minimized to maximize $D(G(z))$, so that the identification module cannot identify the authenticity of the data synthesized by the GM. the GM adopts autoregressive strategy. All output vectors adjusted by the fully connected neural network are merged into a time series in order to generate new and complete time series data.

3.3 Distinguishment Module of MVOTNDA

The DM is used to judge the authenticity of the input sequence data. Take a sequence from the original data sample set or the sequence synthesized by the generation module as input, and let the identification module determine the probability value of its output value between [0, 1]. The objective optimization function of the MVOTNDA algorithm's distinguishment module is shown at below:

$$D_{loss} = E_{s \sim P_{Disk}}[-D(s)] + E_{z \sim P_z(z)}[D(G(z))] \quad (7)$$

In the above equation, it is necessary to minimize the loss value D_{loss} of the identification model. This process mainly maximizes $D(s)$ and minimizes $D(G(z))$. In the other words, in the process of minimizing the objective optimization function, if the sequence in the original sample set is taken as the input, the value of $D(s)$ output by the identification module gradually approaches 1; if the input is the data synthesized by the GM, the output value $D(\tilde{s})$ tends to 0. After the training, the identification module can accurately determine the input sequence data. The specific structure of the identification module is shown in the Figure 5.

4. EXPERIMENTAL CONDITIONS AND RESULTS

The prediction is based on the augmentation and prediction of the traffic data of the British academic backbone network[18]. The data set used in the augmentation and prediction experiment of the traffic data of the British academic backbone network is from The UK Academic Network Backbone. The data traffic is mapped to the preset links in the topology. The experimental simulation topology is shown in Figure 6.

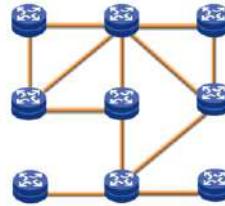


Figure 6. The simulation topology.

MVOTNDA is used to present the performances both of MD-SET and A1H. In addition, LSTM and the Original GAN model are used as the baseline. The Data conducted by the three models are shown in Table 2.

Table. 2 The comparison of three methods' performances.

Performance Index	MVOTNDA	LSTM	Original GANs
Similarity(SS)	0.976	0.826	0.865
Mutual Information(MI)	0.956	0.762	0.799
Accuracy rate	0.980	0.859	0.875

The accuracy of the data set is shown in Figure 7. It can be seen from the figure, most of dirty/redundant data is deleted after the preprocess. Only the performance sequence associated with alarm is left to generate a Data Set called the OD-set (Original Data Set). MVOTNDA augmented data set is MAD-set, with the LSTM-augmented data set is LSTM-set and the data generation of the original GANs model. Figure 8 shows the traffic prediction accuracy curve based on the three generated data. Followed by LSTM, the original GANs model is the worst one, while the accuracy of traffic prediction model trained by enhanced data of MVOTNDA is close to 97.5%.

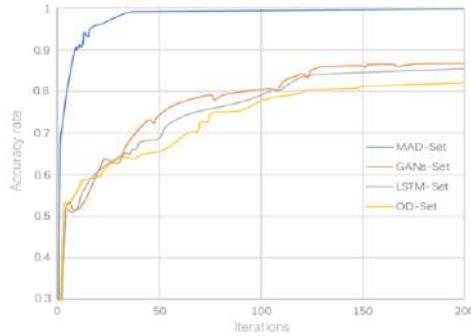


Figure 7. The accuracy of multi-alarm prediction model.

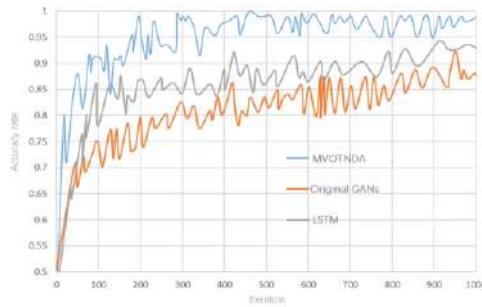


Figure 8. The accuracy of traffic prediction model.

5. Conclusion

In this paper, MVOTNDA is proposed under the background of OTN. Experiments show that augment traffic data in MVOTNDA is based on the specified link and the conditional generation capability. By introducing the label generation module and the auxiliary identification module, MVOTNDA can generate label data under the semi-supervised mechanism and act as the measurement data generation module MLP, which is also supervised by the core identification module, making the synthetic measurement data under the label conform to the mathematical distribution law of the original data set. MVOTNDA can quickly improve the accuracy to more than 90%, and the algorithm has good performance.

Acknowledgments

This work has been supported in part by National Natural Science Foundation of China (NSFC) (62021005, 61822105).

Reference

- [1] Takuya Ohara, Mitsuhiro Teshima, Shigeki Aisawa and Masahiko Jinno, “OTN Technology for Multi-flow Optical Transponder in Elastic 400G/1T Transmission Era”[J], OFC 2012, March 4, 2012- March 8, 2012
- [2] Zhao Y., Yan B., Liu D., et al. “SOON: self-optimizing optical networks with machine learning”[J]. Optics Express, 2018, 26(22):28713.
- [3] Zhang B., Zhao Y., Rahman S., et al. Alarm classification prediction based on cross-layer artificial intelligence interaction in self-optimized optical networks (SOON)[J]. Optical Fiber Technology, 2020, 57:102251.
- [4] Menard, Scott. “Applied logistic regression analysis.[M]. Sage Publications, Incorporated, 2001, Vol.106.
- [5] Breiman, Leo. Random forests[J]. Machine learning, 2001, 45(1): 5-32.
- [6] Jennie, Pearce, and, et al. Evaluating the predictive performance of habitat models developed using logistic regression[J]. Ecological Modelling, 2000.
- [7] S.J.Back, K.M.Sung. “Fast K-nearest-neighbor search algorithm for nonparametric classification”[J]. Ctronics Letters,2000,36(21):1821-1822.
- [8] Yufei Tao, Dimitris Papadias, Nikos Mamoulis, Jun Zhang. “An efficient cost model for K-NN search technical report[J]. HKUST, 2001, 13(1): 1-14.
- [9] Bishop C M, Bishop C, Bishop C M, et al. Neural Network for Pattern Recognition. 1995.
- [10] Schmidhuber, Jürgen. Deep Learning in Neural Networks: An Overview[J]. Neural Netw, 2015, 61:85-117.
- [11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, "Generative adversarial nets", Advances in Neural Information Processing Systems, 2014, pp.2672-2680.
- [12] Mogen O. C-RNN-GAN: Continuous recurrent neural networks with adversarial training. 2016.
- [13] Yin L, Zhang B. Time series generative adversarial network controller for long-term smart generation control of microgrids[J]. Applied Energy, 281.
- [14] Lin Z, Jain A, Wang C, et al. Using GANs for Sharing Networked Time Series Data: Challenges, Initial Promise, and Open Questions[C]// IMC '20: ACM Internet Measurement Conference. ACM, 2020.
- [15] Shuai L, Jin L, Min Z, et al. Adaptive Traffic Data Augmentation Using Generative Adversarial Networks for Optical Networks[C]// Optical Fiber Communication Conference. 2019.
- [16] Zhuang H, Zhao Y, X Yu, et al. Machine-Learning-based Alarm Prediction with GANs-based Self-Optimizing Data Augmentation in Large-Scale Optical Transport Networks[C], 2020 International Conference on Computing, Networking and Communications (ICNC). 2020.
- [17] Arjovsky M, Chintala S, L, et al. Wasserstein generative adversarial networks. 2017.
- [18] Cortez, P., Rio, M., Rocha, M., and Sousa, P. (2012). Multi-scale Internet traffic forecasting using neural networks and time series methods. Expert Systems, 29(2), 143-155.

OSNR Prediction Based on Federal Learning in Multi-Domain Optical Networks

Junhua HUANG^a, Bohan ZHU^{b,1}, Hongxi ZHOU^a, Qiwei ZHENG^a, Zhuo CHEN^a,

Tong LIN^a, Yuanyuan ZHAO^a, Ziqi DONG^a, Xiaofang GAO^a, Yongli ZHAO^b

^a*Information and Communication Branch of State Grid Corporation of China, China*

^b*Beijing University of Posts and Telecommunications, China*

Abstract. With the continuous expansion of the scale of optical communication network and the rapid increase of network traffic demand, the management form of multi-domain optical network has widely existed. OSNR is an important indicator to judge the quality of communication. It is very important to predict OSNR more accurately in a low-cost and energy-saving way in multi-domain optical networks. In this paper, a scheme of federal learning in multi-domain optical networks is proposed to improve the accuracy of the OSNR prediction. The main idea is to train hybrid machine learning model in each single domain, then the strategy of federal learning is used for optimization it in multi-domains. The performance of the proposed scheme is verified by simulation experiments. The strategy can alleviate the problems of data silos and model training set caused by multi-domain optical network. According to simulation result, when the amount of data reaches 5×10^3 , adding this strategy will reduce the mean square error of the prediction model by about 18%. It can improve the performance of machine learning model, the ability of OSNR prediction and the reliability of network operation.

Keywords. OSNR prediction, federal learning in multi-domain, optical networks

1. Introduction

Optical communication networks have the advantages of large bandwidth and low latency. It is an important foundation of the next generation of communication technology [1]. With the increasing scale of optical networks, a large-scale optical communication network is divided into multiple small-scale optical networks according to various dimensions such as geographical locations, managers and equipment types. Each divided small-scale optical network is called an optical network domain, and each optical network domain is managed by its own domain managers.

OSNR has long been widely used in the study of optical communication networks. The value of OSNR directly reflects the transmission quality of signals [2]. OSNR can be considered as a condition in routing and network optimization algorithms. With the increasing diversity of communication services, data silos and model training sets are new challenges to OSNR prediction. Besides, due to the bottleneck of the controller and

¹ Corresponding Author: 1285603132@qq.com

data acquisition mechanism, multi-domain network management constraints and other reasons, the amount of data monitored in their respective management domains is small, the scale of the single domain network will not be too large. Traditional OSNR prediction model cannot exchange data well between domains, and it has high requirements for training samples in a single domain. To sum up, it's more and more difficult for traditional OSNR prediction model to play a role in the multi-domain optical network system.

There are two major directions for OSNR prediction of optical communication equipment. One is based on the prior knowledge scheme of analytical formula. In reference [3], the prediction of OSNR based on power and noise is proposed. The value of OSNR can be calculated by analytical formula through the measured power and noise, which is limited by whether the power and noise can be accurately calculated. In reference [4], it is proposed to monitor OSNR indexes through high-order statistical moments. However, as the evolution process of optical network technology, traditional OSNR monitoring acquisition method does not capture the many important OSNR characteristics. The other direction is based on the posterior knowledge scheme of historical data and machine learning algorithm. Artificial intelligence algorithm is a posteriori knowledge scheme. With the help of historical data, it learns data features through algorithm model and obtains calculation results. In reference [5], a deep learning system based on artificial neural network dynamically predicts the OSNR value of optical communication equipment by monitoring modulation format, channel occupancy rate, power, link length, EDFA noise figure and other parameters. In reference [6], an OSNR calculation method combining cyclic neural network and attention mechanism is proposed. However, it is difficult for these AI models to be directly applied to actual scenarios, because they require many data sets for model training. However, in actual scenarios, there are few effective device performance data, which may fail to meet the requirements of prediction accuracy.

In this paper, an OSNR prediction algorithm based on hybrid machine learning model is proposed. The main idea is to train hybrid machine learning model in each single domain, which combines a priori knowledge with a posteriori knowledge. Then the strategy of federal learning is used for optimization it in multi-domains [7]. Superiority of the OSNR prediction algorithm based on hybrid machine learning model is verified by simulation. Compared with the analytical formula method, the accuracy is higher, and the training complexity is lower than the pure original data-driven machine learning algorithm. Finally, based on the hybrid machine learning model, the strategy of federal learning in multi-domain is introduced, and the OSNR prediction algorithm flow based on federal learning is proposed. The effectiveness of the algorithm is verified by simulation experiments. Compared with acquiring knowledge from a single network domain, the technical scheme of acquiring knowledge from multi-domain and multi-source has better performance. Mean square error (MSE) of OSNR prediction model is 18% smaller when the modal optimized with strategy of federal learning in multi-domain.

2. Hybrid Machine Learning Model

Hybrid machine learning model combines a priori knowledge with a posteriori knowledge. By introducing prior knowledge into the machine learning model based on posterior knowledge, the training complexity of the algorithm can be reduced. It should be noted that in this paper, it is only to prove that the hybrid machine learning model

composed of introducing prior knowledge into the posterior knowledge model has better performance. Different prior knowledge models and a posteriori knowledge models can be selected for different network scenarios to improve performance. In this paper, the OSNR calculation method in reference [3] is used as a prior knowledge, and the calculation is shown in equation (1) and equation (2). The deep neural network model is selected as a posteriori knowledge to learn the -logical relationship of historical data.

$$OSNR = P_{out} - P_{ASE} \quad (1)$$

$$P_{ASE} = -58 + N_F + G \quad (2)$$

$OSNR$ is the value calculated by using the algorithm based on prior knowledge; P_{out} is the output power of the optical communication equipment; P_{ASE} is the noise power of the output end of the optical communication equipment; the amplifier noise power N_F is calculated from the amplifier noise figure and the amplifier gain G .

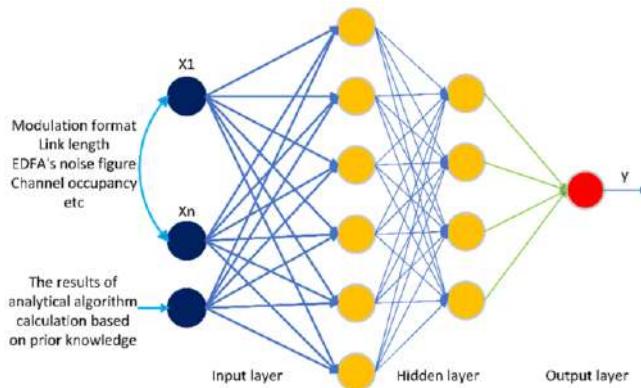


Figure 1. Network structure diagram based on hybrid machine learning model

For the traditional machine learning model, the modulation format, link length, and the noise figure, resource usage and power of the erbium-doped optical fiber amplifier (EDFA) are used as inputs of deep neural networks [8,9], which can be used to extract complex features through multiple nonlinear hidden layers. Then, the nonlinear relationship between input and output is obtained, and the calculated OSNR value is output at the output layer. For the hybrid machine learning model, an additional input parameter is added compared with the traditional machine learning model, which is the result of calculation based on prior knowledge. The outputs are the same. Figure 1 shows the network structure of hybrid machine learning model.

3. Knowledge Fusion Strategy Scheme

In this section, a knowledge fusion strategy is proposed for federal learning of multi-domain in optical networks (as shown in Figure 2). In essence, the scheme uses the local network domain and other domain models to make federal decisions, so that the domain controller can obtain knowledge from a single source to multiple sources. This scheme

eases the problem of data silos between multiple domains in optical network, makes the data exchange between domains smoother.

The machine learning model is represented by a series of parameters (the structure of deep neural network, the number of layers, the number of hidden layer nodes and the weight coefficient after network training, etc.). Each trained model consists of specific vector parameters θ Composed of. Each machine learning model can receive a series of input characteristic data X and output the value of prediction Y , which can be numerical. This machine learning model and additional metadata π (π stores some parameters required for knowledge fusion) are shared among domains as a combination.

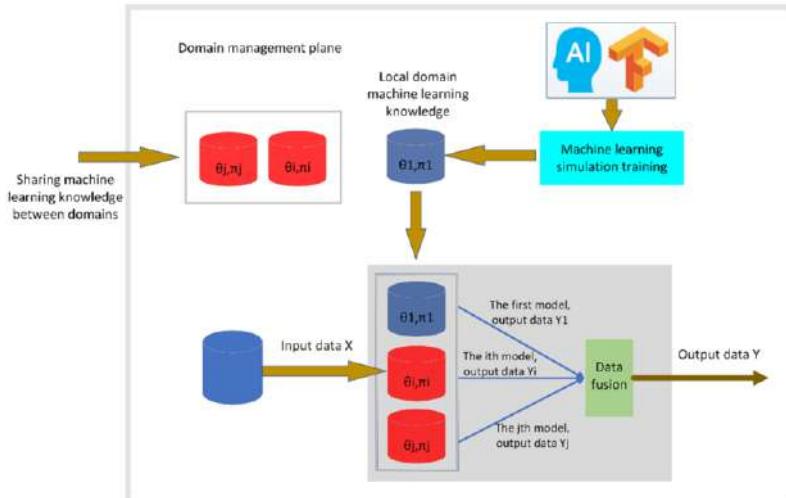


Figure 2. Knowledge fusion strategy for multi-domain optical networks

The specific steps of knowledge fusion are as follows: firstly, the machine learning model trained through the local domain network data in the local domain can be transferred to the local machine learning model. If the model can meet the performance requirements, the network decision can be made, and the knowledge can be shared among alternative domains to other domains. Once the data sample size is insufficient or accuracy of OSNR prediction is too low, you can request high-quality models in other domains for collaborative decision-making; Then, after requesting high-quality models in other domains in the local domain, input the characteristic data into all models to obtain their own model output Y_i . Finally, through the data fusion operation (for the calculation of regression operation equation (3), the result with the highest output frequency can be taken as the final output data for the classification task), the output results Y_i of all models are fused to obtain the final output result Y .

$$Y = \frac{Y_1 f_1 + Y_2 f_2 + \dots + Y_n f_n}{n} \quad (3)$$

Y is the final output data, Y_1, Y_2 and Y_n are the output data of each machine learning model, and n is the number of machine learning models.

4. OSNR Prediction Strategy of Federal Learning in Multi-Domain

This section mainly discusses the process of OSNR prediction algorithm based on federal learning strategy in multi-domain optical networks. It is assumed that the OSNR prediction model in a single domain has been trained, and there will be a trained hybrid machine learning model in each single domain network for OSNR prediction. As shown in the Figure 3, The OSNR prediction technology scheme based on federal learning in multi-domain is mainly divided into three steps:

Step 1 is for the controller to collect historical network data. The collected data types include modulation formats, link lengths, noise coefficients of EDFA, resource occupancy rates, power, and the network characteristic data required by the prior knowledge calculation formula. The results are then calculated by prior knowledge, and the calculated results are combined with other network data into the input feature vector required by the hybrid machine learning model (as shown in Figure 1 input neuron).

Step 2 is the data transmission process, and the input eigenvectors combined in step 1 are input into the trained hybrid machine learning model.

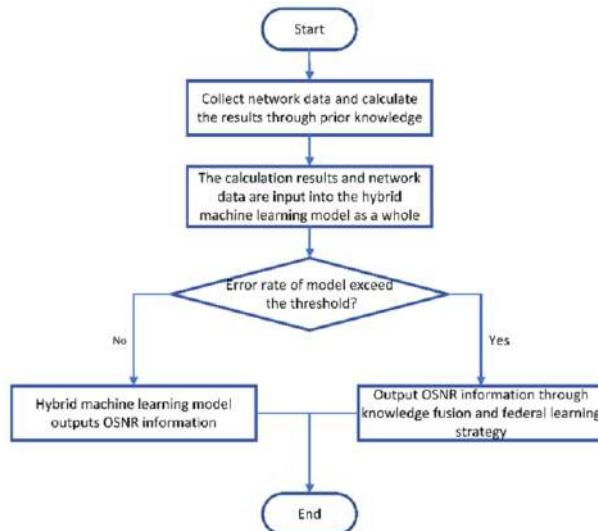


Figure 3. Flow chart of OSNR prediction based on multi-domain in federal learning.

Step 3 is the availability judgment step of hybrid machine learning. The network controller maintains a recent prediction error rate for the original local network domain model. When the prediction error rate reaches the set threshold, the strategy of federal learning in multi-domain will be started. Through knowledge fusion, the strategy of federal learning in multi-domain will replace the output of the original model as optical signal-to-noise ratio information. If the error rate does not reach the set threshold, the multi strategy of federal learning in multi-domain switch will be turned off. The OSNR is predicted directly through the existing single network domain hybrid machine learning model.

So far, through the above three steps, by collecting the network characteristic data required by the hybrid machine learning model and finally outputting the OSNR information, other network applications can obtain the OSNR information and take corresponding network decisions, such as considering the OSNR constraints in routing calculation. Accurate prediction of OSNR can provide a solid foundation for network traffic engineering and enhance the healthy operation of the network.

5. Simulation Analysis of OSNR Prediction Based on Federal Learning in Multi-Domain

To verify the reliability and effectiveness of OSNR prediction algorithm based on hybrid machine learning model, this paper simulates optical communication equipment through Python language and Gnpv optical network planning tool library [10,11]. The system diagram is shown in Figure 4. Different data are generated by changing communication parameters (such as adjusting format, power, amplifier parameters, link length, etc.). The data is expanded by adding white noise to the data, and then the generated data set is divided into two parts, namely, training set and verification set. The hybrid machine learning model is generated through the training set, and the model is verified on the generated verification set.

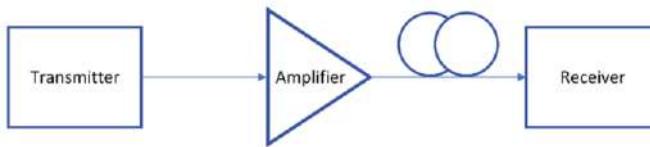


Figure 4. Schematic diagram of the system for generating experimental data.

This paper uses the hybrid machine learning model described in section 2 to predict the OSNR of the generated data. The prior knowledge results are calculated by equations (1) and (2). The posterior knowledge selects the deep neural network model. After a lot of attempts, a group of stable network structure is finally selected.

Firstly, the accuracy of the hybrid machine learning model with full rich data samples and training time at convergence is verified. The MSE predicted by OSNR are 0.295db (analytical formula model), 0.144db (machine learning model) and 0.136db (hybrid machine learning model) respectively. Compared with machine learning model, the error of hybrid machine learning model is reduced by 5%. It is worth noting that in the case of limited data and training time, the hybrid machine learning model has better performance than the machine learning model. In Figure 5(a), we can see the convergence process of the adopted model, in which the hybrid machine learning model has faster speed. More specifically, when the goal of MSE is 0.2db, the time required for hybrid machine learning is 34% less than that of machine learning model (using the same neural network configuration and the same training parameters). It can be found that the training cycle required by hybrid machine learning model is slightly shorter than that of machine learning model.

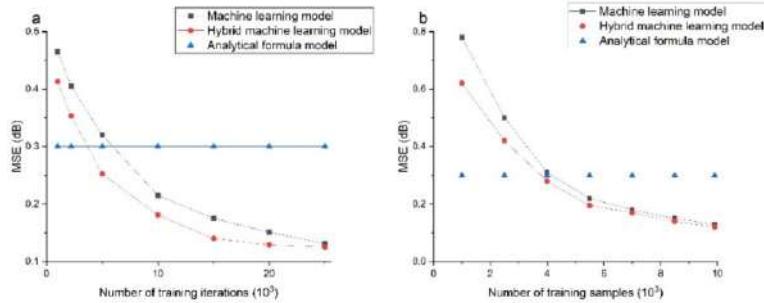


Figure 5. (a) Comparison of training time of different models. (b) Comparison s of training samples in different model.

Figure 5(b) shows the MSE of the three models under different sample numbers. In the case of fewer training samples, the hybrid machine learning model can achieve similar performance to the machine learning model. To finally achieve the result of 0.144 dB of the machine learning model, the hybrid machine learning model can reduce the training samples by 25%. Moreover, to exceed the prediction accuracy of the analytical formula model, the hybrid machine learning model has about 12% less training samples than the traditional machine learning model. The experimental results show that the amount of training data required by the hybrid machine learning model is greater than or equal to the machine learning model to achieve the performance like the machine learning model.

To verify the optimization of the strategy of federal learning in multi-domain on the OSNR prediction scene, two data of the same scale are generated according to the data generation mode through the Gnpv library, which is used to simulate other network domain data in the multi-domain network and use the three data to train three hybrid machine learning models. The control experiment is formed by comparing with the model trained with only full single domain data. Figure 6 shows the MSE of the hybrid machine learning model before and after the introduction of strategy of federal learning in multi-domain. It can be seen from the figure that for the OSNR prediction model with federal learning is introduced, the MSE is 18% smaller when the final model converges, which means that the model optimized with strategy of federal learning in multi-domain

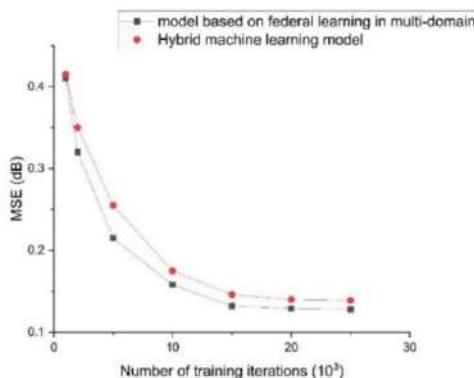


Figure 6. Renderings of fusion strategy of federal learning in multi-domain.

has better performance and higher prediction accuracy. To sum up, the introduction of this strategy can optimize the model performance of the local network domain and obtain a model with better performance.

6. Conclusions

This paper analyzes the problem of OSNR prediction and its related technologies in detail and discusses the scheme evolution of OSNR prediction in single domain network in detail. Starting from the prior knowledge scheme based on analytical formula and the posteriori knowledge scheme based on machine learning. In this paper, a hybrid machine learning model based on prior knowledge and posteriori knowledge is proposed to predict the OSNR, and then the application of strategy of federal learning in multi-domain to OSNR prediction is discussed. Then the algorithm flow of OSNR based on multi-domain optical network federal learning is described. Finally, the performance of the proposed strategy scheme is verified by simulation experiments. This strategy scheme can alleviate the problems of data silos and model training set caused by multi-domain optical network. The experimental results show that in a certain network domain. The algorithm with strategy of federal learning in multi-domain has better performance than the model algorithm obtained only with local data sets.

Acknowledgments

This work has been supported in part by National Natural Science Foundation of China (NSFC) (62021005, 61822105).

References

- [1] F5G Whitepaper rev5 [EB/OL]. 2020-07-30. <https://portal.etsi.org/ngppapp/ContributionCreation.aspx?primarykeys=204101>
- [2] Z. Wang, A. Yang, P. Guo, L. Feng and P. He, "CNN based OSNR estimation method for long haul optical fiber communication systems," 2018 Asia Communications and Photonics Conference (ACP), 2018, pp. 1-3, doi: 10.1109/ACP.2018.8596136.
- [3] Samuel A L. Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress[J]. 1988.
- [4] Hui Yang, Jie Zhang, Yongli Zhao, Jianrui Han, Yi Lin, Young Lee. Cross stratum optimization for software defined multi-domain and multi-layer optical transport networks deploying with data centers[J]. Optical Switching and Networking, 2015.
- [5] Morii M , Ikeda S , Tominaga N , et al. Machine-learning Selection of Optical Transients in Subaru/Hyper Suprime-Cam Survey[J]. Publications of the Astronomical Society of Japan, 2016, 68(6):104.
- [6] Musumeci F, Rottondi C, Nag A, et al. An Overview on Application of Machine Learning Techniques in Optical Networks[J]. IEEE Communications Surveys & Tutorials, 2018:1-1.
- [7] Lili Zhang, Yuxiang Xie, Luan Xida and Xin Zhang, "Multi-source heterogeneous data fusion," 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, 2018, pp. 47-51.
- [8] M. J. da Silva et al., "Investigation of Spectral Tilt and OSNR in Cascade of EDFA with Linear Signal Power Pre-emphasis," 2019 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC), 2019, pp. 1-3, doi: 10.1109/IMOC43827.2019.9317623.
- [9] Y. Yamamoto, M. Hirano, S. Oda, Y. Aoki, K. Sone and J. C. Rasmussen, "Impact of fiber loss and Aeff on OSNR improvement for hybrid-Raman/EDFA amplified systems," 2015 Opto-Electronics and Communications Conference (OECC), 2015, pp. 1-3, doi: 10.1109/OECC.2015.7340235.

- [10] A. Ferrari et al., "The GNPy Open Source Library of Applications for Software Abstraction of WDM Data Transport in Open Optical Networks," 2020 6th IEEE Conference on Network Softwarization (NetSoft), 2020, pp. 386-394, doi: 10.1109/NetSoft48620.2020.9165313.
- [11] A. Ferrari et al., "GNPy: an open source planning tool for open optical networks," 2020 International Conference on Optical Network Design and Modeling (ONDM), 2020, pp. 1-6, doi: 10.23919/ONDM48393.2020.9133027.

Research on 3D Fingerprint Positioning Based on MIMO Base Station

Chang XIA, Yijie REN, Xiaojun WANG¹, Weiguang SUN, Fei TANG and Xiaoshu CHEN

School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University, China

Abstract. The aim of this article is to solve the problem that the accuracy of traditional positioning algorithm decreases in complex environment and to provide some ideas for the few researches of fingerprint localization algorithm in three-dimensional space. This paper builds a system model in a three-dimensional space, provides three reference point distribution methods, and discusses the positioning performance under these distribution methods. After that, based on the high base station deployment density, multi-point fusion positioning method is used to locate the target, which further improves the positioning accuracy and makes more effective use of reference point resources. Finally, a backward-assisted positioning method is proposed, which uses the position information of the positioned points to assist the positioning of the current point. Research shows that this method can improve the positioning accuracy and has good versatility.²

Keywords. Three-dimensional, multi-point fusion positioning, backward-assisted positioning

1. Introduction

In recent years, location-based services have attracted widespread attention in various industries. Different from the traditional positioning algorithm, the fingerprint positioning algorithm uses some corresponding characteristics generated during the signal transmission to achieve positioning [1-3]. In the process of wireless signal transmission, due to the influence of multipath effects, the channel characteristics at different physical locations will reflect some unique characteristics, which called the fingerprints of the signal. When there is a positioning requirement, the fingerprint information of the point to be located is compared with the fingerprint information in the database, and the information of the point is estimated according to a certain similarity criterion. The flow chart of the fingerprint location algorithm is shown in Figure.1 [4-5].

The current fingerprint positioning algorithms mainly focus on the two-dimensional plane, and there are few researches on the positioning in the three-dimensional space. But in modern indoor scenes, especially in environments such as shopping malls, office buildings or hospitals, the positioning of the three-dimensional space becomes more

¹ Corresponding author: Xiaojun WANG, School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University, Nanjing, China; E-mail: wxj@seu.edu.cn.

² Foundation items: Social Development Projects of Jiangsu Science and Technology Department (No.BE2018704).

important. Secondly, with the development of 5G, the combination of fingerprint location algorithms with new generation communication technologies and scenarios is still in progress. Therefore, the research in this article will focus on the above issues. The current fingerprint positioning algorithms mainly focus on the two-dimensional plane, and there are few researches on the positioning in the three-dimensional space. But in modern indoor scenes, especially in environments such as shopping malls, office buildings or hospitals, the positioning of the three-dimensional space becomes more important. Secondly, with the development of 5G, the combination of fingerprint location algorithms with new generation communication technologies and scenarios is still in progress. Therefore, the research in this article will focus on the above issues.

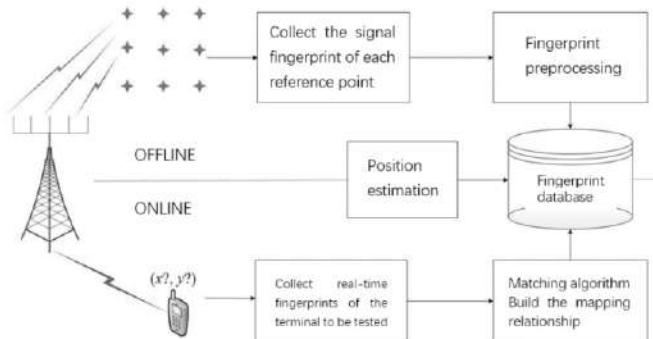


Figure 1. The flow chart of the fingerprint location algorithm.

2. Three-dimensional positioning based on massive MIMO

2.1. Three-dimensional space model

For most open areas outdoors, there is no special requirement for the height of the positioning point, thus in the positioning process, only two coordinate axes are often needed to determine the location. However, people currently spend most of their time in buildings. Although for some specific scenes, they can still be positioned only in the horizontal plane, for most scenes, an additional height coordinate is needed for accurate positioning. Therefore, in the three-dimensional space, the reference point distribution model and the scatterer model require an additional height coordinate axis compared to the two-dimensional space. The three-dimensional model of the scatterer is shown in Figure 2.

For the reference point distribution model in fingerprint positioning, in two-dimensional positioning, the reference points are uniformly distributed. However, considering that the scene in the three-dimensional space is more complicated, and because there is a height coordinate axis, when setting the reference point, it can be considered to introduce a small height float on the Z axis to better show the difference in fingerprint characteristics.

There are three types of reference point distribution models in this article:

- The first is a layered uniform distribution, and the interval of each reference point is 2m, as shown in Figure 3(a).
- The second type is a layered floating distribution, that is, on the basis of the first type of distribution, a random number is introduced in the position coordinates

of each reference point to make the reference point float to different degrees. A random position float of 0.5m is introduced here, as shown in Figure.3(b).

- The third type is full random distribution. The position distribution of reference points no longer has a clear level, but is randomly distributed in the entire three-dimensional space, and there is no fixed grid spacing of reference points. As shown in Figure.3(c).

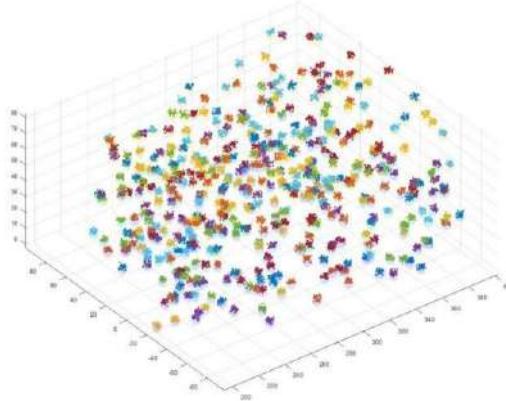


Figure 2. Three-dimensional scatterer model.

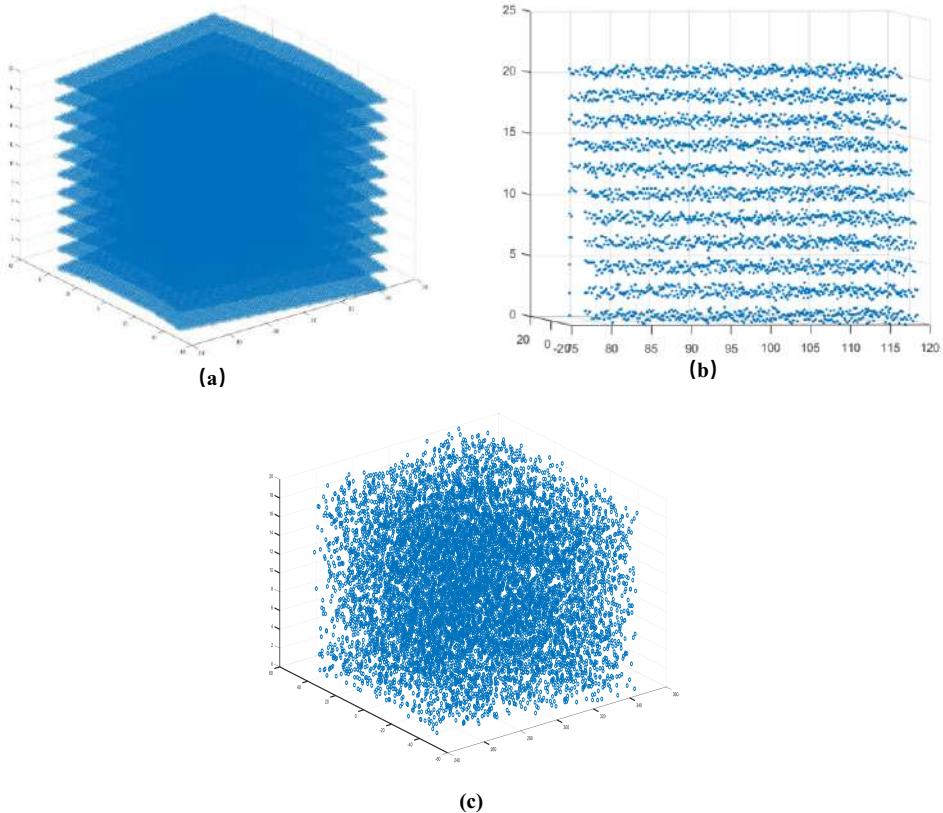


Figure 3. (a) Layered uniform distribution. (b) Layered floating distribution. (c) Full random distribution.

2.2. Antenna array

In the two-dimensional positioning, the antenna array adopts a uniform linear array, but in the three-dimensional positioning, an additional elevation angle needs to be introduced to complete the positioning. In this paper, a uniform planar array is introduced to complete the three-dimensional space positioning. The model of the uniform planar array is shown in Figure 4.

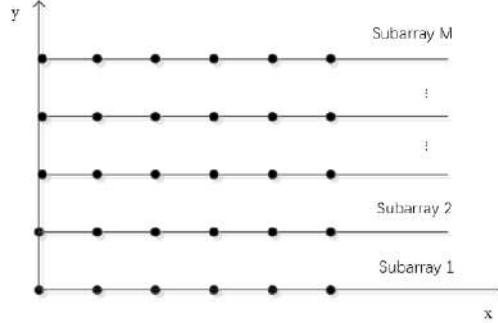


Figure 4. Uniform planar array model.

The XOY plane is an $M \times N$ uniform rectangular area array with M antenna elements and N antenna elements on the X axis and Y axis respectively, and the distance between each adjacent antenna element is d . For a uniform planar array, it can simultaneously estimate two-dimensional angle information, namely the azimuth angle and the pitch angle.

Assuming that K signals are incident on the antenna array, regardless of the noise of the system, the signal received by the i -th element is:

$$x_i(t) = \sum_{i=1}^K s_i(t - \tau_{li}) = \sum_{i=1}^K s_i(t) e^{-j\omega_0 \tau_{li}} \quad (1)$$

Write the above formula as a matrix form :

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j\omega_0 \tau_{21}} & e^{-j\omega_0 \tau_{22}} & \dots & e^{-j\omega_0 \tau_{2K}} \\ \vdots & \vdots & \dots & \vdots \\ e^{-j\omega_0 \tau_{M1}} & e^{-j\omega_0 \tau_{M2}} & \dots & e^{-j\omega_0 \tau_{MK}} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_K(t) \end{bmatrix} \quad (2)$$

Among them, the delay τ_{li} is expressed as:

$$\tau_{li} = \frac{(x_l \cos \theta_i \sin \varphi_i + y_l \sin \theta_i \sin \varphi_i)}{c} \quad (3)$$

The i -th phase shift signal can be expressed as:

$$\phi_{(m,n,i)} = \frac{2\pi}{\lambda} (md \cos \theta_i \sin \varphi_i + nd \sin \theta_i \sin \varphi_i) \quad (4)$$

The response vector can be written as:

$$\mathbf{e}(\theta_i, \varphi_i) = \begin{bmatrix} 1 & e^{-j\phi_{(0,1,i)}} & \dots & e^{-j\phi_{(M-1,N-1,i)}} \end{bmatrix}^T \quad (5)$$

2.3. Fingerprint similarity criteria

For the calculation of the distance between fingerprints, due to the needs of the positioning itself and the consideration of the practicality of the algorithm, the overall distance function between the fingerprints should show a monotonous increasing or monotonous decreasing trend. In this section, the Joint Angle Delay Similarity Coefficient (JADSC) is introduced to measure the similarity between fingerprints [6].

For the joint angle delay similarity coefficient at position i and position j, it is defined as:

$$J_{ij} = \max_{n \in (-L+1, L-1)} \left(\frac{1}{N_t^2} [C_i]_t^T [C_j]_{t-n} \right) \quad (6)$$

Notation: Matrix C_i represents the fingerprint matrix at position I, $[C_i]_t$ represents the t-th column of the fingerprint matrix, n represents the delay compensation between two fingerprints, and L represents the maximum delay compensation.

In three-dimensional positioning, there may be a value of zero when calculating the similarity of two fingerprints. Here, such a point is called a blind spot. Generally, if the similarity is zero, it can be considered that the similarity between the two fingerprints is very low, and the fingerprint can be discarded during calculation. However, for fingerprints with strong sparsity, some blind spots may be adjacent to the points to be located, which greatly affects the effectiveness of the positioning algorithm.

Considering the expansion of the positioning dimension, in view of the strong sparsity of ADCPM, this article improves JADSC. When calculating the similarity, a very small positive value is introduced into the fingerprint, so as to ensure that the value of JADSC will not be zero. The improved JADSC is as follows:

$$J_{ij} = \max_{n \in (-L+1, L-1)} \left(\frac{1}{N_t^2} [C_i]_t^T [C_j]_{t-n} \right) + \Delta J \quad (7)$$

ΔJ is a minimal positive value, here is 10^{-10} .

2.4. Simulation and analysis

The simulation is based on MATLAB to build a massive MIMO-OFDM single station system. The radius of the area is $R_{\text{loc}}=100\text{m}$, the width is $W_{\text{loc}}=40\text{m}$, the height is $H_{\text{loc}}=20\text{m}$, and the compression ratio ρ is 1000. The reference point distribution model is the above three distributions. Other relevant parameters are shown in Table 1 [7].

Table 1. System Simulation Parameters

OFDM parameters	Transmission bandwidth B_W	5-80MHz
	Subcarrier spacing Δf	15kHz
	Pilot interval T_g	4.7 μ s
	Sampling interval T_s	1/(2 · BW)
Antenna parameters	Number of antennas N_t	128
	The location of the base station BS_{pos}	[0;0]
Other parameters	Scatterer radius R_{Scat}	20m
	Scatterer density	0.02/ m^2
	Subpath N_{Sub}	40
	Test points N	1000

The three-dimensional space positioning is performed under the three reference point distribution modes, and the number of reference points in the three methods remains the same. The positioning results are shown in Figure.5(a).

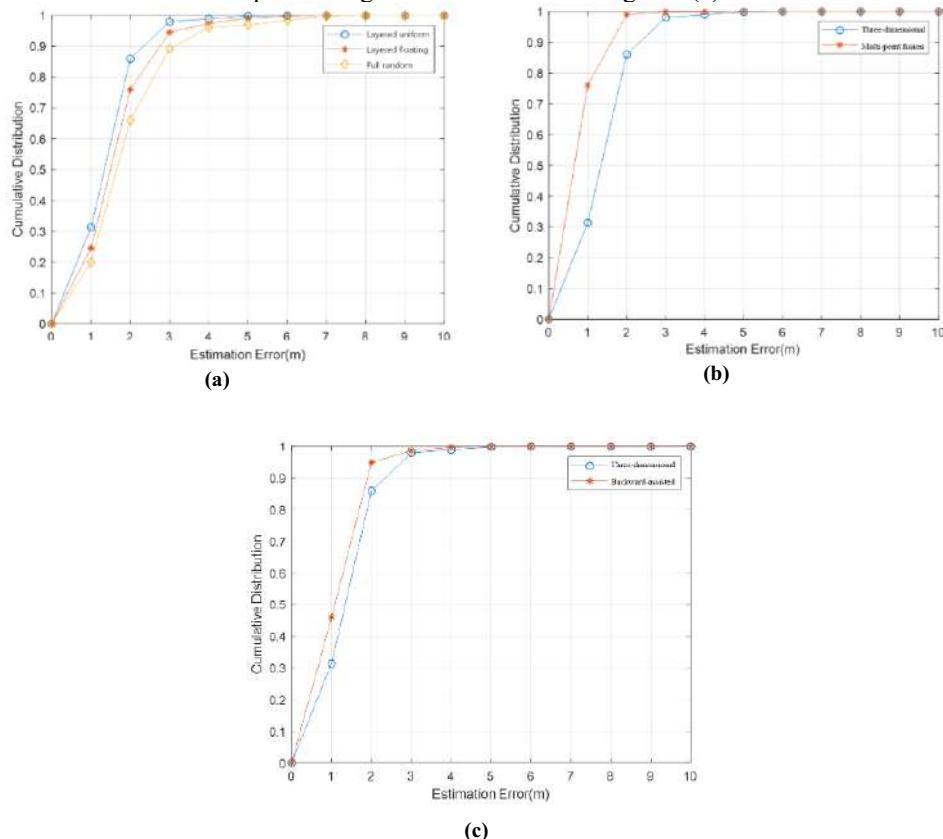


Figure 5. (a) Positioning result diagram under three reference point distributions. (b) Multi-point fusion positioning result. (c) Backward-assisted positioning result.

Simulation analysis: From the figure, in the three reference point distribution modes, the layered uniform distribution has the highest positioning accuracy, and the fully random distribution has the lowest accuracy. In the fully random distribution, although the characteristics between fingerprints at each reference point will be more obvious, this method will make some information missing, and this imbalance will have a negative impact on the matching algorithm. However, each reference point is uniformly distributed in a layered uniform distribution, and the transition of different fingerprint features is gentle, which is conducive to the matching algorithm.

3. Multi-point fusion positioning

3.1. Theoretical basis

In fingerprint positioning, the fingerprint library collects the fingerprint information of the reference points set within the communication range of the base station. Under high base station deployment density, the communication areas of each base station will overlap to different degrees, and the corresponding reference point network may also produce various interlaces. The schematic diagram is shown in Figure.6.

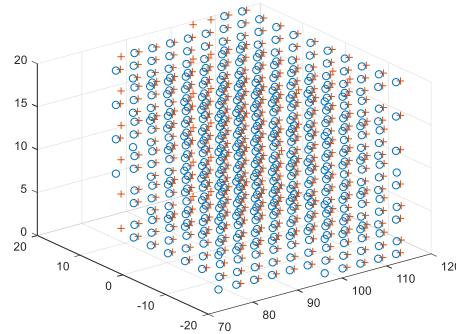


Figure 6. Reference point grid.

The reference points between different base stations are staggered, and the points to be located in the staggered area can receive positioning services from two or more base stations. If these base stations all perform separate positioning for the positioning point, the positioning result is consistent with the normal positioning. However, when the reference point networks overlap, these overlapped reference points can be used, and multiple base stations can cooperate with each other to share reference point information, and these interleaved reference points are collectively regarded as another new reference point network. Compared with the reference point network for single-point positioning, the new reference point network has a smaller grid spacing and can provide higher positioning accuracy.

3.2. Simulation and analysis

The simulation parameters under multi-point fusion positioning are consistent with the parameter settings during three-dimensional positioning. In order to simplify the model

and calculations, it is assumed that two base stations have generated the overlap of the reference point network, and the positioning results are shown in Figure.5(b).

Simulation analysis: It can be seen from the above figure that multi-point fusion positioning can effectively improve the positioning accuracy, because the overlapped reference point grid is equivalent to reducing the grid spacing after being regarded as a whole. If more base stations coincide with reference points, the positioning accuracy can be further improved. In theory, the reference point network overlap of n base stations can bring up to n times the increase in reference point density. This not only achieve higher positioning accuracy in the matching stage, but also in the acquisition stage can use high grid density based on machine learning algorithms to extract higher-order fingerprint features and improve fingerprint resolution.

4. Backward-assisted positioning

4.1. Theoretical basis

As shown in the figure, the actual position of the user is on one side of the line connecting several anchor points. Assuming that the positioning is performed at a fixed time interval T_B , the positions of the positioning correspond to the five points in the figure. The actual positioning result and the first i points are averaged to obtain the optimized result of the positioning point. For $(i+1)$ points including the actual positioning point of the user, they are distributed symmetrically around the positioning point. Compared with directly positioning the positioning point, the positioning result obtained after taking the average value can be optimized with a probability of $(1-(1/2)i)$.

As shown in Figure.7, the actual position of the user is on one side of the line connecting several anchor points. Assuming that the positioning is performed at a fixed time interval T_B , the positions of the positioning correspond to the five points in the figure. The actual positioning result and the first i points are averaged to obtain the optimized result of the positioning point. For $(i+1)$ points including the actual positioning point of the user, they are distributed symmetrically around the positioning point. Compared with directly positioning the positioning point, the positioning result obtained after taking the average value can be optimized with a probability of $(1-(1/2)i)$.

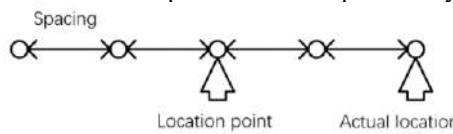


Figure 7. Backward-assisted positioning.

For the distance error between the positioning point and the actual position, a correction value is introduced to correct it. As mentioned above, most users in a short distance are moving in a straight line at a uniform speed, and the speed is v_B , so the correction value is $v_B * T_B$. For example, if the speed of human walking is 1.5m/s, the correction value is 1.5 T_B .

4.2. Simulation and analysis

The simulation parameters of the backward-assisted positioning are consistent with the above. Assuming that the user is walking in the positioning scene, the average walking speed is 1.5m/s. The fixed positioning time interval between two points is 0.4s, the number of auxiliary points is 4, and the positioning result is shown in Figure.5(c).

In this simulation test, the average positioning accuracy has increased by 16%. On this basis, the simulation results based on different positioning time intervals and the number of auxiliary positioning points are shown in Table 2.

Table 2. Positioning results based on different positioning intervals and auxiliary positioning points.

System positioning time (s)	Number of auxiliary positioning points	Improved average positioning accuracy
0.3	3	6%
0.3	5	16%
0.3	7	15%
0.6	3	13%
0.6	5	16%

Simulation analysis: If the time interval is too small or the auxiliary points are too few, the error will be reduced less, but the error caused by the additional correction value will be smaller. As the time interval and the auxiliary points increases, the accuracy of the backward-assisted positioning is gradually improved, but the error caused by the introduced correction value will become larger. And as the time interval further increases, the user's motion trajectory will become unpredictable, and the positions of the auxiliary points may not be symmetrically distributed, thus an additional error will be introduced at this time. Therefore, when performing backward-assisted positioning, it is necessary to find a balance between the above two situations. In addition, this method has strong versatility and can be used in positioning algorithms other than fingerprint positioning.

5. Conclusion

This paper studies the fingerprint 3D positioning algorithm, expands the antenna array, provides three reference point distribution models, and improves the fingerprint similarity standard JADSC, which improves the stability of the algorithm. After that, in the scenario of high base station deployment density, using the coincidence of the reference point network, a multi-point fusion positioning technology is proposed to further improve the positioning accuracy of the algorithm. Finally, a backward-assisted positioning method is proposed, which uses the position information of the positioned points to assist the positioning of the current position, and uses statistical laws to improve the positioning accuracy. This method has strong versatility.

References

- [1] Wang X, Gao L, Mao S, et al. CSI-based Fingerprinting for Indoor Localization: A Deep Learning Approach[J]. IEEE Transactions on Vehicular Technology, 2016, 66(1):763-776.

- [2] Wang X, Liu L, Lin Y and Chen X. A Fast Single-Site Fingerprint Localization Method in Massive MIMO System[C]// 2019 11th International Conference on Wireless Communications and Signal Processing.
- [3] Kupper A. Location-based Services: Fundamentals and Operation[J]. Jcms Journal of Common Market Studies, 2005, 49(5):923-947.
- [4] Kushki A, N Plataniotis K, N Venetsanopoulos A. Kernel-based Positioning in Wireless Local Area Networks[J]. IEEE Transactions on Mobile Computing, 2007, 6(6): 689-670.
- [5] S. S. Moosavi and P. Fortier, A Fingerprint Localization Method in Collocated Massive MIMO-OFDM Systems Using Clustering and Gaussian Process Regression[C]// 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall), Victoria, BC, Canada, 2020, pp. 1-5.
- [6] Sun X, Gao X, Ye G L, et al. Single-Site Localization based on a New Type of Fingerprint for Massive MIMO-OFDM Systems[J]. IEEE Transactions on Vehicular Technology, 2018, 67(7):6134-6145.
- [7] Evolved Universal Terrestrial Radio Access, Further Advancements for E-UTRA Physical Layer Aspects (Release 9) 3GPP Technical SpecificationT [J]. 3GPP TR36.814 v1.4.1, 2010,36(v2):1-107.

High-Performance Fingerprint Localization in Massive MIMO System

Yijie REN, Zhixing XIAO, Yuan TANG, Fei TANG, Xiaojun WANG¹ and Xiaoshu CHEN

School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University

Abstract. Location-based service (LBS) for both security and commercial use is becoming more and more important with the rise of 5G. Fingerprint localization (FL) is one of the most efficient positioning methods for both indoor and outdoor localization. However, the positioning time of previous research cannot achieve real-time requirement and the positioning error is meter level. In this paper, we concentrated on high-performance in massive multiple-in-multiple-out (MIMO) systems. Principal Component Analysis (PCA) is applied to reduce the dimension of fingerprint, so that the positioning time is about tens of milliseconds with lower storage. What's more, a novel fingerprint called Angle Delay Fingerprint (ADF) is proposed. Simulation result of the positioning method based on ADF shows the positioning error is about 0.3 meter and the positioning time is about hundreds of milliseconds, which is much better than other previous known methods.²

Keywords. Fingerprint localization, High-performance, Data dimension reduction, Massive MIMO

1. Introduction

Recently, with the rise of 5th Generation Mobile Communication Technology (5G), the characteristics such as high speed, low latency, large bandwidth and massive Internet of things make people enter the age of the Internet of everything fast. Whether it is auto-driving or smart home, or emergency rescue, they all rely on location base service deeply. However, tall buildings and trees rise up in the city, which makes the propagation environment of wireless signals complicated and changeable. Thus, high positioning performance of mobile terminals is required.

The traditionally employed positioning methods can be divided into two kinds. One is satellite positioning, such as Global Positioning System (GPS) and Big Dipper. The other is based on the received signal strength (RSS), the time difference of arrival (TDOA), angel of arrival (AOA) which assumes the wireless signal transmit along the line-of-sight (LOS) path [1-3]. However, the classic location techniques mentioned above are sensitive to surrounding environments and cannot perform well in dense outdoor circumstances.

¹ Corresponding author: Xiaojun WANG, No. 2, Southeast University Road, Nanjing, China; E-mail: wxj@seu.edu.cn.

² Foundation items: Social Development Projects of Jiangsu Science and Technology Department (No.BE2018704).

Compared with the traditional positioning methods, fingerprint localization is simple and efficient. It uses multipath information efficiently. The idea of fingerprint localization is that the propagation of wireless signal is strongly dependent on wireless environment, so that the positioning problem can be thought as a pattern recognition issue [4-7]. Fingerprint scheme can be divided into two phases. In offline phase, a database including fingerprints of reference points in target localization area is established. In online phase, we extract fingerprints and match them with the fingerprints in the database to get the estimated location.

There are two types of fingerprint. The first one is receiving signal strength information measured from access points (APs), which is simple and cost-effective. It is usually used in indoor positioning. The second type of parameters is multipath characteristics from wireless channels, including angel of arrival (AOA) [8], channel state information (CSI) [9], power delay profile (PDP) [10].

Massive multiple-in-multiple-out (MIMO) [11] is considered as one of the leading technologies of 5G, which can provide more multipath information for fingerprint localization. Traditional CSI fingerprint is extracted from time domain. Due to the application of massive MIMO, the information of time domain can be transformed into angle-delay domain, which improve the stability of fingerprint localization [12]. Angel delay Channel Power Matrix (ADCPM) is a kind of fingerprint based on massive MIMO system [13]. A deep learning approach based on ADCPM and a fast fingerprint localization method based on ADCPM were proposed in [14]. These researches based on ADCPM improve the performance of fingerprint positioning to some extent, but they are also limited.

In this paper, we focus on high-performance fingerprint localization problem in massive MIMO system. First, Principal Component Analysis (PCA) is used to reduce the dimension of ADCPM, which accelerates the online positioning process and reduces storage overhead. Second, Angle Delay Fingerprint (ADF) is proposed based on ADCPM. The average positioning error is about one-third of ADCPM. The average positioning time is about tens of milliseconds. The positioning performance is much better than previous methods based on ADCPM. Finally, many simulations are conducted to demonstrate the performance of the two new fingerprints.

The rest of the paper is organized as follow. In Section 2, the system model and fingerprint extraction method is stated. The compressed fingerprint localization method based on PCA and the positioning method based on ADF are presented in section 3. In section 4, simulation results are provided to demonstrate the performance of the proposed localization methods.

2. Fingerprint Extraction

In this section, the channel model and fingerprint extraction in the Angle-Delay domain are presented. The base station (BS) is equipped with N_t antennas in the form of uniform linear array (ULA), and by uplink channel estimation at the base station, CSI is considered to be known.

2.1 Channel Model

We assume that users are randomly distributed in the localization area, which is simple and without loss of generality. The channel model is the same as [13]. Firstly, the

Channel Impulse Response (CIR) of each user is given, where the multipath characteristic is included. To combine the AOA information, the array response vector corresponding to the AOA is considered. The scatters are independent and uniform in the target area. Then, the overall channel frequency response (CFR) of each user is generated by discrete Fourier transform (DFT). Detailed derivations are presented in [13].

2.2 Angle-delay Fingerprint Extraction

The fingerprint used for positioning is extracted from massive MIMO system, which indicates much more multipath information in both angle and delay domain. Using the CFR generated from previous section, the angle-delay fingerprint is easily extracted. By multiplying the DFT and unitary DFT matrix on the left and right, the CFR matrix is mapped from the frequency domain to the angle-delay domain, which is called angle-delay domain channel response matrix (ADCRM). Finally, in order to get more visualized data, we use the expectation of the Hadamard product of two ADCRM as the angle-delay channel power matrix (ADCPM). The derivation procedure and proof of ADCPM property is omitted, which can be referred to [13] for details.

3. Proposed Fast Fingerprint Localization Algorithms

In this section, two new fingerprint localization algorithms will be proposed. The process of fingerprint positioning includes online phase and offline phase. The overall process is shown as Figure.1.

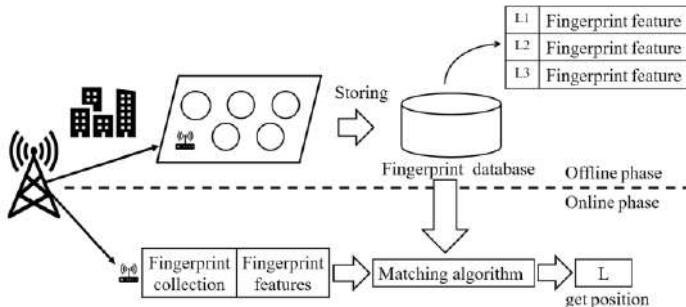


Figure 1. The process of fingerprint localization.

After establishing a fingerprint database in offline phase, the fingerprints extracted from unknown points to be positioned are matched with the fingerprints in database in online phase. Different fingerprint similarity criteria are available.

In this paper, we use Center Angle of Arrival (CAOA) clustering algorithm for data preprocessing, and use Weighted K Nearest Neighbor (WKNN) as the basic matching algorithm [13]. The similarity of two fingerprints is measured using the Manhattan Distance (MD) criterion which is shown in the following formula:

$$dist_{man}(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\|_1 = \sum_{u=1}^n |x_{iu} - x_{ju}| \quad (1)$$

In the formula, N is the dimension of the fingerprint, x_i and x_j are two n -dimensional vectors. The similarity of two fingerprints is higher, as the value of $dist_{man}(x_i, x_j)$ is smaller and vice-versa.

3.1 The Compressed Fingerprint Localization Method Based on PCA

ADCPM is a high-dimensional data. Using ADCPM in fingerprint positioning will bring high storage overhead and high computation cost in online matching phase. Data dimension reduction can extract efficient information from original data.

PCA is a widely used method for data dimension reduction. The main idea of PCA is that the original data can be reconstructed in feature space by eigenvalue decomposition. Thus, the n -dimensional data can be represented by k -dimensional eigenvalue. N is greater than k , so the original data can be compressed. The flowchart of PCA algorithm is shown in Table 1.

Table 1. The Flowchart of PCA Algorithm

1: Input: Original data $\{x_1, x_2, \dots, x_m\}$ 、 Low dimensional spatial dimensions d' ;
2: Centralize all samples: $x_i \leftarrow x_i - \frac{1}{m} \sum_{i=1}^m x_i$;
3: Calculate the sample covariance matrix XX^T ;
4: Decompose the covariance matrix XX^T by eigenvalue;
5: Pick the eigenvectors corresponding to the largest eigenvalues $w_1, w_2, \dots, w_{d'}$;
6: Output: The basic vectors of reconstructed plane $W = (w_1, w_2, \dots, w_{d'})$.

After extracting ADCPM from wireless channel, PCA algorithm is used for preprocessing. The PCA compressed fingerprint will be stored in database. The flowchart of PCA-based fingerprint positioning is shown in Table 2. The simulation will be presented in section 4.

Table 2. The Flowchart of PCA-based Compressed Fingerprint Localization

1	Offline phase:
1)	Fingerprint collection: Divide the target area into grids by definite intervals, collect ADCPM of every grid, use PCA algorithm to compress ADCPM;
2)	CAOA clustering: Calculate CAOA of every fingerprint in database, divide the fingerprint with same CAOA into the same set to form several sets and store them in database;
2	Online phase:
1)	Calculate CAOA of real-time fingerprint of the target terminal;
2)	Find the corresponding set according to CAOA;
3)	Match the fingerprint of target terminal using WKNN in the set and calculate the position of target terminal;

3.2 The Positioning Method Based on ADF

Angle delay fingerprint can be extracted from ADCPM easily. Compared with ADCPM, the localization performance based on ADF is much better. The idea of ADF-based localization will be introduced in this section.

3.2.1 ADF Extraction

ADCPM is the mapping of CIR in angle-delay domain, which is a high-dimensional matrix. Figure.2 shows the distribution of ADCPM channel power. The number of antennas is 64 and the number of cyclic prefix is 754.

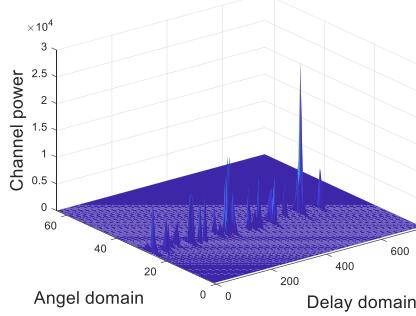


Figure 2. The distribution of ADCPM channel power

The channel power distributes mainly in coordinates 25~30 in angle domain. The peak power can be seen in coordinates 400 and 600 in delay domain. When ADCPM is used for fingerprint positioning, the positioning time is long with high computation cost. However, the positioning error is about several meters.

If all the elements in ADCPM are used for positioning, every single element will be calculated by Eq.(1). The elements of fewer channel power which are the majority in ADCPM will be treated the same as the elements of strong channel power which are the minority. Therefore, the positioning error will be affected.

In order to use the positioning information of ADCPM efficiently, the elements which have stronger channel power are selected for positioning, and those elements which have fewer channel power are omitted. The process of ADF extraction from ADCPM is shown in Figure.3.

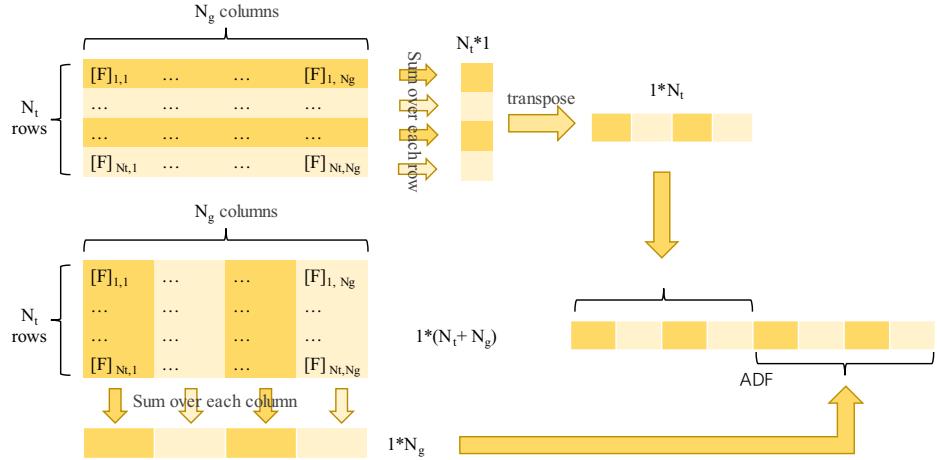


Figure 3. The process of ADF extraction from ADCPM

The extraction of ADCPM is introduced in Section 2 and can be referred to [14] for details. An ADCPM has N_t rows and N_g columns. Firstly, sum over each row and each column, separately. Then, transpose the $N_t * 1$ vector into $1 * N_t$ vector. Finally,

concatenate two vectors into $1*(N_t + N_g)$ which is the final ADF. The process of ADF extraction from ADCPM is easy. The positioning performance of ADF will be introduced in section 4.

3.2.2 The Characteristics of ADF

Compared with ADCPM, the data size of ADF is much smaller. For example, if an ADCPM is $64*754$, the ADF will be $(64+754)*1$. The Similarity performance of ADCPM and ADF with Manhattan Distance is shown in Figure.4, which shows that ADF has almost the same similarity performance with ADCPM.

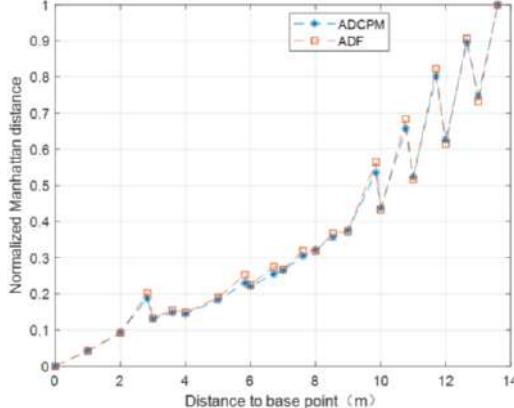


Figure 4. Similarity performance of ADCPM and ADF with Manhattan Distance

4. Simulation Result

4.1 Positioning Time and Storage

Average positioning time in online phase of three fingerprints is shown in Figure.5. The storage of fingerprint database of three fingerprint is shown in Figure.6. The number of antennas is 64. The interval of reference points is 1 meter. The density of scatters is $0.02/m^2$.

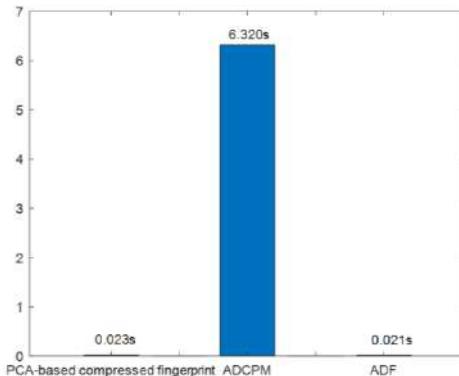


Figure 5. Average positioning time

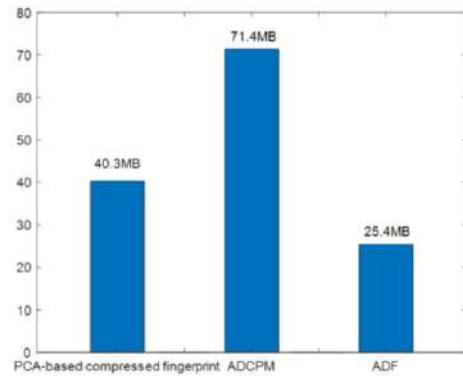


Figure 6. The storage of fingerprint database

Simulation analysis: The average positioning time of ADCPM is about 6 seconds, which cannot meet real-time requirements of Location-based service (LBS). However, the average positioning time of ADF and PCA-based compressed fingerprint is almost the same, which is about 0.02 seconds. The storage of ADCPM is the largest. However, the storage of ADF is about one-third of ADCPM and the storage of PCA-based compressed fingerprint is about 55% of ADCPM.

4.2 Simulations under Different Wireless Environments

Wireless environment is very complicated, which has great influence on positioning technology. In this section, we concentrated on positioning performance under different wireless environments, including different number of antennas, bandwidth and density of scatters.

In massive MIMO system, large number of antennas will be used. Average positioning error with different numbers of antennas is shown in Figure 7. The number of antennas is 8, 16, 32, 64, 128, 256, 512, respectively. The bandwidth of Orthogonal Frequency Division Multiplexing (OFDM) is 80MHz. The density of scatters is $0.02/m^2$. The interval of reference points is 1 meter.

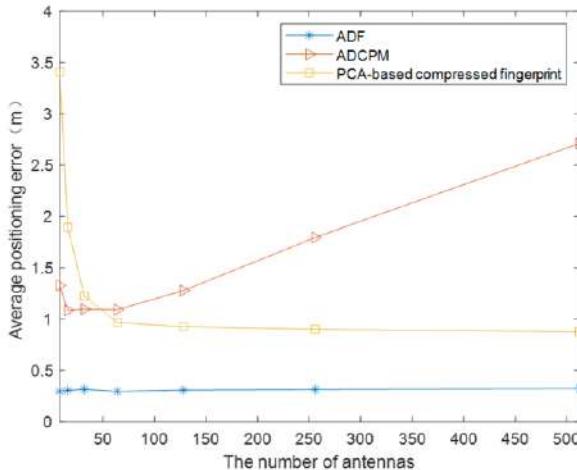


Figure 7. Average positioning error with different number of antennas

Simulation analysis: ADF has the best performance. When the number of antennas changes, average positioning error of ADF is about 0.3 meter and almost unchanged. However, with the increase of antennas' number, the positioning accuracy of ADCPM becomes worse, while the positioning accuracy of PCA-based compressed fingerprint becomes better. With more antennas, the dimension of ADCPM becomes bigger. The similarity performance of Manhattan distance becomes worse with the increase of weak-energy elements.

Average positioning error with different bandwidth is shown in Figure 8. The number of antennas is 64. The density of scatters is $0.02/m^2$. The interval of reference points is 1 meter.

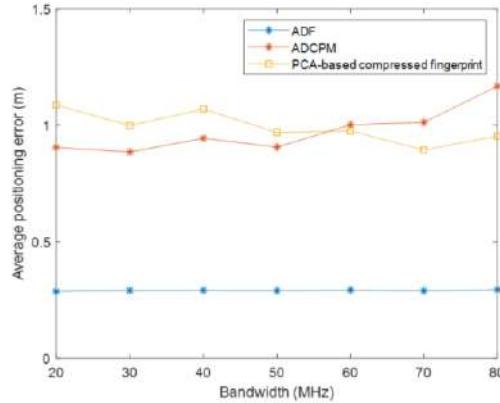


Figure 8. Average positioning error with different bandwidth

Simulation analysis: ADF has the smallest positioning error among these three fingerprints. The average positioning error of ADCPM and PCA-based compressed fingerprint is about 1 meter. With the increase of bandwidth, the positioning accuracy of ADCPM becomes worse and the positioning accuracy of PCA-based compressed fingerprint becomes better. This is the same situation as different number of antennas. With bigger bandwidth, elements of weak channel power take up a larger proportion of ADCPM. Therefore, the similarity performance of Manhattan distance becomes worse.

Finally, we investigate average positioning error with different density of scatters, and the simulation results is shown in Figure.9. The bandwidth of OFDM is 80MHz. The number of antennas is 64. The interval of reference points is 1 meter. The density of scatter changes from $0.005/m^2$ to $0.2/m^2$.

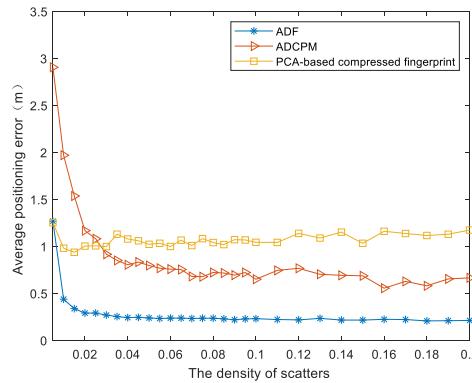


Figure 9. Average positioning error with different density of scatters

Simulation analysis: With the increase of the density of scatters, more CSI information will be provided due to multipath effect. Hence, the positioning error of three fingerprints becomes smaller. Within these three fingerprint, ADF has the best positioning accuracy. When the density of scatters is smaller than $0.03/m^2$, PCA-based compressed fingerprint is better than ADCPM. When the density of scatters is bigger than $0.03/m^2$, PCA-based compressed fingerprint performs worse. With more multipath information, ADCPM has better positioning accuracy, but in PCA algorithm, the compressed fingerprint remains in a low-dimensional space and the loss of effective positioning information is more.

5. Conclusion

In this paper, we have investigated high-performance positioning algorithms based on massive MIMO system. Firstly, the channel model is introduced and the extraction of ADCPM fingerprint is introduced. Then, two kinds of new fingerprint are proposed. One is PCA-based compressed fingerprint, which use PCA algorithm to extract efficient information from ADCPM. Simulation results show that PCA-based compressed fingerprint positioning can reduce average positioning time and the storage overhead. However, the positioning performance of ADF is much better. The average positioning error is about 0.3 meter, which is one third of ADCPM. What's more, the positioning time is about tens of milliseconds and the data size of fingerprint database is about one third of ADCPM. Finally, numerical results demonstrate the performance of those two new fingerprints.

References

- [1] G. Wang, H. Chen, Y. Li and N. Ansari, "NLOS Error Mitigation for TOA-Based Localization via Convex Relaxation," in *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4119-4131, Aug. 2014, doi: 10.1109/TWC.2014.2314640.
- [2] X. Guo, Y. Huang, B. Li and L. Chu, "DOA estimation of mixed circular and non-circular signals using uniform circular array," 2014 7th International Congress on Image and Signal Processing, Dalian, 2014, pp. 1043-1047, doi: 10.1109/CISP.2014.7003933.
- [3] G. Wang, Y. Li and N. Ansari, "A Semidefinite Relaxation Method for Source Localization Using TDOA and FDOA Measurements," in *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 853-862, Feb. 2013, doi: 10.1109/TVT.2012.2225074.
- [4] Honkavirta, Ville , et al. "A comparative survey of WLAN location fingerprinting methods." *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on IEEE*, 2009.
- [5] Zhang, Lei , et al. "A novel fingerprinting using channel state information with MIMO-OFDM." *Cluster Computing* (2017).
- [6] X. Wang, L. Gao, S. Mao and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," 2015 IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, 2015, pp. 1666-1671, doi: 10.1109/WCNC.2015.7127718.
- [7] Y. Lin, P. Tseng, Y. Chan and G. Wu, "Super-Resolution-Aided Positioning Fingerprinting Based on Channel Impulse Response Measurement," 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, 2017, pp. 1-6, doi: 10.1109/WCNC.2017.7925866.
- [8] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich and M. Coulon, "Direct Localization for Massive MIMO," in *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475-2487, 15 May 15, 2017, doi: 10.1109/TSP.2017.2666779.
- [9] H. Xie et al., "Accelerating Crowdsourcing Based Indoor Localization Using CSI," 2015 IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS), Melbourne, VIC, 2015, pp. 274-281, doi: 10.1109/ICPADS.2015.42.
- [10] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," in *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, November 2010, doi: 10.1109/TWC.2010.092810.091092.
- [11] Larsson E G, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems[J]. *IEEE Communications Magazine*, 2014, 52(2):186-195.
- [12] Alamu O, Iyaomolore B, et al. An overview of massive MIMO localization techniques in wireless cellular networks: Recent advances and outlook[J]. *Ad Hoc Networks*, 2021, 111:102353.
- [13] Sun X, Gao X, et al. Single-Site Localization Based on a New Type of Fingerprint for Massive MIMO-OFDM Systems[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(7): 6134-6145.
- [14] Sun X, Wu C, et al. Fingerprint-based Localization for Massive MIMO-OFDM System with Deep Convolutional Neural Networks[J]. *IEEE Transactions on Vehicular Technology*, 2019, PP (99):1-1.

Scattering Clustering Method for Terminal Fingerprint Positioning Based on MIMO Base Station

Fei TANG, Yijie REN, Xiaojun WANG¹, Weiguang SUN and Xiaoshu CHEN

School of Information Science and Engineering, National Mobile Communications Research Laboratory, Southeast University, China

Abstract. With the rapid development of wireless communication technology, location-based services are playing an increasingly important role in people's lives. However, as the living environment becomes more and more complex, the existence of obstructions and various scatterers makes the accuracy of traditional positioning algorithms decrease, thus, fingerprint positioning has gradually become a research hotspot in the field of positioning.

This paper researches the 5th Generation (5G) fingerprint location method based on machine learning. A massive multiple-in multiple-out (MIMO) channel is constructed on the MATLAB simulation platform, from which the fingerprint information is extracted to establish a fingerprint database. Considering the huge amount of data in the fingerprint database, and under the multipath effect, the channel characteristics are mainly affected by the scatterers near the point to be located. This paper proposes a scattering-based clustering method that combines the particularity of multipath propagation for clustering. Research shows that this method has excellent clustering effects, which can effectively improve algorithm efficiency and reduce data storage pressure on the base station side.²

Keywords. Massive MIMO, Fingerprint extraction, Clustering

1. Introduction

In recent years, with the development of wireless communications, location-based services have attracted widespread attention in various industries. In outdoor scenes, especially in open areas, traditional positioning algorithms such as Global Positioning System (GPS) [1] have sufficient positioning accuracy. However, in some urban areas with tall buildings, or inside buildings, due to the influence of occlusion and non-line-of-sight transmission, the positioning effect of the traditional positioning algorithm is very limited [2]. In addition, in modern society, most of people's activities are concentrated indoors. New user needs make traditional positioning algorithms face new challenges, and fingerprint positioning technology emerges at the historic moment.

Different from the traditional positioning algorithms, the fingerprint positioning algorithm uses some corresponding characteristics generated during the signal

¹ Corresponding author: Xiaojun WANG, No. 2, Southeast University Road, Nanjing, China; E-mail: wxj@seu.edu.cn

² Foundation items: Social Development Projects of Jiangsu Science and Technology Department (No.BE2018704).

transmission to achieve positioning. In the process of wireless signal transmission, due to the influence of multipath effects, the channel characteristics at different physical locations will reflect some unique characteristics, which called the fingerprints of the signal. When there is a positioning requirement, the fingerprint information of the point to be located is compared with the fingerprint information in the database, and the information of the point is estimated according to a certain similarity criterion. Through the above-mentioned series of processing, the fingerprint positioning algorithm transforms a positioning problem into a pattern recognition problem. Literature [3-4] use channel state information to achieve reliable fingerprint positioning, and literature [5-6] combine the 5th Generation (5G) massive multiple-in multiple-out (MIMO) scenarios to explore the application prospects of fingerprint positioning in the new context. The flow chart of the fingerprint location algorithm is shown in Figure.1 [7-8].

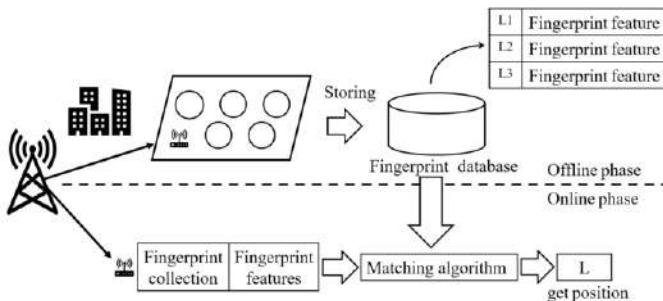


Figure 1. The flow chart of the fingerprint location algorithm.

This paper builds a massive multiple-in multiple-out-Orthogonal Frequency Division Multiplexing (MIMO-OFDM) system based on a simulation platform, and uses a scattering-based clustering method to complete fingerprint location. The main work is as follows:

- STEP 1: Build a massive MIMO-OFDM system, extract information from the channel and perform further processing to obtain fingerprints. The fingerprints make full use of the multipath effect and are easy to extract;
- STEP 2: Propose a clustering method based on scattering, which can not only improve the efficiency of the algorithm, but also ensure the positioning accuracy of the algorithm;
- STEP 3: Based on the similarity criterion of the Joint Angle Delay Similarity Coefficient (JADSC), the Weighted K Nearest Neighbor (WKNN) algorithm is used to improve the positioning accuracy of the algorithm.

2. Massive MIMO Channel Model

In the massive MIMO-OFDM single station system, the channel is constructed based on the ray tracing model, as shown in Figure.2.

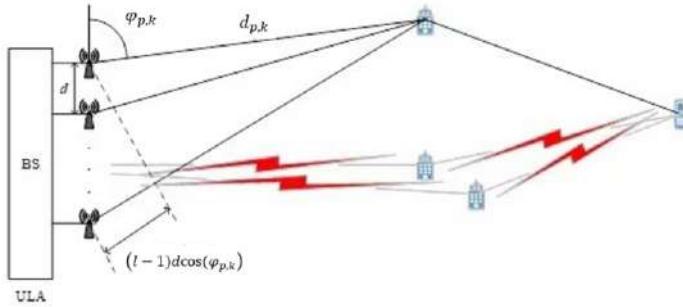


Figure 2. Massive MIMO-OFDM single station system.

The base station is equipped with N_t antennas in the form of Uniform Linear Array (ULA), the distance between each antenna element is d , the number of users is K , and its distribution law obeys random distribution, which satisfies $N_t \gg K$. Due to the influence of the scatterer, the signal will propagate in the channel in the form of multipath, and the base station estimates the information from the user k to the base station through the uplink channel in the communication link. In summary, the Channel Impulse Response (CIR) of user k on the p -th path can be expressed as:

$$q_{p,k} = a_{p,k} e(\varphi_{p,k}) \exp\left(-j2\pi \frac{d_{p,k}}{\lambda_c}\right) \quad (1)$$

Notation: $a_{p,k} \sim \mathcal{CN}(0, \delta_{p,k})$ represents the complex channel gain of the k -th user on the p -th path, $\varphi_{p,k} \in (0, \pi)$ represents the angle of arrival of the signal, $d_{p,k}$ represents the physical distance from the antenna element of the transmitting signal on the p -th path to the first receiving antenna element of the base station, λ_c represents the carrier wavelength, $e(\varphi_{p,k})$ represents the corresponding array response vector, as shown in the following formula:

$$e(\varphi_{p,k}) = \left[1, \exp\left(-j2\pi \frac{d \cos(\varphi_{p,k})}{\lambda_c}\right), \dots, \exp\left(-j2\pi \frac{(N_t-1)d \cos(\varphi_{p,k})}{\lambda_c}\right) \right]^T \quad (2)$$

Define a discrete Fourier phase shift matrix $V \in \mathbb{C}^{N_t \times N_t}$, which satisfies the following formula:

$$[V_{i,j}] \triangleq \frac{1}{\sqrt{N_t}} \exp\left(-j2\pi \frac{i(j-N_t/2)}{N_t}\right) \quad (3)$$

The phase shift matrix can map the CIR from the antenna domain to the corresponding angle domain. Considering that the Time of Arrival (TOA) of each path in multipath transmission is different, the total CIR of user k can be expressed as the sum of the CIR of all paths:

$$\mathbf{q}_k(\tau) = \sum_{p=1}^P \mathbf{q}_{p,k} \cdot \delta(\tau - \tau_{p,k}) \quad (4)$$

In the formula, $\tau_{p,k} = d_{p,k}/v$ represents the TOA of each path in the multipath, and v represents the speed of light. Sampling the antenna array output based on the sampling interval T_s , then the symbol interval of OFDM can be expressed as $T_c = N_c T_s$.

Through the modulation of the OFDM system, the frequency selective fading channel caused by the multipath effect can be transformed into a frequency flat channel, and then time domain sampling is performed to obtain the TOA of each path. According to the mathematical relationship between channel responses, the Channel Frequency Response (CFR) of the l -th subcarrier can be expressed as the Fourier transform of its CIR:

$$\begin{aligned} \mathbf{h}_{k,l} &= \sum_{p=1}^P \sum_{n=0}^{N_c-1} a_{p,k} e(\varphi_{p,k}) \exp(-j2\pi f_l \tau_{p,k}) \cdot \delta(nT_s - \tau_{p,k}) \\ &= \sum_{p=1}^P a_{p,k} e(\varphi_{p,k}) \exp\left(-j2\pi \frac{l \cdot n_{p,k}}{N_c}\right) \end{aligned} \quad (5)$$

In the formula, $n_{p,k}$ represents the propagation delay of the p -th path in multipath propagation, and $[x]$ represents the integer closest to x . The overall CFR matrix of user k is composed of the corresponding CFRs of all its subcarriers, as shown in the following formula:

$$\mathbf{H}_k = [\mathbf{h}_{k,0}, \mathbf{h}_{k,1}, \dots, \mathbf{h}_{k,L-1}] \quad (6)$$

3. Clustering Algorithm and Terminal Positioning

In fingerprint positioning, because the fingerprint database is generally large, traversing the entire fingerprint database will consume a lot of resources. Therefore, before online matching, clustering algorithms are used to classify reference points. Considering that fingerprint features are mainly affected by adjacent scatterers, this chapter proposes a scattering-based clustering method to improve the clustering effect, and finally uses WKNN in the terminal positioning part to improve the accuracy of the algorithm.

3.1 Clustering method based on scattering

As one of the classic machine learning, clustering algorithm mainly includes K-means and fuzzy C-means method, among which the fuzzy C-means method belongs to the fuzzy clustering algorithm. It can be seen from the above that the fingerprint feature at a certain location is mainly determined by the nearest scatterers around it. Taking into account the complexity of the multipath scene, the number of scatterers that affect the reference point can be correlated with its physical location, and the number of scatterers affected by the signal at a certain location is used as a standard to classify the fingerprint data.

The **scattering-based clustering algorithm** refers to the fuzzy clustering algorithm. The same reference point can belong to different categories. The membership degree in the fuzzy clustering algorithm is the number of scatterers that will affect the signal here,

which is called m_c . Record the membership degree of each reference point, establish the mapping between the membership degree and the physical location of the reference point, and set the cluster center according to a certain step:

$$O_i = O_0 + i^* s_{CBS} \quad (7)$$

Notation: O_i is the i -th cluster center, O_0 is the initial cluster center, which satisfies $O_0 = \min(m_c) + s_{CBS}$, and s_{CBS} is the step size of the cluster center.

According to the recorded membership degree, set a certain numerical diffusion for the membership degree center to perform clustering:

$$m_C = [m_c - \Delta m, m_c + \Delta m] \quad (8)$$

In the formula, Δm is the numerical diffusion of the degree of membership. The fingerprint matrix of Clustering Based on Scattering (CBS) is as follows:

$$\mathbf{F}_{CBS} \triangleq \mathbb{E} \left\{ \mathbf{M}_{CIR} \odot \mathbf{M}_{CIR}^* \right\} \in \mathbb{R}^{N_r \times N_s} \quad (9)$$

Notation: \mathbf{M}_{CIR} is the total CIR matrix of the user, and \mathbf{F}_{CBS} is the final fingerprint matrix in the CBS.

3.2 Terminal positioning

3.2.1 Fingerprint Similarity Criteria.

For the calculation of the distance between fingerprints, due to the needs of the positioning itself and the consideration of the practicality of the algorithm, the overall distance function between the fingerprints should show a monotonous increasing or monotonous decreasing trend. In this section, the Joint Angle Delay Similarity Coefficient (**JADSC**) is introduced to measure the similarity between fingerprints [5].

For the joint angle delay similarity coefficient at position i and position j , it is defined as:

$$J_{ij} = \max_{n \in (-L+1, L-1)} \left(\frac{1}{N_i^2} [\mathbf{C}_i]^T [\mathbf{C}_j]_{t-n} \right) \quad (10)$$

Notation: Matrix \mathbf{C}_i represents the fingerprint matrix at position i , $[\mathbf{C}_i]_t$ represents the t -th column of the fingerprint matrix, n represents the delay compensation between two fingerprints, and L represents the maximum delay compensation.

From the above, in the massive MIMO-OFDM system, the base station is usually deployed in a high physical location. When the number of antennas increases, the JADSC is as follows:

$$J_{ij} = \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \sigma_{p_1,i} \sigma_{p_2,j} \delta(\varphi_{p_1,i} - \varphi_{p_2,j}) \delta(n_{p_1,i} - n_{p_2,j}) \quad (11)$$

In the formula, P represents the nearest P scatterers that affect the fingerprint characteristics, and σ represents the channel energy of the corresponding scatterers. It can be further inferred from the above formula that the JADSC is mainly determined by the overlap of the scatterers of the fingerprints at the two positions. The simulation diagram is shown in Figure 3.

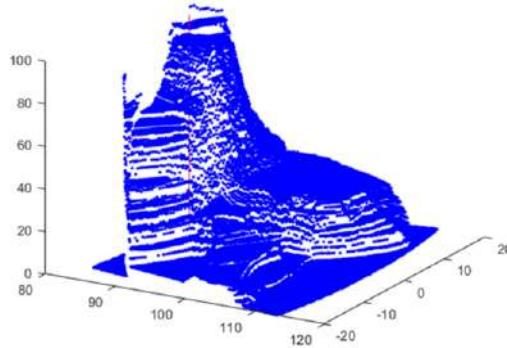


Figure 3. Schematic diagram of JADSC.

In the figure, each point corresponds to a reference point, the X-axis and Y-axis coordinates represent the physical position of the reference point, and the Z-axis value represents JADSC. The straight line perpendicular to the XOY plane is a randomly selected point, that is, the actual point to be located.

The simulation test proves that the reference points with larger JADSC are all near the points to be located, and the overall distribution of JADSC shows a monotonous decreasing trend, thus the JADSC meet the conditions of similarity standards between fingerprints.

3.2.2 Terminal positioning algorithm.

Nearest Neighbor (NN) is a basic classification and regression method in machine learning, mainly including K-Nearest Neighbor (KNN) and weighted K-Nearest Neighbor (WKNN). The KNN method contains two basic elements: distance measurement and the choice of K value. This article uses JADSC as the distance measurement standard. In the selection of the K value, if the value is small, it may lead to over-fitting; if the value is too large, it will cause too many points to be substituted into the calculation, resulting in a decrease in accuracy and causing under-fitting. The calculation formula for positioning by KNN method is as follows:

$$\hat{p}_{x,y} = \frac{1}{K} \sum_{k=1}^K p_{x,y}^k \quad (12)$$

In the formula, $\hat{p}_{x,y}$ represents the position coordinates of the point to be located, and $p_{x,y}^k$ represents the coordinates of the k-th reference point.

The operation of the KNN algorithm is simple and easy to implement. When a more reasonable K value is selected, the positioning accuracy is also guaranteed to a certain extent. However, the selected K reference points have the same weight, which makes the KNN algorithm susceptible to some outliers.

Therefore, this paper proposes a better algorithm: WKNN algorithm. The steps of the WKNN algorithm are roughly the same as KNN, but when substituting the reference point position coordinates, a weight is set for each reference point. The more similar the reference point to the point to be located, the greater the impact on the positioning result, which ensures the stability of the algorithm. The calculation formula of WKNN is:

$$\hat{p}_{x,y} = \sum_{k=1}^K w_k \cdot p_{x,y}^k, \quad w_k = \frac{1}{\sum_{k=1}^K \frac{1}{d_k}} \quad (13)$$

In the formula, w_k represents the weight corresponding to the k -th reference point, and its value satisfies $\sum w_k = 1$, and d_k is the JADSC of the k -th reference point and the point to be located.

The nearest neighbor method and related improved algorithms do not need to train the model in advance, and can be used directly during matching, but it will consume a lot of time to traverse the fingerprint database during online matching. In specific use, it can be combined with the clustering algorithm to greatly improve the efficiency of the algorithm.

4. Simulation and Analysis

4.1 Simulation parameter setting

The simulation is based on MATLAB to build a massive MIMO-OFDM single station system. The cell radius is $R=500m$, the radius of the area to be located is $R_{loc}=300m$, the width is $W_{loc}=100m$, and the angle range is $A_{loc}=9/\pi$. The base station is located in the center of the cell. The position of the scatterer is shown in Figure.4.

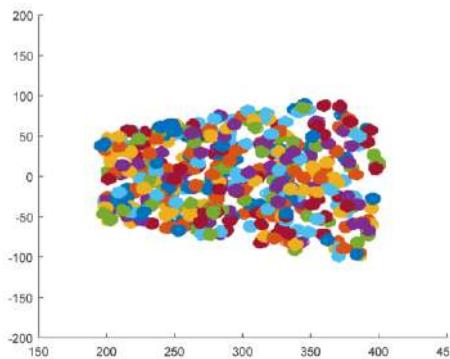


Figure 4. Scatterer model.

The reference points are divided at an interval of 2m, evenly distributed in the area to be located. When collecting fingerprint information, it is necessary to traverse all reference points. Then compress the fingerprint information, and finally store it in the

fingerprint database together with the corresponding position information. The reference point distribution model is shown in Figure.5.

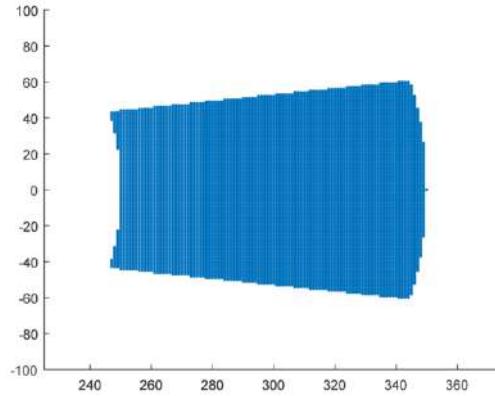


Figure 5. The reference point distribution model.

Other relevant parameters are shown in Table 1.

Table1. System Simulation Parameters

OFDM parameters	Transmission bandwidth B_W	5-80MHz
	Subcarrier spacing Δf	15kHz
	Pilot interval T_g	4.7 μs
	Sampling interval T_s	1/(2*BW)
Antenna parameters	Number of antennas N_t	128
	The location of the base station BS_{pos}	[0;0]
Other parameters	Scatterer radius R_{Scat}	20m
	Scatterer density	0.02/ m^2
	Subpath N_{Sub}	40
	Test points N	1000

4.2 Simulation results

Perform a simulation test based on the WKNN algorithm to select the best K value for subsequent comparison and analysis. The simulation result based on different K values in the WKNN algorithm is shown in Figure.6.

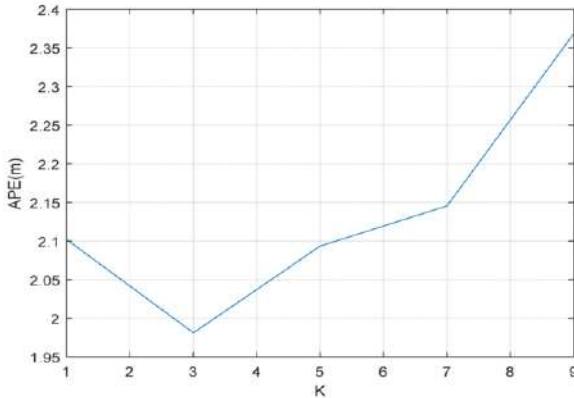


Figure 6. Average positioning accuracy based on different K values in WKNN.

Simulation analysis: In the WKNN algorithm, when the K value is small, the number of reference points is small, which brings greater uncertainty to the result, resulting in loss of some positioning accuracy. When the K value is large, there are too many reference points, and some reference points that are not particularly correlated are also included in the calculation of WKNN, resulting in a decrease in the accuracy of the algorithm. It can be seen from the figure above that the optimal K value should be 3.

The simulation test is carried out around the clustering algorithm to verify the performance of the clustering algorithm proposed in this paper. The positioning result of the two clustering algorithms is shown in Figure 7.

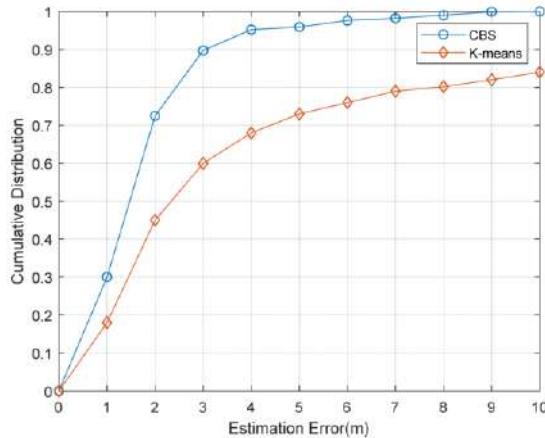


Figure 7. Simulation result of the clustering algorithm.

Simulation analysis: As can be seen from the figure above, the scattering-based clustering algorithm proposed in this paper has more superior performance. The K-means algorithm is a hard-clustering algorithm. When calculating JADSC, due to the large number of reference points, the JADSC value is approximately continuous, and the decrease in certain distances is not obvious, which leads to the low similarity of the searched similar points. In addition, the K-means algorithm uses the reference point selected by the cluster as the standard. The clustering effect of points near the reference point is acceptable, but for other reference points, there may be some errors in the

clustering division. In the simulation data, there are many high error points, and the error value may be close to 10m.

5. Conclusion

This paper constructs a massive MIMO-OFDM system model, extracts fingerprints based on CIR information, and associates location information with fingerprint information to establish a fingerprint database. Due to the huge amount of data in the fingerprint database, after introducing the classic clustering algorithm and combining with specific scenarios, this paper proposes a scattering-based clustering algorithm to improve the performance of the algorithm. In the terminal positioning part, the proposed clustering algorithm is verified by simulation test based on WKNN algorithm, which proves its superiority.

References

- [1] Angjelicinoski M, Atanasovski V, Gavrilovska L. Comparative analysis of spatial interpolation methods for creating radio environment maps[C]// Telecommunications Forum. IEEE, 2012:334-337.
- [2] Parkinson B W, Gilbert S W. NAVSTAR: global positioning system-ten years later[J]. Proceedings of the IEEE, 1983, 71(10):1177-1186.
- [3] Wang X, Gao L, Mao S, et al. CSI-based Fingerprinting for Indoor Localization: A Deep Learning Approach[J]. IEEE Transactions on Vehicular Technology, 2016, 66(1):763-776.
- [4] Wang X, Gao L, Mao S. CSI Phase Fingerprinting for Indoor Localization with a Deep Learning Approach[J]. IEEE Internet of Things Journal, 2016, 3(6):1113-1123.
- [5] Sun X, Gao X, Ye G L, et al. Single-Site Localization based on a New Type of Fingerprint for Massive MIMO-OFDM Systems[J]. IEEE Transactions on Vehicular Technology, 2018, 67(7):6134-6145.
- [6] Wang X, Liu L, Lin Y and Chen X. A Fast Single-Site Fingerprint Localization Method in Massive MIMO System[C]// 2019 11th International Conference on Wireless Communications and Signal Processing.
- [7] Kupper A. Location-based Services: Fundamentals and Operation[J]. Jcms Journal of Common Market Studies, 2005, 49(5):923-947.
- [8] Kushki A, N Plataniotis K, N Venetsanopoulos A. Kernel-based Positioning in Wireless Local Area Networks[J]. IEEE Transactions on Mobile Computing, 2007, 6(6): 689-670.

A Novel OLTC Fault Diagnosis Method Based on Optimized Long Short-Term Memory Parameters

Yan YAN¹, Hongzhong MA

College of Energy and Electrical Engineering, Hohai University, China

Abstract. Recently, long short-term memory (LSTM) networks have been widely adopted to help with fault diagnosis for power systems. However, the parameters of LSTM networks are determined by prior knowledge and experience and thereby not capable of dealing with unexpected faults in volatile environments. In this paper, we propose and apply an improved grey wolf optimization (IGWO) algorithm to optimize the parameters of LSTM networks, aiming to circumvent the drawback of empirical LSTM parameters and enhance the fault diagnosis accuracy for on-load tap changers (OLTCs). The composite multiscale weighted permutation entropy and energy entropy yielded by the grasshopper optimization algorithm and variational mode decomposition (GOA-VMD) method are used as the inputs of LSTM networks. The IGWO algorithm is applied in an iterative manner to optimize the relevant super arithmetic of the LSTM. In this way, an IGWO-LSTM combination model is constructed to classify different faults diagnosed in OLTCs. Experimental results verify the diagnosis performance superiority of the proposed method over several widely used comparison benchmarks

Keywords. Long short-term memory (LSTM) network, fault diagnosis, on-load tap changer (OLTC), parameter optimization, improved grey wolf optimization (IGWO)

1. Introduction

On-load tap chargers (OLTCs) are important components in transformers but are relatively fault-prone compared to other components. In recent years, it has been reported that the faults in OLTCs have accounted for more than 20% of transformer faults, and this percentage keeps increasing [1]. OLTCs are mechanical components in transformers, and therefore when switching OLTCs, complex friction and collision among components generate random and nonlinear vibration signals [2]. As a result, one can process and study these vibration signals to extract features, summarize patterns, and finally diagnose operational statuses [3]. Time-frequency methods are suitable for the analysis and processing of non-stationary signals and can characterize the relationship between signal frequency series and time [4]. Among them, the wavelet analysis is widely used for OLTC fault diagnosis by virtue of its multi-resolution, high accuracy, and small reconstruction error [5]. However, due to the arbitrary selection of wavelet bases, the analytical results by time-frequency analysis can hardly characterize mechanical statuses in an accurate manner. As an alternative, mode decomposition methods, e.g., empirical

¹ Corresponding author; Email: yanyan503986706@163.com

mode decomposition [6] and variational mode decomposition [2], can extract vibration characteristic information of OLTCs quickly and efficiently. However, it is still challenging to cope with randomness and real-time diagnosis requirements for OLTCs by these mode decomposition methods. Recently, with the advances in artificial intelligence (AI) technologies, researchers also resort to hidden Markov model (HMM) [7], support vector machine (SVM) [8], convolutional neural network (CNN) [9] to help with OLTC fault diagnosis. Among these technologies, HMM cannot yield synchronized data, and thereby features are easily submerged in noise. SVM can be greatly affected by internal parameters and relies on an extra optimization process with higher computational complexity. The main issues with CNN are the low training efficiency and manually adjusted parameters. Different from the aforementioned AI technologies, the long short-term memory (LSTM) is a special architecture of recurrent neural network (RNN) and has been proven effective for fault diagnosis of power systems [10]. On the other hand, the LSTM parameters are determined empirically based on previous knowledge and experience and thereby not capable of dealing with unexpected faults in volatile environments. In this regard, we propose and apply an improved grey wolf optimization (IGWO) algorithm to optimize the parameters of LSTM networks, aiming to circumvent the drawback of empirical LSTM parameters and enhance the fault diagnosis accuracy for OLTCs. In particular, the proposed IGWO algorithm inputs composite multiscale weighted permutation entropy and energy entropy yielded by the grasshopper optimization algorithm and variational mode decomposition (GOA-VMD) method to LSTM networks and optimizes relevant super arithmetic of the LSTM on an iterative basis. In this way, we can construct an IGWO-LSTM model to efficiently classify different faults diagnosed in OLTCs. The rest of the paper is organized as follows. We first propose the novel OLTC fault diagnosis method and expatiate on the optimization of LSTM parameters in Section II. Experimental setups are presented in Section III, analysis of experimental results in Section III. Finally, the paper is concluded in Section V.

2. IGWO-LSTM fault diagnosis based on multi-feature fusion

2.1. Feature fusion based on CMWPE

The GOA-VMD method lies on the core of the proposed the OLTC fault diagnosis method. Specifically, the GOA [11] is used to optimize the number of intrinsic mode function (IMF) components and a penalty factor yielded by VMD, which are denoted as K and γ . It is stipulated by that the minimum envelop entropy of a vibration signal detected from the OLTC denoted as $\min E_p$ is adopted as the fitness function the fault diagnosis process

In order to further verify the effectiveness of the proposed method, CMWPE is introduced into the multi-scale analysis. CMWPE [12] is also a measure of system complexity. The specific calculation steps are as follows:

- ① input vibration signal $x(i)(i=1,2,3,\dots,N)$, the number of embedded bits is τ , and the delay time is m ;
- ② initialization $CMWPE = \emptyset$, $s = 1$;
- ③ for $x(i)(i=1,2,3,\dots,N)$ coarse-grained time series $y^{s,d} = \{y^{s,d}(y(j))\}_{j=1}^{[(N+1)/s]-1}$ can be expressed as:

$$y^{s,d}(j) = \frac{1}{s} \sum_{i=(j-1)s+d}^{js+d-1} x(i) \quad (1)$$

Where, s is scale factors, $[N + 1]/s$ represents the a ceiling function, $j = 1, 2, \dots, [(N+1)/s] - 1$, and $d < s \leq s_{\max}$, $d = 1, 2, \dots, s$, $i = 1, 2, \dots, js + d - 1$

④ CMWPE can be expressed as:

$$\text{CWMPE}(Y, m, \tau, s) = \frac{1}{s} \sum_{d=1}^s \text{WPE}(y^{s,d}, m, \tau) \quad (2)$$

Where, CMWPE is the average WPE value

⑤ CMWPE = CMWPE \cup CMWPE(Y, m, τ, s), $s = s + 1$, repeat steps 2-4, until $s > S_{\max}$.

The vibration signal of OLTC is decomposed by GOAVMD to obtain 1 IMF. The CMWPE value and energy value are calculated respectively, and then two of them are fused to form the feature vector. The matrix of CMWPE is as follow:

$$H = \begin{bmatrix} H_{1,1} & H_{1,2} & \cdots & H_{1,s} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ H_{L,1} & H_{L,2} & \cdots & H_{L,s} \end{bmatrix} \quad (3)$$

Where, $H_{\eta,\eta}$ represents the η value of the IMF element with scale η . $\forall \eta \in \{1, 2, \dots, l\}$ and $\forall \eta \in \{1, 2, \dots, s\}$, then add an average of 1 to l elements in each H column vector, which is expressed as follows: $\bar{H} = [\bar{H}_1, \bar{H}_2, \dots, \bar{H}_s]$, $\bar{H}_\eta = \frac{1}{l} \sum_{\eta=1}^l H_\eta$. The energy eigenvector is expressed as $B = [E_{l,1}, E_{l,2}, \dots, E_{l,J}]$. Where $E_{l,j}$ represents the i -th energy eigenvalue corresponding to the j -th IMF component, and $E_i = \int_{-\infty}^{+\infty} |c_i(t)|^2 dt$. where $c_i(t)$ represents the magnitude of each scale coefficient or natural mode function at time t . The energy characteristic B is standardized, and is expressed as follows:

$$B' = [E_{l,1} / E_\Sigma, E_{l,2} / E_\Sigma, \dots, E_{l,J} / E_\Sigma] \quad (4)$$

$$\text{Where, } E_\Sigma = \sqrt{\sum_{j=1}^l |E_j|^2}$$

All the features are fused to form the feature matrix $D = [\bar{H}, B']$. In this paper, the sample signal with length N is divided into z segments. Each segment has 1024 samples so that the characteristic samples can be expressed as

$$D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_z \end{bmatrix} = \begin{bmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,J} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ d_{z,1} & d_{z,2} & \cdots & d_{z,J} \end{bmatrix} \quad (5)$$

Where, $J = s + l$ represents the dimension of solving the problem, and Z represents the number of the Grey Wolf population, which is applied in the next scene.

2.2. The improved Grey Wolf Optimization

The Grey Wolf Optimization(GWO) [13] algorithm simulates the social system and hunting behavior of a wolf pack, and the optimal solution is set as α wolf; the second

optimal solution is β wolf, the third optimal solution is σ wolf, and the remaining solution is ω wolf. The hunting process of wolves can be divided into three stages: the stage of tracking and approaching the prey; surrounding and harassing; and the stage of attacking the prey, and the mathematical model of surround behavior is:

$$Q(t+1) = Q_p - A |C \cdot Q_p(t) - Q(t)| \quad (6)$$

Where t is the current iteration number, C and a are coefficients, a is the distance control parameter, $C = 2r_1$, $a = 2ar_1 - a$, Q_p is the position vector of the prey, Q is the position of the grey wolves, and r_1 and r_2 are random variables uniformly distributed in $[0,1]$. The mathematical model of the tracking target is as follows:

$$\begin{cases} Q_1(t+1) = Q_\alpha(t) - A_1 |C_1 \cdot Q_\alpha(t) - A_1 \cdot Q(t)| \\ Q_2(t+1) = Q_\beta(t) - A_2 |C_2 \cdot Q_\beta(t) - A_2 \cdot Q(t)| \\ Q_3(t+1) = Q_\delta(t) - A_3 |C_3 \cdot Q_\delta(t) - A_3 \cdot Q(t)| \end{cases} \quad (7)$$

$$Q(t+1) = \frac{Q_1(t) + Q_2(t) + Q_3(t)}{3} \quad (8)$$

In the hunting process, α , β and σ wolves locate the position of the prey and form a enclosure. Under the guidance of α , β and σ wolves, ω wolf wolves surround the prey. According to the literature, the GWO algorithm can easily fall into the local optimum [14]. This paper improves the GWO algorithm. Inspired by particle swarm optimization (PSO), a strategy of adjusting parameters by random distribution is proposed according to the literature [15]: Assuming that the distance parameter a obeys a random number of a certain distribution, the value of a control parameter is adjusted by using the characteristics of random variables to make the algorithm jump out of the local optimum; if the optimal individual is near, the random distribution control parameter produces a relatively small value, which is conducive to accelerating the convergence speed of the algorithm. The specific formula is as follows:

$$a(t) = a_{\text{initial}} - (a_{\text{initial}} - a_{\text{final}}) \times \text{rand}() + \lambda \times \text{randn}() \quad (9)$$

Where, a_{initial} and a_{final} represent the initial value and the end value of the distance control parameter, respectively. $\text{rand}()$ is $[0, 1]$, $\text{randn}()$ is a random number that obeys the normal distribution, λ (variance) is used to measure the deviation degree between the control parameter and the random variable and the mathematical expectation (mean value). It is to control the error of parameter a in the value selection and make the control parameter evolve to the expected control parameter a .

2.3. Parameter optimization of optimized LSTM based on IGWO

Long short-term memory (LSTM) neural network is a variant of recurrent neural network (RNN). It can learn the long-term dependence of input data and alleviate the problems of gradient disappearance and gradient explosion in model training. It has obvious advantages in processing data with nonlinear time series.

As multiple parameters need to be adjusted in the process of using LSTM, the parameter values have a significant impact on the model prediction and classification effect, and the manual search of parameters is time-consuming and laborious, so the goal programming algorithm is needed to assist the parameter optimization. Therefore, the IGWO algorithm in this paper is used to optimize the LSTM. In the neural network, the number of hidden layer neurons, and the learning rate are optimized as follows:

① The number and range of optimization parameters are determined. Set the values of the parameters to be optimized in LSTM, and set the upper and lower limit arrays ub and lb .

The parameters are expressed as follows:

$$LB = [lb_1, lb_2, lb_3, \dots, lb_{v-1}, lb_v] \quad (10)$$

$$UB = [ub_1, ub_2, ub_3, \dots, ub_{v-1}, ub_v] \quad (11)$$

Where v is the number of parameters to be optimized, lb_v is the lower limit of the v th parameter to be optimized, and ub_v is the upper limit of the v th parameter to be optimized.

② The model index accuracy of LSTM is used as the fitness function output, and the fitness function is as follows:

$$Fitness = \frac{\varepsilon}{J} \times 100\% \quad (12)$$

Where ε represents the number of correct classifications.

③ Set algorithm parameters: population size m , maximum iteration parameters t_{\max} , a_{initial} , a_{final} ;

④ Initialize the grey wolf population $\{d_i, i = 1, 2, \dots, Z\}$, and calculate the fitness value of each grey wolf individual using equation (12);

⑤ The fitness value is sorted in non-decreasing order, and the first, second and third individuals of the wolf population are found, which are α , β and δ . In the iterative process, the position of each individual is updated using (7) and (8) in the wolf population;

⑥ Use formula (9) to generate random parameters a , A and C ;

⑦ Judge whether the algorithm reaches the termination condition ($t \leq t_{\max}$). If the conditions are not met, repeat steps ④~⑧. If the termination condition is reached, the iteration is stopped, and the optimal solution is output. The test flow chart is shown in as Fig. 1.

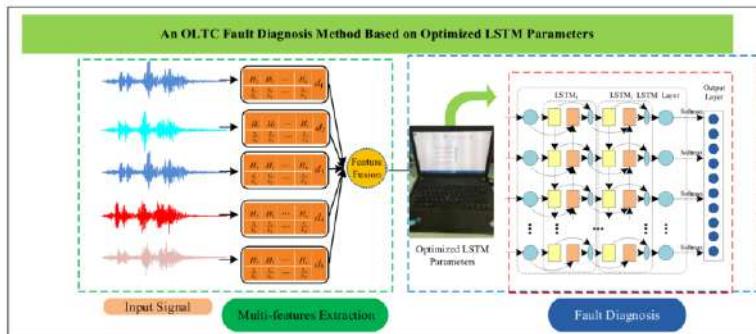


Fig. 1: Test flow chart



Fig. 2: Contact fault setting



Fig. 3: Acceleration sensor location

In this paper, the model CM111-50-63B-10193W Huaming on-load tap changer was adopted, and lc0151 was used in the experiment. The advantage of this collector is that the traditional piezoelectric sensor and charge amplifier are integrated, which simplifies the test system and improves the test accuracy and reliability. A typical mechanical fault of the OLTC switch was simulated. Specifically, it was divided into to normal operation marked as 1, contact burning marked as 2, contact loosening marked as 3, and contact shedding marked as 4. for comparison. A total of 60 groups were collected in each state, and the sampling frequency was 50kHz. The contact fault setting is shown in Fig. 2. The acceleration sensor is shown in Fig. 3.

3. Analysis of Experimental Results

To verify the superiority of this method, the optimal value (the maximum accuracy) and the standard deviation are used as the evaluation indexes. Comparing the support vector machine(SVM)assigned marked as 1. kernel extreme learning machine(KELM) marked as 2,Convolutional Neural Networks(CNN) marked as 3, LSTM marked as 4.GWO-LSTM marked as 5 compared with the proposed method in this paper which marked as 6.The proposed method network simulation is composed of the input layer, one hidden layer, and the output layer, and the Adaam algorithm is used to train the internal parameters of LSTM. The activation function in the hidden layer adopts the tanh function, and the abandonment rate of the network node is 0.2. In order to prevent excessive fitting, the number of iterations is selected as 100. Fig. 4 shows the parameter iteration diagram of IGWO-LSTM. As can be seen from Fig. 4a, the convergence of IGWO is earlier than that of GWO, indicating that the convergence performance of IGWO is stronger than that of GWO, and the convergence value is finally stabilized at 0.0038. Fig. 4b shows the change of hidden neurons with the number of iterations and finally stabilized at 128. LSTM network, SVM network, and KLEM network are constructed by training set samples. CNN network, combined with literature, the learning rate of the LSTM model is 0.01, and the number of hidden neurons is 152. Fig. 4c shows that the learning rate is stable at 0.0012 with the change of iteration times. The kernel parameter g of SVM is 1, and the penalty factor c is 2. The number of network layers of CNN is seven layers (3 convolution layers, three pooling layers, and one full connection layer), the kernel parameter c of KELM is 0.001.

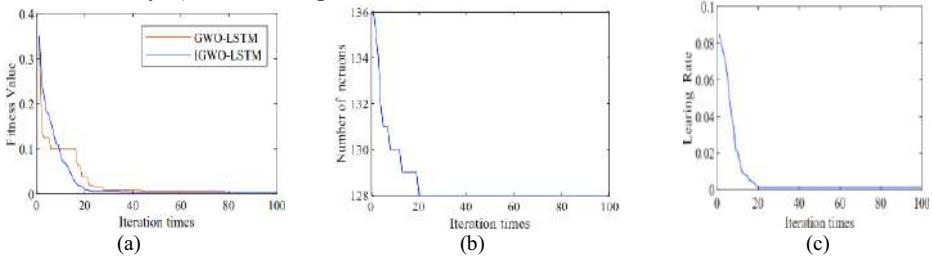


Fig. 4: Changes of c , RLR and T_{\max} in IGWO-LSTM optimal grey wolf individual

It is compared with the other five methods in the case of CMWPE only. It can be seen from Fig. 5 that the average value of IGWO-LSTM is the highest when $\tau = 3; 4; 5; 6$. It can be seen from Fig. 7 that the STD value of the proposed is the smallest of the six methods, which indicates that the accuracy of OLTC fault diagnosis based on LSTM can be improved by optimizing LSTM parameters by IGWO.

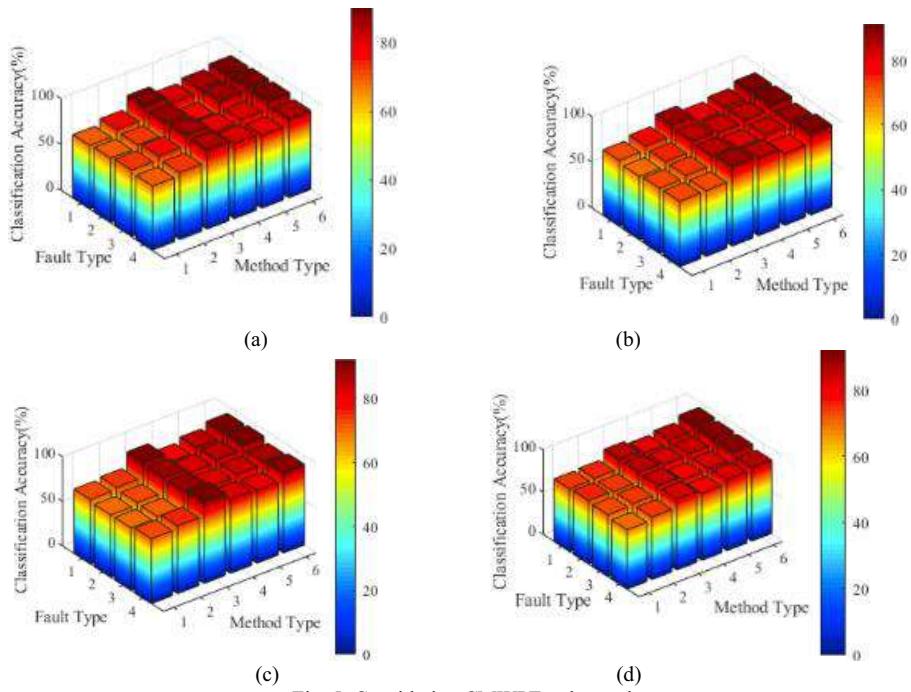


Fig. 5: Considering CMWPE only results

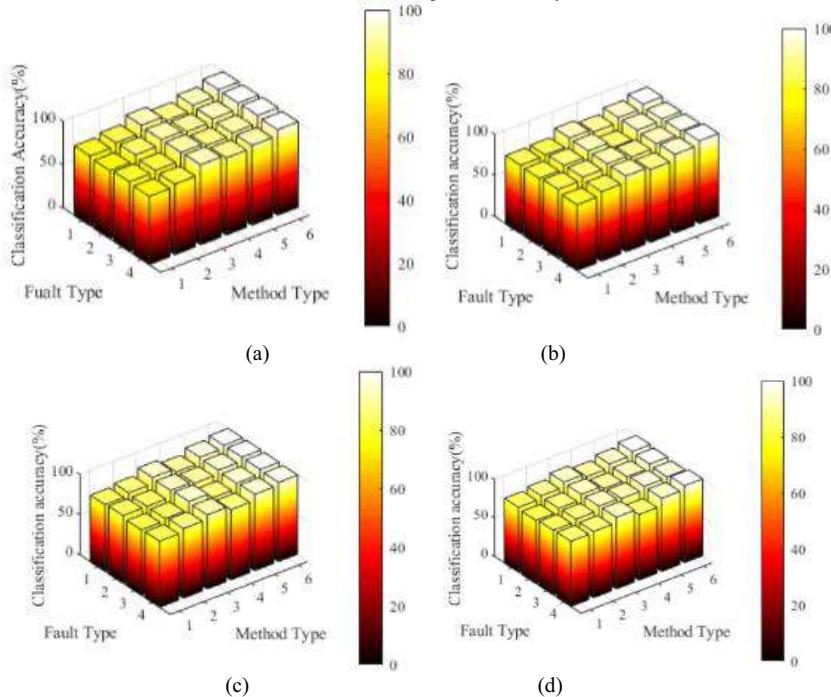


Fig. 6 Feature fusion results

But because based on single scale proposed method there is still a false negative, there can still be a misjudgment in OLTC fault diagnosis. To further improve the fault diagnosis accuracy of OLTC, CMWPE and energy entropy of fusion, the proposed

method in the same place with six kinds of method is compared. As can be seen from Fig. 6, compared with proposed method based on a single feature, when $\tau = 3, 4, 5, 6$, the average value of the proposed method is 8%-10% higher than that of GWO-LSTM and about 20%-30% higher than that of the other 4 methods. As can be seen from Fig. 8, the STD value of proposed method is the smallest, so it can be concluded that the method proposed in this paper has the highest fault diagnosis accuracy and the best classification effect among the six methods. This shows that the fault diagnosis accuracy of OLTC based on IGWO-LSTM can be greatly improved by feature fusion.

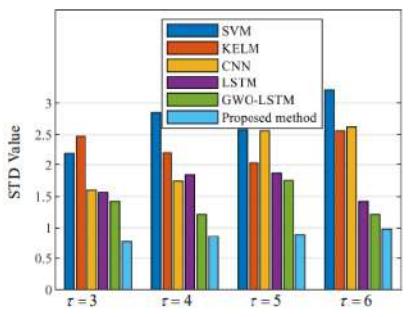


Fig. 7: Considering CMWPE only stds

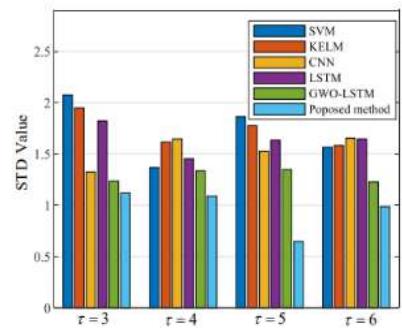


Fig. 8: Feature fusion stds

4. Conclusion

In this paper, an improved grey wolf optimization algorithm is proposed to optimize the relevant parameters of LSTM, and it is applied to OLTC fault diagnosis. The test is carried out by single feature and multi-feature fusion. The specific results are as follows: under the condition of CMWPE only, the parameters of LSTM can be optimized by IGWO, which can improve the fault diagnosis accuracy of OLTC based on LSTM. Under the condition of CMWPE and energy entropy fusion, the effect of the method proposed in this paper is obvious, compared with the accuracy of OLTC fault diagnosis based on single feature IGWO-LSTM. The stability of the method is better, which shows that the method proposed in this paper is relatively good, and further provides a theoretical basis for online OLTC fault monitoring.

References

- [1] P. Kang and D. Birtwhistle, "Condition monitoring of power transformer on-load tap-changers. part 1: Automatic condition diagnostics," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 148, no. 4, pp. 301–306, 2001.
- [2] G. Qian and S. Wang, "Fault diagnosis of on-load tap-changer based on the parameter-adaptive vmd and sa-elm," in *2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, 2020, pp. 1–4.
- [3] Y. Yan, H. Ma, M. Wen, S. Dang, and H. Xu, "Multi-feature fusion based mechanical fault diagnosis for on-load tap changers in smart grid with electric vehicles," *IEEE Sensors Journal*, 2020.
- [4] R. Duan and F. Wang, "Fault diagnosis of on-load tap-changer in converter transformer based on time-frequency vibration analysis," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 6, pp. 3815–3823, 2016.
- [5] P. Kang and D. Birtwhistle, "Characterisation of vibration signals using continuous wavelet transform for condition assessment of on-load tap changers," *Mechanical Systems and Signal Processing*, vol. 17, no. 3, pp. 561–577, 2003.

- [6] Y. Xu, C. Zhou, J. Geng, S. Gao, and P. Wang, "A method for diagnosing mechanical faults of on-load tap changer based on ensemble empirical mode decomposition, volterra model and decision acyclic graph support vector machine," *IEEE Access*, vol. 7, pp. 84 803–84 816, 2019.
- [7] Q. Li, T. Zhao, L. Zhang, and J. Lou, "Mechanical fault diagnostics of onload tap changer within power transformers based on hidden Markov model," *IEEE Transactions on Power Delivery*, vol. 27, no. 2, pp. 596– 601, 2012.
- [8] H. Gu, Y. Wang, B. Li, Y. Chen, and Z. Li, "Mechanical fault diagnosis of transformer on-load tap-changer based on improved variational mode decomposition and support vector machine," pp. 1–4, 2020.
- [9] X. Liang, Y. Wang, and H. Gu, "A mechanical fault diagnosis model of on-load tap changer based on same-source heterogeneous data fusion," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2021.
- [10] G. S. Lee, S. S. Bang, H. A. Mantooth, and Y. J. Shin, "Condition monitoring of 154 kv hts cable systems via temporal sliding LSTM networks," *IEEE Access*, vol. 8, pp. 144 352–144 361, 2020.
- [11] D. Wang, H. Chen, T. Li, J. Wan, and Y. Huang, "A novel quantum grasshopper optimization algorithm for feature selection," *International Journal of Approximate Reasoning*, vol. 127, pp. 33–53, 2020.
- [12] H. A. Cheng, W. B. Tao, B. Cl, and C. C. Tong, "A novel method of composite multiscale weighted permutation entropy and machine learning for fault complex system fault diagnosis - sciencedirect," *Measurement*, vol. 158.
- [13] Y. Fu, H. Xiao, L. H. Lee, and M. Huang, "Stochastic optimization using grey wolf optimization with optimal computing budget allocation," *Applied Soft Computing*, vol. 103, p. 107154, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494621000776>
- [14] M. H. Nadimi-Shahroki, S. Taghian, and S. Mirjalili, "An improved grey wolf optimizer for solving engineering problems," *EXPERT SYSTEMS WITH APPLICATIONS*, vol. 166, MAR 15 2021.
- [15] W. Long, S. Cai, J. Jiao, and et al, "An improved grey wolf optimization algorithm," *ACTA ELECTRONICA SINICA*, vol. 47, no. 1, pp. 169–175, 2019. [Online]. Available: <http://www.cnki.com.cn/Article/CJFDTotal-DZXU201901022.htm>

Automated Contextual Anomaly Detection for Network Interface Bandwidth Utilisation: A Case Study in Network Capacity Management

Esmaeil ZADEH^a, Stephen AMSTUTZ^b, James COLLINS^b, Craig INGHAM^b,
Marian GHEORGHE^a and Savas KONUR^a

^a*Department of Computer Science, University of Bradford, UK*

^b*Xalient Holdings Limited, UK*

Abstract. We present a contextual anomaly detection methodology utilised for the capacity management process of a managed service provider that administers networks for large enterprises. We employ an ensemble of forecasts to identify anomalous network traffic. Stream of observations, upon their arrival, are compared against these baseline forecasts and alerts generated only if the anomalies are sustained. The results confirm that our approach significantly reduces false alerts, triggering rather more accurate and meaningful alerts to a level that could be proactively consumed by a small team. We believe our methodology makes a useful contribution to the applications enabling proactive capacity management.

Keywords. contextual anomaly detection, ensemble of timeseries forecasts, data gathering, machine learning, alerting

1. Introduction

The ability to effectively maintain the performance of a network relies on careful management of the capacity of its constituent paths. Anomalous behaviours can potentially result in losses to the business – in both revenue and long-term reputation [1]; i.e., an unusually overutilised path will increase latency and/or create packet-loss resulting in degradation of the user experience.

Traditionally capacity management would be accomplished by monitoring interface traffic and setting up static threshold alerting (i.e., point anomaly [2]) based on a percentage of the configured bandwidth. While this global outlier detection approach can be used to capture large global anomalies, it is ill-suited for detecting seasonal anomalies [1,3] and can lead to a high frequency of alerts, especially during busy periods making it difficult to identify genuine issues within the noise.

An alternate and more effective approach is contextual anomaly detection [3]. Chandola et al. [2] define the term as where a data point is only anomalous in a specific context; i.e. given time context, an interface traffic might be anomalous in weekends but not in weekdays. It is also suggested that modelling the structure of the data and using it

* Corresponding Author. Email: e.habibzadeh1@brad.ac.uk

for anomaly detection is one of the techniques that is usually applied to time series data with seasonality. Rather than a single forecast model, an ensemble of forecasts where each model can be optimised to address a very specific problem at hand is widely accepted for its improved performance [4,5].

Herein we present a contextual anomaly detection methodology for the capacity management process of a managed service provider that administers networks for large enterprises. Taking into account the trend and seasonality, this approach sets dynamic thresholds based on the variations from the forecasts, and therefore tends to significantly reduce false alerts generated to allow a much smaller team to respond to them. One of the challenges in this case study was accurately generating a small number of meaningful alerts (i.e., eliminating false positives and false negatives), and ensuring these alerts were delivered in a timely manner, ideally within 15 minutes. We discuss the investigations into how data granularity and smoothing functions play a role and can help. We also consider algorithms and parameters used for alert dampening.

The rest of the paper is organised as follows: Section 2 briefly overviews the existing related work; Section 3 describes our methodology; Section 4 presents experimental results; Section 5 summarises the findings; and Section 6 concludes with future work.

2. Related Work

The existing research in anomaly detection is valuable and quite abundant [1,2,6,7,8,9]; nevertheless, the field cannot claim maturity yet, due to the lack of an overall, integrative framework to understand the nature and different manifestations of its focal concept, the anomaly [6]. Hochenbaum et al. [1] argue that detection of anomalies in time series with multimodal distribution, seasonality and/or underlying trend is non-trivial; the authors highlight that a large body of anomaly detection techniques, e.g., the Three-Sigma (i.e., 3σ) rule, can potentially capture global anomalies, but is not applicable for data with seasonality characteristics mentioned above. They propose detection techniques that involve decomposition of time series into trend and seasonality, followed by point anomaly detection applied to the noise component. Chandola et al. [2] conduct an extensive survey on anomaly detection; they classify anomalies into three types: a) global (point), b) contextual (conditional), and c) collective; they also broadly classify techniques on contextual anomaly detection into two main categories:

- reducing the problem to a global anomaly detection problem, and
- modelling the structure in the data and using it to detect contextual anomalies, which is typically applied to time series data. They provide and discuss, among others, auto-regressive models (i.e. ARMA and ARIMA) that have been widely applied to various application domains including statistics, financial markets, security, networking and graphs.

Ensemble forecasting, however, is widely accepted for its improved performance. Bates and Granger [4] suggest the idea of using multiple forecasts rather than choosing a perfect or even significantly superior forecast model as a way to improve forecast performance. They argue that any forecast model nearly always contains some useful independent information; as such, it is not a wise procedure to discard any of them if the objective is to make as good a forecast as possible.

3. Methodology

Figure 1 illustrates our methodology and the data journey, from the point where it is acquired, to the point where it is evaluated against forecasts and generates alerts.

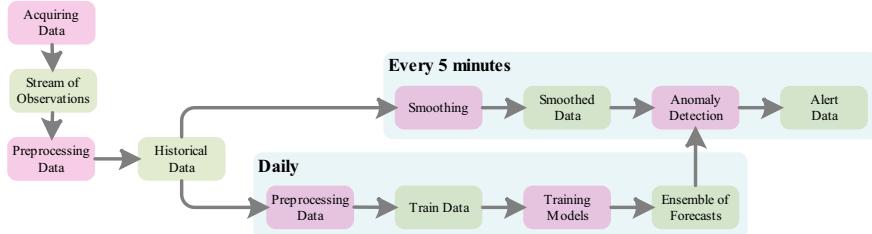


Figure 1. Methodology and data journey block diagram.

Data is sourced from a selection of network devices via a collocated proxy agent. These devices securely report on a regular basis flow records (metadata about each IP conversation traversing a device such as a router, switch, or host) to a proprietary cloud-based agent, storing the records in a cloud data storage. The data here, however, is quite huge for various reasons, but mainly due to its fine granularity (i.e., per second) and that it contains a lot more dimensions and metrics than one would normally require for a certain analytics task at hand. Here, we give focus on interface bandwidth utilisation data.

Per second data is an invaluable asset and sometimes crucial, e.g., for monitoring and analysing min/max network traffic periods for capacity planning; however, it is less suitable for pattern recognition algorithms due to the presence of excessive amount of irregular roughness. As part of data pre-processing, we downsample and take moving average to smooth the data as required. It is worth mentioning, however, that unlike moving average, downsampling tends to reduce granularity, leading to less visibility.

Here, we are particularly interested in detecting contextual anomalies; but we need a mechanism (e.g., a model) to help identify what should be considered normal within certain confidence levels, given the time context. We use past data to learn past behaviours, e.g., seasonality patterns and/or long-term trends, and based on which fit an ensemble of forecast models [10] to predict the future.

It is important to note that granularity of forecast data should be greater than or equal to that of the training data; i.e. given a 5-minute granularity of training data, this suggests 5, 10, 15, ... minutes for the forecast granularity. Otherwise, the fitted models will be less confident on predicting values for sub-intervals.

The last stage of the methodology involves the anomaly detection algorithm, which evaluates the stream of actual observations against their respective forecasts and flags them, if significantly deviated, as:

- low anomaly (i.e., underutilised) if it is below the lower boundary
- high anomaly (i.e., overutilised) if it is above the upper boundary

We take high/low anomalies as disparate classes and treat them independently. The annotated observations are then fed to alerting engine, which in turn generates alerts on breach. Combination of contextual anomaly detection and ensemble forecasting alone tends to significantly reduce false alerts. We also define algorithms to help eliminate duplicate alerts and suppress an alert being triggered multiple times before being cleared.

4. Experimental Results

In this section, we demonstrate how our methodology works and its effectiveness in generating timely and more meaningful alerts. In particular, we briefly look into the rationale for using EWM (Exponentially Weighted Mean) as opposed to SMA (Simple Moving Average) to smooth our data. We then demonstrate how ensemble forecasting can help address two common network issues more reliably. Throughout we use real life ingress/egress interface traffic.

4.1. SMA vs EWM

We use moving average to smooth out irregular variations in interface traffic. All these functions apply to a subset of data, sliced by a moving window, with a side-effect of introducing a lag time, proportional to the window size. That is, smoother outputs can only be achieved at the cost of longer lags and therefore increased resolution times. We look into the two common moving average functions (i.e. SMA and EWM) to find out which can better help achieve reasonably smoothed data while minimising the lag times.

Figure 2 confirms that both functions tend to produce very similar smoothed outputs as long as variations are slow and less significant; even the output of SMA_30m does not look too different, despite using a half-size window.

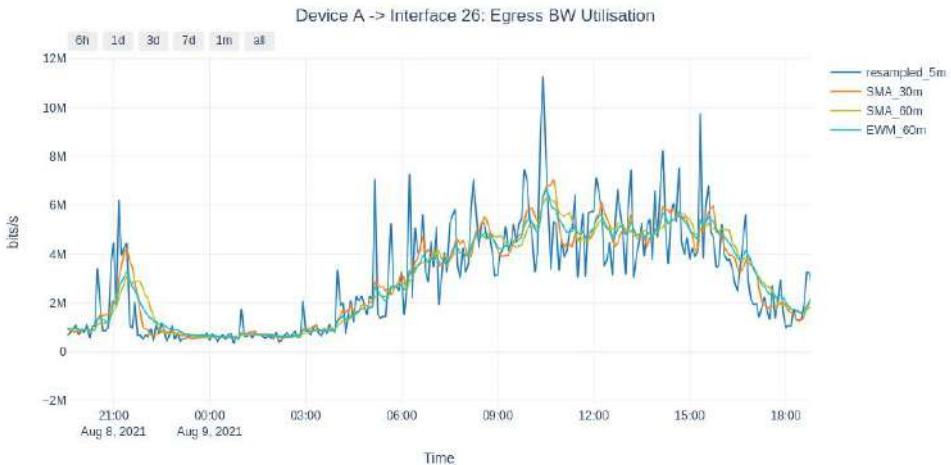


Figure 2. SMA vs EWM: in response to slow.

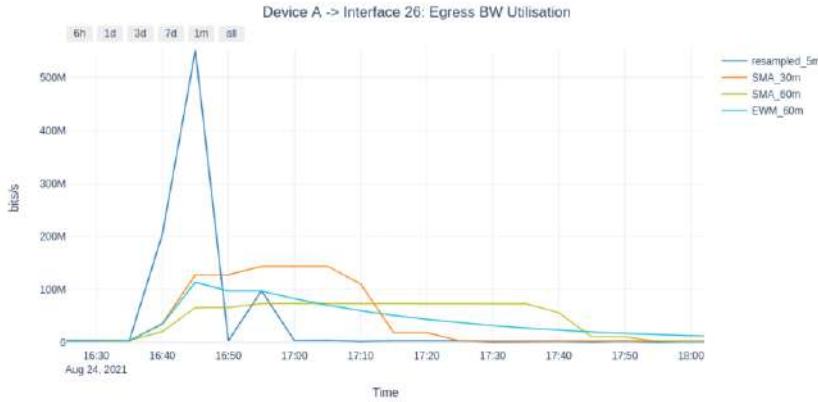


Figure 3. SMA vs EWM: in response to rapid variations.

However, in the presence of rapid changes that are also significant, SMA and EWM tend to behave quite differently. As illustrated in Figure 3, in response to a sharp spike, SMA produces much longer lasting pulses, albeit lower in amplitudes. We also observe that SMA_60m is slower (i.e., longer lag) and left behind the most. EWM, on the other hand, rises and falls so quickly; in fact, despite having double window size, it tends to perform better than SMA_30m, producing smoothed outputs with less distortion. Nevertheless, one may argue that EWM tends to introduce far longer tails; this, however, is very unlikely to have serious implications since the tail dies exponentially.

It is crucial to realise that the distortions due to these smoothing functions can, sometimes, manifest themselves as false persistency, leading to false positive alerts, whereas the original signals with narrow spikes would probably have no chance to trigger any alerts no matter how high in amplitude.

4.2. Ensemble Forecasting

Figure 4 shows actual traffic versus the forecast data, achieved using a single forecast model. Apart from daily and weekly seasonalities, it is also apparent that a daily scheduled job is performed at a specific time (~3-4am).

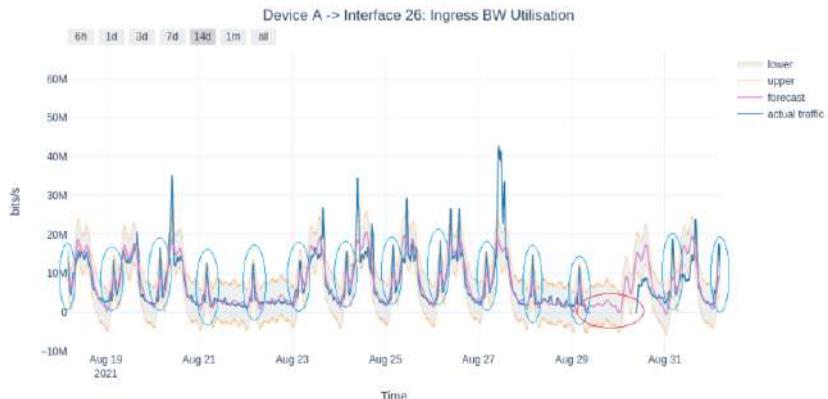


Figure 4. Actual traffic vs forecasts.

At a glance, the model appears to be a reasonably good fit, following major seasonality patterns quite well. However, we realise that the model is partially impaired and ineffective, with two major drawbacks as follows:

- It performs poorly and is less confident around daily scheduled jobs, suggesting way too wide uncertainty intervals to allow for detection of missed jobs! (Figure 4, marked by blue ellipses).
- Also interface failures, e.g. disconnections, or when the interface traffic is zero, will go unidentified since uncertainty intervals include zero values; i.e. no traffic at all are not considered unusual! (Figure 4, marked by red ellipse).

We employ an ensembleforecasts to address these issues and improve overall performance as follows.

4.2.1. Alert on missed scheduled jobs

As mentioned above, the uncertainty intervals, in particular around daily scheduled jobs, are way too wide. It is therefore highly likely that a missed job will go unidentified since there is not enough margin between model's lower bound and the level of dropped traffic due to the absence of a job.

With this problem in mind, we optimize the model such that it is more confident and follows the daily scheduled jobs more accurately. Figure 5 illustrates the new model, which confirms improved margins between presence and absence of a scheduled job; i.e. in the event of a missed job, the dropped traffic will be well below the lower bound and for long enough to allow for timely detection.

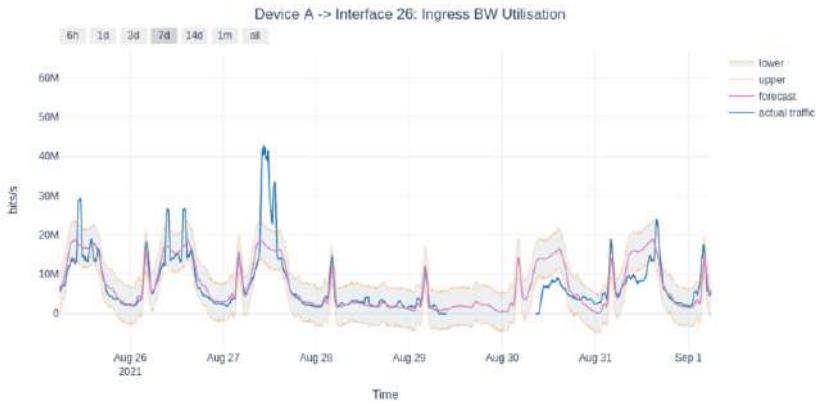


Figure 5. Improved forecast model for daily scheduled jobs

4.2.2. Alert on interface failure

It is not uncommon for an interface traffic to drop to very low levels during evenings and weekends, hence zero is within the expected range. However, due to the type and nature of interfaces used in this study, no traffic at all for quite long periods¹ would be abnormal and timely detection of such cases will be desirable. To achieve this, we apply a log transform prior to fitting the model, which helps to dramatically increase the margin and distinguish zero traffic from the rest quite easily (see Figure 6).

¹ Recall that we aggregate and smooth the traffic data over 60-minute moving windows.

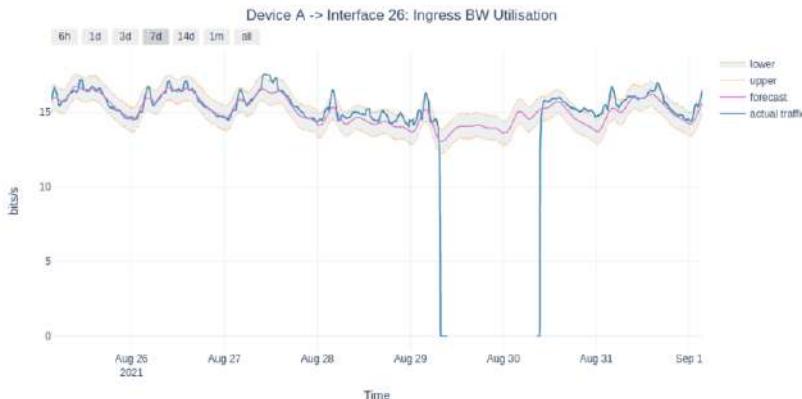


Figure 6. Improved forecast model for interface failure and zero traffic

4.3. Alert Dampening

To reduce false positives, a distinction is drawn between anomalies and alerts. An alert will only be generated when there is a sustained anomaly which is determined using the following parameters:

- Persistent anomalies: where we see n consecutive anomalies.
- Impersistent anomalies: where we see n anomalies across a specific window size (number of samples).

In addition, once an alert has been generated, we suppress further alerts until there has been a period of stability. This period is determined using the following parameter:

- Alert cleared: where we see n consecutive non-anomalous samples.

5. Discussions

Careful selection of the moving average function enabled the smoothed data to better represent the original data resulting in greater accuracy in both the detection and persistency of anomalies, without reducing detection time:

- improved detection: EWM_60m was more likely to reflect an anomaly than SMA_60m due to its higher amplitude.
- quicker detection: EWM_60m raised as quickly as SMA_30m and much quicker than SMA_60m resulting in quicker detection (i.e. shorter lag).
- reduction of false persistency: EWM_60m fell quicker than both SMA_60m and SMA_30m resulting in reduced persistency. This would in turn create less consecutive anomalies therefore reducing the likelihood of an alert being raised.

Using an ensemble of forecasts (versus a single forecast model) improved the accuracy in the detection of anomalies by:

- reducing the uncertainty interval where important anomalies could otherwise be missed (under-utilisation), or

- increasing the uncertainty interval where false-positives are more likely due to bursty traffic (over-utilisation).
- in the case of under-utilisation we found an improvement in detecting the absence of small but regular traffic (scheduled jobs) in some cases from a 10% likelihood of detection to a 100% likelihood of detection.

And finally, utilising alert dampening parameters reduced the quantity of anomalies being escalated to alerts, or duplicate alerts being raised.

6. Conclusions

The exhibition of heavy seasonality in interface bandwidth utilisation is apparent; as such, global anomaly detection techniques are not effective and can lead to too many false alerts. We have, instead, taken a contextual anomaly detection approach together with ensemble forecasts.

Stream interface traffic, upon their arrival, are evaluated against forecast values and flagged anomalous if they are significantly deviating from predicted values. Anomalies get marked as alerts if they satisfy sustainability rules. We also discuss algorithms used to prevent duplicate alerts and suppress any alerts being raised multiple times before they are declared cleared.

The results show that the methods presented can meaningfully improve the alert generation process both in quantity and quality to a level that could be proactively consumed by a small team. We believe our methodology makes a useful contribution to the applications enabling proactive capacity management.

The next move is to consider multi-variate (i.e., latency, loss, utilisation) timeseries where channels are correlated and can inform/share useful knowledge about each other.

Acknowledgements

The authors acknowledge the Innovate UK support [grant number KTP012139].

References

- [1] Hochenbaum J, Vallis O and Kejariwal A. Automatic anomaly detection in the cloud via statistical learning. *arXiv*. 2017; abs/1704.07706.
- [2] Chandola V, Banerjee A and Kumar V. Anomaly detection: A survey. *ACM computing surveys (CSUR)* 2009; 41(3): 1-58.
- [3] Koosha G. and Osmar Z R. Time series contextual anomaly detection for detecting market manipulation in stock market. *IEEE International Conference on Data Science and Advanced Analytics (DSAA)* 2016; 1-10.
- [4] Fox A J. Outliers in time series. *Journal of the Royal Statistical Society: Series B (Methodological)* 1972;
- [5] Bates J M and Granger C W J. *The Combination of Forecasts*. OR 1969; 4(20): 451-468.
- [6] Foorthuis R. On the nature and types of anomalies: a review of deviations in data. *International Journal of Data Science and Analytics* 2021;12(4): 297-331.
- [7] Nielsen A. Practical time series analysis: Prediction with statistics and machine learning. O'Reilly Media 2019.
- [8] Song X, Wu M, Jermaine C and Ranka S. Conditional anomaly detection. *IEEE Transactions on knowledge and Data Engineering* 2007; 19(5): 631-645.
- [9] Hodge V J and Austin J. A survey of outlier detection methodologies. *Artificial intelligence review* 2004; 22(2): 85-126.
- [10] Taylor S J and Letham B. Forecasting at scale. *The American Statistician* 2018; 72(1): 37-45.

An Algorithm of Vocabulary Enhanced Intelligent Question Answering Based on FLAT¹

Jing Sheng LEI^a, Shi Chao YE^a, Sheng Ying YANG^{a,2}, Wei SONG^a and Guan Mian LIANG^b

^a*Department of School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou, PR China*

^b*Cancer Hospital of the University of Chinese Academy of Science, Hangzhou, PR China*

Abstract. The main purpose of the intelligent question answering system based on the knowledge graph is to accurately match the natural language question and the triple information in the knowledge graph. Among them, the entity recognition part is one of the key points. The wrong entity recognition result will cause the error to be propagated, resulting in the ultimate failure to get the correct answer. In recent years, the lexical enhancement structure of word nodes combined with word nodes has been proved to be an effective method for Chinese named entity recognition. In order to solve the above problems, this paper proposes a vocabulary-enhanced entity recognition algorithm (KGFLAT) based on FLAT for intelligent question answering system. This method uses a new dictionary that combines the entity information of the knowledge graph, and only uses layer normalization for the removal of residual connection for the shallower network model. The system uses data provided by the NLPCC 2018 Task7 KBQA task for evaluation. The experimental results show that this method can effectively solve the entity recognition task in the intelligent question answering system and achieve the improvement of the FLAT model, and the average F1 value is 94.72

Keywords. Intelligent question and answer, Knowledge Graph, Entity recognition, Vocabulary enhancement, Residual connection

1. Introduction

With the continuous growth of data on the Internet, a large-scale knowledge base was born, including Freebase [1] and NLPCC KBQA. These promote the development of knowledge base question answering system (Question Answering over Knowledge Base, KBQA), which can help people naturally retrieve answers from large-scale semi-structured knowledge base. Entity recognition and attribute mapping are the main components of a ques-

¹This work is supported by Natural Science Foundation of China (No. 61672337, 61972357), Zhejiang Key R&D Program (No. 2019C03135), Basic Public Welfare Research Project of Zhejiang (LGF19F020003).

²Corresponding author: Shengying Yang, Department of School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou, 310023, PR China; E-mail: syyang@zust.edu.cn.

tion answering system based on large-scale knowledge base. Because there are many different ways to ask the same question, the deep understanding of the input question is the key to realize an intelligent question answering system based on a knowledge base. The intelligent question answering system of knowledge graph based on information extraction method can effectively locate the key entities in the question and match the correct answer from the knowledge graph.

In May 2012, Google proposed the concept of Knowledge Graph [2], which is a new type of data representation. The knowledge graph model can understand the relationship between objective things by storing entities, relationships between entities and attributes. The basic unit is a triplet, such as "entity-attribute-attribute value" or "entity-relation-entity". Therefore, the knowledge graph is a high-quality corpus, which can efficiently organize and manage data information.

The knowledge graph intelligent question answering system based on the information extraction method first needs to obtain the entity information in the question sentence through the named entity recognition technology. If the entity-related information is queried in the knowledge graph, we can obtain the subgraph of the knowledge graph centered on the entity node. The candidate answer is the node or edge in the subgraph of the knowledge graph. Then, the question sentence is transformed into a question feature vector through information extraction technology. Finally, the problem feature vector and candidate answers are filtered through the classifier to find the correct answer. HUANG [3] proposed a named entity model based on LSTM and CRF, which are widely used in the entity recognition part of intelligent question answering systems based on knowledge graphs. However, this method can only be based on characters or vocabulary in Chinese, and cannot solve the Chinese entity boundary problem well.

In this article, our contributions are as follows:

1. The KBQA framework based on information extraction is proposed, and the Flat-Lattice model is applied to the entity recognition part of the knowledge graph intelligent question answering system.

2. Designed a vocabulary enhancement method based on FLAT model and combined with knowledge graph entities. This method can effectively solve the problem that the intelligent question answering entity recognizes the boundary of some Chinese entities. While based on characters, the vocabulary information of entities in the knowledge graph is introduced, which enriches the high-level features of the text information more specifically, thereby improving the recognition of question entities in the Chinese intelligent question answering system based on the knowledge graph.

3. By designing the entity tagging algorithm of question-and-answer pairs based on remote supervision, the data set of NLPCC2018-KBQA is labeled, so that it can be used for named entity evaluation.

4. A large number of experiments were conducted on the NLPCC 2018 Task7 KBQA task, and our model achieved good accuracy, recall rate and f1 score.

2. Related Work

The method is based on information extraction first needs to obtain the entity information in the question through the named entity recognition technology. If the entity-related information is queried in the knowledge graph, we can obtain the subgraph of the knowl-

edge graph centered on the entity node. The candidate answer is the node or edge in the subgraph of the knowledge graph. Then, the question sentence is transformed into a question feature vector through information extraction technology. Finally, the problem feature vector and candidate answers is filtered through the classifier to find the correct answer.

The deep neural network model has become a research trend in named entity recognition tasks because it does not require manual feature engineering and expert knowledge in related fields [4]. Liu [5] proposed a sequence tagging framework LM-LSTM-CRF. In addition to the word-level knowledge contained in pre-trained word embeddings, this framework also integrates a character-aware neural language model, so the language model can be effectively used Extract character-level knowledge from a self-contained sequence information. Yan [6] proposed a NER architecture-TENER that uses transformer encoders to simulate character-level features and word-level features. It greatly improves performance by introducing orientation perception, distance perception and scale-free attention. Gui [7] proposed a Chinese NER model LRCNN based on convolutional neural network. This method combines vocabulary and reflection mechanism to make parallel modeling of all characters and potential words in a sentence.

BERT [8] is a language representation model based on a pre-trained deep bidirectional Transformer [9] structure proposed by Google in October 2018. Use the context information of the word to do feature extraction, dynamically adjust the word vector according to the different context information, and solve the problem of word2vec[10]. To use the pre-trained BERT, you only need to load the pre-trained model as the word embedding layer of your current task, and then build the subsequent model structure for the specific task.

According to the language characteristics of English, identifying the categories of entities is the main task of English named entity recognition task. Compared with English, Chinese has no linguistic characteristics that can determine the boundaries of entities, such as the spaces between words and words and the obvious capital letters of entity words. Therefore, determining the entity boundary is one of the research focuses of Chinese named entity task. Character-based entity recognition and word-based entity recognition after word segmentation are two common ways of Chinese named entity recognition at present. However, the semantic meaning of a sentence in Chinese will change due to different results of word segmentation, resulting in errors. Therefore, the character-based method is generally better than the vocabulary-based method in Chinese named entity recognition. Recent studies have shown that lexical boundaries are beneficial to the determination of entity boundaries in Chinese named entity tasks. Therefore, on the basis of the character-based method, the use of vocabulary information, through the way of vocabulary enhancement, so as to improve the performance of the Chinese named entity recognition system. Zhang [11] applied lexical enhancement method to Chinese named entity recognition for the first time, and proposed a lattice structure LSTM model-Lattice LSTM for Chinese NER. The model encodes the input character sequence and all the potential words matched with the dictionary, and fuses the vocabulary information into the original cyclic neural network model. Gui [12] proposed a global semantic graph neural network (LGN), based on dictionaries, in which the graph structure is composed of characters as nodes and matched lexical information as edges. The network uses dictionary knowledge to connect the characters to obtain local combinations, while global relay nodes can capture global sentence semantics and long-range dependencies, so that

the model of rnn is easily affected by word ambiguity. Liu [13] proposed the WC-LSTM model, which adds word information to the beginning or end of a word, which reduces the impact of word segmentation errors on the acquisition of word boundary information. At the same time, through four different strategies, the lexical information are encoded into a fixed size representation, so that effective parallelization training can be carried out. Ma [14] introduce lexical information in the embedding layer through the Soft-lexicon method. For different sequence frameworks, this method avoids the complex design of the sequence modeling architecture, and can introduce dictionary information by fine-tuning the character representation layer. Li [15] proposed FLAT, which transforms the grid structure into a plane structure composed of spans. Each span corresponds to a character or latent word and its position in the original lattice. At the same time, according to the fully connected self-attention mechanism in Transformer, it can well capture the long-distance dependence and use it to introduce the position vector to maintain the position information. Thus it can be seen that it is of great research significance to apply vocabulary enhancement technology to the entity recognition steps in the intelligent question answering system of Chinese knowledge graph.

3. Methodology

The system is mainly composed of two parts: named entity recognition and attribute mapping. Among them, the purpose of the entity recognition step is to find the entity information in the question, and the purpose of the attribute mapping step is to determine the related attributes of the question after finding the question entity. This paper will introduce the model of named entity recognition step in detail.

3.1. *Named Entity Recognition*

The entity recognition part is mainly to find the entity information in the question, which is used for the subsequent steps of attribute mapping in the intelligent question and answer. The model framework of this part of this article is shown in Figure 1. First, combine the entity vocabulary information in the knowledge graph to construct a lattice structure composed of characters and vocabulary, flatten it, and then encode the relative position of awakening according to the span of the lattice. Finally, the coding information combined with the model is used to identify the correct entity in the question.

3.2. *Dictionary Combining Entity Information Of Knowledge Graph*

This model does named entity recognition in question sentences. Except for byte points, the added word nodes is added according to the dictionary. The FLAT model uses the word segmentation information of the giga dictionary. As long as the word segmentation information appears in the question sentence, it is added as a new word node. For this task, there is more redundant information. The appearance of the entity in the question is related to but not exactly the same as the entity vocabulary in the mention2id file, such as participles, abbreviations or abbreviations. The new dictionary mainly focuses on the word segmentation, acronyms or abbreviations of the entity information in the knowledge graph through the mention2id file. The new vocabulary is more accurate, and the entity is more targeted on the vocabulary. At the same time, simplifying the vocabulary helps to improve the efficiency of the model.

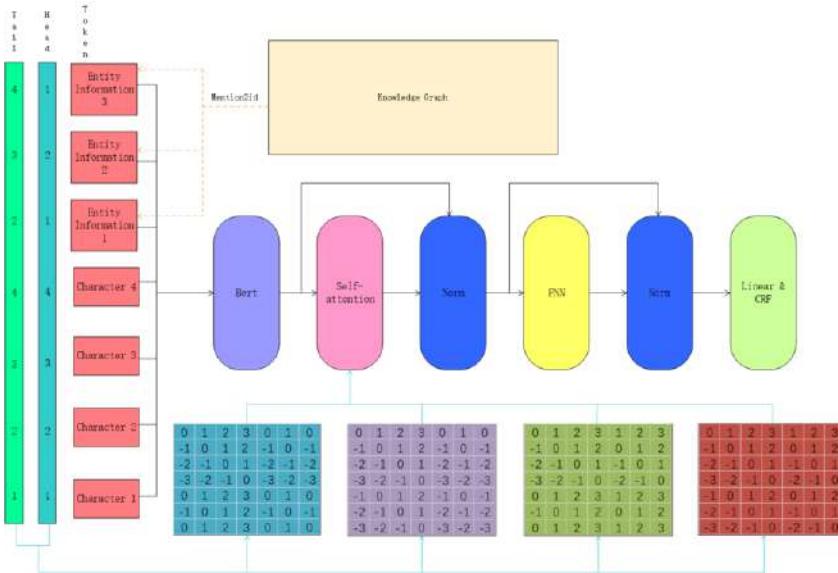


Figure 1. an algorithm of vocabulary-enhanced intelligent question-answer entity recognition based on FLAT

3.3. Converting Lattices To FLAT Structures

The sentence is matched with the new dictionary in the form of characters, and the obtained information is the grid. Convert it to a flat structure as shown in Figure 1. The plane grid is displayed in the form of a span. A span consists of token, head and tail. In this form, a character or vocabulary information can be marked. In order to express the position information of each token, the model uses head and tail to indicate the position index of the first and last characters of each token in the input sequence. If the token information is a single character, its head and tail are the same. Therefore, byte points are constructed by the same symbols of head and tail, and word nodes are marked by different head and tail by constructing jumping paths. And the change is recoverable, and the original structure of the plane grid can be restored.

3.4. Relative Position Coding Of Span

By encoding the relative position of the spans, we encode the interactions between spans of different lengths in a planar lattice structure. According to the span of head and tail, the two spans Y_a and Y_b of a planar lattice have three relationships: separation, intersection and inclusion. We use the continuous transformation calculation of head and tail information, and use the form of a dense vector to model their relationship. This method not only represents the relationship between the two markers, but also contains information such as the distance between words. The four relative distances of Y_a and Y_b are calculated by the following formula, they can represent the relationship between the two, where $h[a]$ and $t[a]$ represent the head and tail information about the span Y_a , and $l_{ab}^{(hh)}$ represents The distance between the head of Y_a and the tail of Y_b , the rest of $l_{ab}^{(ht)}$, $l_{ab}^{(th)}$ and $l_{ab}^{(tt)}$ can be deduced by analogy.

$$l_{ab}^{(hh)} = h[a] - h[b] \quad (1)$$

$$l_{ab}^{(ht)} = h[a] - t[b] \quad (2)$$

$$l_{ab}^{(th)} = t[a] - h[b] \quad (3)$$

$$l_{ab}^{(tt)} = t[a] - t[b] \quad (4)$$

As shown in the following formula, referring to the method of Vaswani et al. [9], and through the nonlinear transformation of the above four distance formulas, the final relative position coding of the span is obtained.

$$P_d^{(2k)} = \sin(d/10000^{2k/d_{model}}) \quad (5)$$

$$P_d^{(2k+1)} = \cos(d/10000^{2k/d_{model}}) \quad (6)$$

$$M_{ab} = \text{ReLU}(W_m(P_{l_{ab}^{(hh)}} \oplus P_{l_{ab}^{(ht)}} \oplus P_{l_{ab}^{(th)}} \oplus P_{l_{ab}^{(tt)}})) \quad (7)$$

Where d is $l_{ab}^{(hh)}$, $l_{ab}^{(ht)}$, $l_{ab}^{(th)}$, $l_{ab}^{(tt)}$, W_m is a learnable parameter, and k represents the position-coded dimension index. Then, according to the method of Dai et al. [16], a variant of self-attention is used to encode the relative span position, and the calculation formula is as follows:

$$A_{a,b}^* = W_q^T E_{Y_a}^T E_{Y_b} W_{k,E} + W_q^T E_{Y_a}^T M_{ab} W_{k,M} + u^T E_{Y_b} W_{k,E} + v^T M_{ab} W_{k,M} \quad (8)$$

Among them, $W_q, W_{k,M}, W_{k,E} \in \mathbb{R}^{l_{model} \times l_{head}}$ and $u, v \in \mathbb{R}^{l_{head}}$ are the learnable parameters.

3.5. Remove Residual Connection

The purpose of residual connection is to solve the degradation problem that when the layers of the training network are getting deeper and deeper, the training effect is not as good as the relatively shallow network. By adding shortcut connections, the residual network makes the training model easier to optimize. The residual block consists of a multi-layer network containing shortcut connections. A residual block is represented by the following formula, and the residual network is composed of residual blocks.

$$X_{l+1} = X_l + F(x_l, W_l) \quad (9)$$

The additional X term can solve the problem of gradient disappearance in the process of back propagation. Because when the network calculates the partial derivative of X , there is an extra constant 1.

Transform uses the structure of a six-layer encoder-decoder, which consists of a self-attention mechanism and a feed-forward network (FFN), with residual connections and layer normalization added after each sublayer. The calculation formula is as following:

$$Att(A, V) = \text{softmax}(A)V \quad (10)$$

$$A_{ab} = \frac{Q_i K_j^T}{\sqrt{d_{head}}} \quad (11)$$

$$[Q, K, V] = E_x [W_q, W_k, W_v] \quad (12)$$

Among them, $W_q, W_k, W_v \in \mathbb{R}^{l_{model} \times l_{head}}$ are the learnable parameters, and the model in this paper replaces A_{ab} with $A_{a,b}^*$ in formula 8. FLAT inherited this structure directly, but he actually only used one layer of encoder. We think this is a shallower network compared to transform, so the residual connection is removed and only layer normalization is used. Finally, the character representation is put into the output layer, and then the entity recognition task is completed through the Conditional Random Field (CRF) [17].

4. Experiments

4.1. Experimental Data And Preprocessing

This experiment uses data provided by the NLPCC 2018 Task7 KBQA task. The task data contains 24,479 one-to-one corresponding question and answer data pairs and the required knowledge graph. The knowledge graph provided by this data contains 6,502,738 entities, 587,875 attributes, and 43,063,769 triples. It is currently one of the large-scale open domain knowledge graphs in Chinese. Each row in the knowledge graph stores a triple (entity, attribute, attribute value/entity).

Because the NLPCC2018-KBQA dataset is not a dataset in the task area of named entity recognition, the 24479 question-and-answer pairs of data it provides do not clearly indicate the location of the entity. Therefore, this experiment is based on the idea of remote supervision. According to the triples of the provided knowledge graph, remote supervision annotates the entity sequence of the question in the question and answer pair. In this way, a data set with entity sequence annotations that can be used in the entity recognition step is obtained.

The key idea of the question-and-answer entity labeling algorithm based on remote supervision is to deduce the corresponding triples of each question-and-answer pair from the triplet knowledge base. For the knowledge graph triplet set, when there is a triple (e, r, s), s is included in the answer of a certain question and answer pair, and e is included in the corresponding question, then the question is considered question. The sentence entity matches successfully. Find from the question whether the entity in the corresponding triple is included. If the entity is not included, use the mention2id file to replace the entity name with its corresponding synonym to continue the search. After finding the entity, it is written in the NER training corpus format, and each word is given a corresponding label such as B-ENT, as shown in Figure 2.

The collection of 24479 questions and answer data pairs, mention2id files and 43063796 knowledge graph triples provided by the NLPCC 2018 Task7 KBQA task are preprocessed through 2.1.2 steps. First, a mapping dictionary AnswerDict with a length of 21269 is formed, that is, the size of the dictionary key when the question and answer



Figure 2. training expectation format after labeling

Table 1. KGFLAT model parameter list

batch	8
learning rate	0.0005
optimizer	SGD
-momentum	0.9
head	8
head_dim	16
output dropout	0.3
warmup	0.1

pair is built, which is the number of answers that are not repeated in the data. Find from the question whether it contains the entity in the corresponding triple. If it does not contain the entity, replace the entity name with its corresponding synonym through the mention2id file to continue the search. Even then, there are still entities that cannot be found in the corresponding question. So, discard this part of the data. Finally, 15640 valid questions and answers are labeled successfully by the question-and-answer entity labeling algorithm based on remote supervision.

Split the dataset 80 % as the training set, 10 % as the validation set, and 10 % as the test set. The number of training sets, verification sets and test sets are 12512, 1564 and 1564 respectively. At the same time, in order to obtain a reliable and stable model, the ten-fold cross-validation (10-fold cross-validation) method [18] is used for training. Models that work well on the validation set are evaluated on the test set. According to the common evaluation criteria of intelligent question and answer, the accuracy, recall rate and average F1 value were used to evaluate the experimental results.

4.2. Experimental Settings And Parameter Optimization

The experiment was carried out in the Ubuntu 18.04 system, where the CPU was an Intel i5-7500 processor with 4 cores, the GPU was NVIDIA TITAN XP-12GB, and the memory was 16GB. The model training process is carried out on the GPU, and the experimental code adopts the Pytorch deep learning framework.

The dimension of the input word vector in the model is 50, the Batch Size is set to 8, the learning rate is set to 0.0005, and the parameter optimizer uses the SGD optimizer. More model parameters are shown in Table 1:

4.3. Experimental Results And Analysis

In order to verify the advancement of this model, the classical models BI-LSTM[3], Lattice LSTM [11], TENER [6], LRCNN [7], FLAT [15] and BERT-LSTM are tested on this dataset. The experimental results are shown in table 2.

Comparing the classic model and the current Chinese named entity recognition model with better performance, as shown in Table 2, the model proposed in this paper has

Table 2. Comparison of KGFLAT with the classic model on the NLPCC 2018 Task7 KBQA task data set

Model	P	R	F1
LSTM+CRF	88.00	86.88	87.44
Lattice LSTM	89.76	89.82	89.79
TENER	90.50	90.27	90.38
LRCNN	91.32	90.91	91.11
FLAT	91.99	91.99	91.99
BERT+LSTM+CRF	93.22	92.64	92.93
BERT +FLAT	93.92	93.70	93.81
BERT+FLATND	94.50	94.56	94.52
BERT+FLATNR	94.29	94.47	94.38
BERT+KGFLAT	94.75	94.69	94.72

achieved better results. As shown in Table 2, compared with the classic models based on recurrent neural networks and combined with Lattice lattice structure, such as Bi-LSTM, LatticeLSTM and BERT-LSTM, the average F1 value of the model in this paper increased by 7.28, 4.93, and 1.79 percentage points, respectively; compared with the model TENER, LRCNN and FLAT increased by 4.34, 3.61, and 2.73 percentage points respectively. The performance of this model in this data is significantly improved, which is better than the performance of the above model in this data.

The FLAT model uses a character vector and word vector which are trained by Giga Dictionary. In this paper, the pre-training model Bert is used to train the embedding. On this basis, the performance of the model FLATND after modifying the dictionary combined with the knowledge graph entity information and the model FLATNR after de-residual connection on the data is the same as the performance of the original FLAT model on the data Promote. KGFLAT not only uses the new dictionary but also removes the residual connections. Compared with the original FLAT model, the average F1 value of this model is improved by 0.91 %. The model in this paper verifies the effectiveness of the new dictionary and de-residual connection to improve the ability of the FLAT model, and the average F1 value reaches 94.72.

5. Conclusion

Entity recognition is a difficulty in building an intelligent question answering system based on knowledge graph. The model proposed in this paper solves this problem well. Traditional entity recognition algorithms is mainly based solely on character information or word information. If it is a character -based approach, words are often rarely considered, so a large amount of information contained in the vocabulary can not be used. Lexical boundaries usually play an important role in entity boundaries, and the entity information in the knowledge graph can make entity recognition algorithms more targeted. Based on the knowledge corpus of NLPCC2018-KBQA, this paper proposes and applies the question-answer entity labeling algorithm based on remote supervision and the word-enhanced entity recognition algorithm based on FLAT. The above data and algorithms is used to build an intelligent question answering system. In the entity recognition step,

the model proposed in this paper uses vocabulary enhancement technology, based on the FLAT model and through the creation of a new dictionary and de-residual connection to achieve the promotion of the FLAT model.

References

- [1] Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. Freebase: a collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1247–1250, 2008.
- [2] Thomas Steiner, Ruben Verborgh, Raphaël Troncy, Joaquim Gabarro, and Rik Van de Walle. Adding realtime coverage to the google knowledge graph. In *11th International Semantic Web Conference (ISWC 2012)*. Citeseer, 2012.
- [3] Zhiheng Huang, Wei Xu, and Kai Yu. Bidirectional lstm-crf models for sequence tagging. *arXiv preprint arXiv:1508.01991*, 2015.
- [4] Joey Tianyi Zhou, Hao Zhang, Di Jin, Xi Peng, Yang Xiao, and Zhiguo Cao. Roseq: Robust sequence labeling. *IEEE transactions on neural networks and learning systems*, 31(7):2304–2314, 2019.
- [5] Liyuan Liu, Jingbo Shang, Xiang Ren, Frank Xu, Huan Gui, Jian Peng, and Jiawei Han. Empower sequence labeling with task-aware neural language model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.
- [6] Hang Yan, Bocao Deng, Xiaonan Li, and Xipeng Qiu. Tener: Adapting transformer encoder for named entity recognition. *arXiv preprint arXiv:1911.04474*, 2019.
- [7] Tao Gui, Ruotian Ma, Qi Zhang, Lujun Zhao, Yu-Gang Jiang, and Xuanjing Huang. Cnn-based chinese ner with lexicon rethinking. In *IJCAI*, pages 4982–4988, 2019.
- [8] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.
- [10] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- [11] Yue Zhang and Jie Yang. Chinese ner using lattice lstm. *arXiv preprint arXiv:1805.02023*, 2018.
- [12] Tao Gui, Yicheng Zou, Qi Zhang, Minlong Peng, Jinlan Fu, Zhongyu Wei, and Xuan-Jing Huang. A lexicon-based graph neural network for chinese ner. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 1039–1049, 2019.
- [13] Wei Liu, Tongge Xu, Qinghua Xu, Jiayu Song, and Yueran Zu. An encoding strategy based word-character lstm for chinese ner. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2379–2389, 2019.
- [14] Ruotian Ma, Minlong Peng, Qi Zhang, and Xuanjing Huang. Simplify the usage of lexicon in chinese ner. *arXiv preprint arXiv:1908.05969*, 2019.
- [15] Xiaonan Li, Hang Yan, Xipeng Qiu, and Xuanjing Huang. Flat: Chinese ner using flat-lattice transformer. *arXiv preprint arXiv:2004.11795*, 2020.
- [16] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V Le, and Ruslan Salakhutdinov. Transformer-xl: Attentive language models beyond a fixed-length context. *arXiv preprint arXiv:1901.02860*, 2019.
- [17] John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. 2001.
- [18] Peng Chen, HK Feng, CC Li, GJ Yang, JS Yang, WP Yang, and SB Liu. Estimation of chlorophyll content in potato using fusion of texture and spectral features derived from uav multispectral image. *Transactions of the Chinese Society of Agricultural Engineering*, 35(11):63–74, 2019.

Detection Method of Computer Room Personnel Based on Improved DERT

Dan SU^a, Qiong-lan NA^{a1}, Hui-min HE^a and Yi-xi YANG^b

^a*State Grid Jibei Information and Telecommunication Company, Beijing 100053, China*

^b*State Grid Information and Telecommunication Branch, Beijing 100761, China*

Abstract. Recently developed methods such as DETR [1] apply Transformer [2] structure to target detection. The performance of using Transformers for target detection (DETR) is similar to that of two-stage target detector. First of all, this paper attempts to apply Transformer to computer room personnel detection. The contributions of the improved DETR include: 1) in order to improve the poor performance of small target detection. Embed Depthwise Convolution in the encoder. When the coding feature is reconstructed, the channel information is retained. 2) in order to solve the problem of slow convergence in DETR training. This paper improves the cross-attention in DECODE and adds the spatial query module. It can accelerate the convergence of DETR. The convergence speed of the improved method is six times faster than that of the original DETR, and the mAP0.5 is improved by 3.1%.

Keywords. Transformer, DETR, Personnel detection, Cross attention, Depthwise Convolution

1. Introduction

DETR proposes an end-to-end framework based on codec Transformer architecture and two-part matching, which directly predicts a set of bounding boxes without post-processing. However, there are several problems: (1) it takes longer training time to converge. DETR needs 10 to 20 times more training time than modern mainstream detectors to converge. (2) the performance of DETR in small target detection is relatively poor. Current target detectors usually use multi-scale features to detect small targets from high-resolution feature images. For DETR, high-resolution feature maps mean high complexity. The improved DETR can improve these two problems: (I) improve the cross-attention module of the decoder, better learn high-quality content embedding [3], and add spatial query to speed up the convergence speed of training. (II). Embed Depthwise Convolution in the encoder. When the coding feature is reconstructed, the channel information is retained. This can not only improve the detection accuracy, but also improve the detection effect of small targets.

This paper summarizes the research status of general object detection and DETR variants in section 2. Every part of the improved DETR method is described in section 3.1, especially the detailed improvement method is given in the Transformer Encoder

¹ Corresponding Author: Qionglan Na, E-mail: 81885883@qq.com ; This work is supported by the Science and Technology Project of State Grid Jibei Power Company Limited (No. B3018E210000).

part. In section 3.2, we analyze the reasons for the slow convergence of DETR, and introduce the improvement of Transformer Decoder in detail. At section 4.1 we introduced the dataset and the parameter settings during training. In section 4.2, we compare the detection performance of the improved DETR and other algorithms, which shows the effectiveness of the proposed method. In section 4.3, we visualize the output of the cross-attention module. In section 5, we summarize this paper and look forward to the future.

2. Related Work

General object detection. General object detectors can be divided into two categories, namely, two-stage detection and one-stage detection, and one-stage detector is mainly based on anchor point. Anchor points are generated at the center of each sliding window position, providing candidates for objects. For example, SSD[4], RetinaNet[5], YOLO [6] and FCOS[7], the network directly predicts the types and offsets of anchor points in the whole feature graph. Usually, multiple label targets are generated for each anchor point, and then the repeated prediction of the object is removed by post-processing techniques such as NMS. The two-stage detector does not directly output the final prediction. Such as Faster RCNN[8] and Mask RCNN[9], they first generate prospect proposals through the Regional proposal Network (RPN). The ROI Pool[8] or ROI Align [10] layers are used to extract the features generated by RPN from the backbone features. Then NMS and other processing techniques are used to remove repeated predictions to get the final results. In general, the accuracy of the two-stage detector is higher than that of the first-stage detector.

DETR and its variants. DETR successfully applies Transformer to object detection without the need for additional manual design components and performs object detection end-to-end. For example, non-maximum suppression or the generation of an initial anchor. DETR has two problems: slow convergence and poor small target detection. To solve these problems, multi-scale features are applied to help detect small objects in Deformable DETR[11]. And the use of deformable attention module can solve the problem of high computational complexity caused by the self-attention of the global encoder. The adaptive clustering converter[12] clusters the key in the self-attention mechanism and improves the convergence speed by reducing the complexity. The TSP (transformer-based ensemble prediction) method[13] eliminates the cross-attention module and combines FCOS and R-CNN detection heads. The spatial modulation common attention (SMCA) method[14] uses a Gaussian graph around several (shift) centers learned from the decoder embedding to modulate the global crossover attention of the DETR multiple heads to pay more attention to several areas within the estimation frame. Conditional DETR[3] learns conditional space queries from decoder content embedding and predicts spatial attention weights without manual attention decay, highlighting the four extremes of box regression and different regions within the objects used for classification.

this paper attempts to apply Transformer to personnel detection in the computer room. In order to solve the problem of slow convergence of DETR, the spatial query module is introduced to make the algorithm converge quickly. Embed Depthwise Convolution into the Encoder of Transformer. The coding features are reconstructed. Retain channel information as much as possible. This can improve the detection accuracy of the algorithm. The improved method has high convergence speed and detection

accuracy, and has achieved excellent results in the detection of personnel in the computer room.

3. Proposed Method

3.1. Depthwise Convolution in transformer network

The improved DETR, as shown in figure 1, consists of CNN backbone for feature extraction, location coding, six Transformer encoders for feature coding and reconstruction, six transformer decoders for object query, and a simple feedforward network for detection and prediction.

Feature extractor. ResNet50 is used as the backbone of the feature extractor. We use ResNet50 as the feature extraction backbone network. First of all, the image of $x \in \mathbb{R}^{3 \times H_0 \times W_0}$ is used as input. Then Outputting an advanced feature map $h \in \mathbb{R}^{B \times C \times H \times W}$. Where $C = 2048$ and $H, W = \frac{H_0}{32}, \frac{W_0}{32}$. The input feature map is down sampled using the convolution of 1x1 to reduce the dimension of the number of channels to get the feature map $m \in \mathbb{R}^{256 \times \frac{H_0}{32} \times \frac{W_0}{32}}$. Then the feature graph m is serialized and transformed into a serialized feature with a shape size of $(\frac{H_0 \times W_0}{1,024}, 256)$.

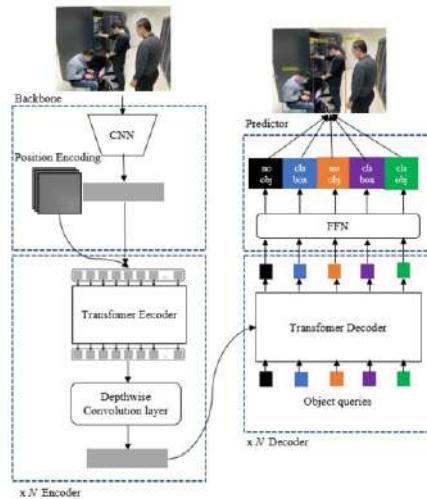


Figure 1. improved overall structure of DETR

Positional Embedding. The author provides two methods of absolute position coding. One is trigonometric function position coding, which is also known as Sinusoidal position coding; the other is training position coding, which directly takes the position coding as a trainable parameter, which is initialized randomly and updated with the training process. Sinusoidal location coding is used in this article.

Encoder. Contains six Transformer encoders with built-in Depthwise Convolution. Our encoder consists of a standard Transformer encoder and Depthwise Convolution layer, each Transformer encoder follows the standard architecture, and consists of a multi-head self-attention module and a feedforward network. Positional embedding is introduced into the input of each self-attention layer, and each Depthwise Convolution

layer consists of a 3×3 convolution, batch normalization, and correction linear unit. As shown in figure 1, the location embedded and extracted features are used as inputs to the Transformer encoder. After Transformer Encoder, the sequence characteristics of $(\frac{H_0 \times W_0}{1,024}, 256)$ are obtained. The feature is reconstructed and deformed to get the shape $(256, \frac{H_0}{32}, \frac{W_0}{32})$ of feature graph. Using 256 convolution kernel of 3×3 for Depthwise Convolution. In this way, the feature information of different channels in the same spatial location can be preserved to the maximum extent. Then through batch normalization and GELU activation unit. Finally, the feature graph m is serialized and transformed into $(\frac{H_0 \times W_0}{1,024}, 256)$ serialized features.

Decoder. Contains six Transformer decoders. The main difference from the original DETR is the input query and cross-attention module. It will be introduced in Section 3.2.

Object predictor. The object predictor is a feedforward network with two fully connected layers, namely, a box regression layer to predict the location of the target (x, y, w, h) and a frame classification layer to predict the target score. Therefore, N object queries are independently decoded into frame coordinates and class labels by the feedforward network, resulting in N final predictions, including object (defect) and no object (background) prediction.

3.2. Cross-attention module based on spatial Query

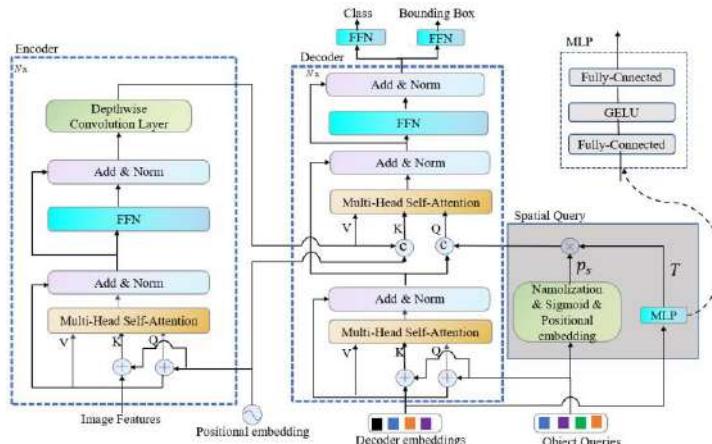


Figure 2. improved DETR codec structure

Due to the slow convergence of DETR, various variants of DETR have been improved. We think that Conditional DETR has given the answer to the reason why DETR converges slowly. DETR highly relies on high-quality content embedding to locate the extremity region of an object, which is the key to locating and identifying objects. In order to make our algorithm converge quickly, as shown in figure 2, we first modify the DETR Cross-Attention, and the output of the original Attention module is shown in formula (1).

$$c_q^T c_k + c_q^T p_k + p_q^T c_k + p_q^T p_k \quad (1)$$

Where c_q is content query, p_q is spatial query, c_k is content key, p_k is spatial key. Among them, c_q and c_k contain more information on the image (color, texture, etc.). p_q and p_k contain more spatial information.

As shown in figure 2, the Query of the decoder is stacked by the c_q and the p_q . The key is stacked by c_k and p_k . The output of the decoder becomes the result of the formula (2).

$$c_q^T c_k + p_q^T p_k \quad (2)$$

Where $c_q^T c_k$ represents the similarity of image information, and $p_q^T p_k$ represents the similarity of computing spatial information.

Depu Meng et al found that decoder embedding and reference point contain spatial location information. This paper also introduces Conditional spatial query prediction and makes simple improvements, as shown in the gray module in the figure.

First of all, the reference point corresponding to query is normalized and mapped to the same bit coding space as spatial key to get reference p_s in formula (3).

$$p_s = \text{sinusoidal}(\text{sigmoid}(s)) \quad (3)$$

Then, we map the offset information contained in decoder embedding to high-dimensional space through an MLP. Get the "offset" for p_s . MLP, we're using the GELU activation function. GELU is not only nonlinear activation, but also introduces the idea of random regularity. The experimental effect is better than that of RELU. Finally, we get p_q by doing the inner product of p_s and T. It is important to note that when we use T as the inner product, we use the values on the diagonal matrix to calculate.

4. Results and Discussion

4.1. Experiment setup

Dateset: A total of 4192 images were marked by professionals, including 10674 tags in one category. The ratio of training sets to test sets is 8:2.

Training: the improved DETR structure uses ResNet-50 pre-trained by ImageNet[15] as the backbone network, the batch norm layer is fixed[16], and the Transformer parameters are initialized using Xavier initialization scheme[17]. Weight falloff is set to 10-4. Use the ADAMW[18] optimizer[19]. The learning rates of backbone and converter are initially set to 10-5 and 10-4, respectively. The shedding rate of the transformer is 0.1. For 50 training cycles, after 40 cycles, the learning rate is reduced 10 times; object query is set to 300. for other default parameters that use the original DETR, and two 12G GTX-1080 are used to complete the training.

4.2. Comparison with DETR

The improved DETR uses the same architecture as DETR, and the improved DETR is negligible in terms of computational cost and training time for each period. Compared with DETR, the improved DETR is due to DETR in accuracy and convergence speed. The improved DETR can achieve better overall performance. the accuracy of our method for training 75 epoch is a little higher than that of the original DETR training 300 epoch. Moreover, the convergence rate of the improved DETR is 6 times faster than that of the original DETR method. As shown in Table 1, DETR needs to train 300 epoch to achieve good performance. In the case of significant improvement in convergence speed and performance, the increase in the number of FLOPS of the improved DETR is very small. Flops increased by 4G. The FPS is reduced by 2 frames. Compared with the Faster RCNN using FPN, the improved method can achieve the performance comparable to the two-

stage algorithm. However, the detection effect of DETR algorithm for small targets is not as good as that of two-stage algorithm, and the detection effect of medium targets and large targets is better than that of Faster RCNN.

Table 1. performance comparison between DETR and improved DETR

Method	epoch	AP	AP50	AP75	APs	APm	API	FPS	FLOPs
Faster RCNN+FPN	110	61.8	89.3	67.6	26.7	48.2	72.1	13	181
DETR	100	56.2	86.9	60.1	14.2	46.8	69.5	15	86
DETR	300	61.4	88.1	66.5	24.4	50.8	74.1	15	86
Ours	75	61.4	90.0	67.1	24.6	51.1	74.2	13	90

4.3. Visualization

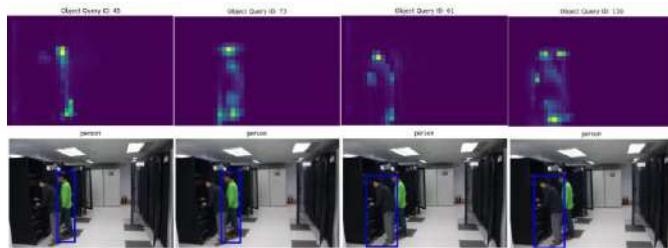


Figure 3. Multi-Attention visualization of DETR Decoders and improved DETR Decoder



Figure 4. Multi-Attention visualization of DETR Decoders and improved DETR Decoder

As shown in figure 3, the first line is the id of the object query, and the second line is the attention matrix diagram of the decoder's Multi-Attention. The third line is the currently detected category, and the fourth line is the corresponding bounding box. Figure 3 shows a photo that detects two targets. The left (DETR) and the right (Ours) are visualization results of the same location. Interesting things can be found by comparison. The improved DETR seems to focus on more areas, which can improve the detection effect of small targets. This is more obvious in figure 4, where the visualization methods of the improved DETR and the original DETR are shown above and below, respectively.

5. Results and Discussion

For the first time, we use Transformer to detect the personnel in the computer room. The detection accuracy can be improved by embedding Depthwise Convolution in the encoder and preserving the channel information when the coding features are reconstructed. In order to make the DETR algorithm converge faster, we improve the cross-attention module of the decoder and mine the location information of the bounding box of the previous decoder as much as possible before spatial query. The improved DETR is used to achieve the detection performance comparable to that of the two-stage detector. The improved detection method has better detection accuracy and convergence speed. The detection method of computer room personnel in this paper provides a good idea to improve the migration to other fields. In the future, we will continue to study the detection of dense occlusion personnel and model compression.

References

- [1] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-End object detection with transformers. InECCV, 2020.1,2,5,6,7,9
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser,L., Polosukhin, I.: Attention is all you need. In: NeurIPS (2017).
- [3] Meng D , Chen X , Fan Z , et al. Conditional DETR for Fast Training Convergence. 2021.
- [4] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. SSD: Single shot multibox detector. InECCV, 2016.2
- [5] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. InICCV,2017.1,2,6,8,9
- [6] Joseph Redmon, Santosh Divvala, Ross Girshick, and AliFarhadi. Y ou only look once: Unified, real-time object detection. InCVPR, 2016.1,2,6
- [7] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He. FCOS:Fully convolutional one-stage object detection. InICCV,2019.2,4,8,9
- [8] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun.Faster R-CNN: Towards real-time object detection with region proposal networks. InNeurIPS, 2015.1,2,6,7,9
- [9] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask R-CNN. InICCV, 2017.2,6,7
- [10] Zhaowei Cai and Nuno V asconcelos. Cascade R-CNN: Delving into high quality object detection. InCVPR, 2018.2
- [11] Xizhou Zhu, Weijie Su, Lewei Lu, Bin Li, Xiaogang Wang, and Jifeng Dai. Deformable DETR: deformable transformers for end-to-end object detection.CoRR, abs/2010.04159, 2020.1,2,6,7,8
- [12] Minghang Zheng, Peng Gao, Xiaogang Wang, Hongsheng Li, and Hao Dong. End-to-end object detection with adaptive clustering transformer.CoRR, abs/2011.09315, 2020.2
- [13] Zhiqing Sun, Shengcao Cao, Yiming Yang, and Kris Kitani.Rethinking transformer-based set prediction for object detection.CoRR, abs/2011.10881, 2020.2,7
- [14] Jifeng Dai, Haozhi Qi, Y uwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. InProceedings of the IEEE international conference on computer vision, pages 764–773, 2017.3
- [15] Xavier Glorot and Y oshua Bengio. Understanding the difficulty of training deep feedforward neural networks. InAISTATS, 2010.6
- [16] Sobti, P ., Nayyar, A., & Nagrath, P . (2021). EnsemV3X: a novel ensembled deep learning architecture for multi-label scene classification. PeerJ Computer Science, 7, e557.
- [17] Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In Journal of physics: conference series (Vol. 1142, No. 1, p. 012012). IOP Publishing.
- [18] Ilya Loshchilov and Frank Hutter. Fixing weight decay regularization in adam, InICLR, 2017.6
- [19] Jain, A., & Nayyar, A. (2020). Machine learning and its applicability in networking. In New age analytics (pp. 57-79). Apple Academic Press.

Design of Speech Denoising Algorithm Based on Wavelet Threshold Function and PSO

Lanyong ZHANG^a, Ruixuan ZHANG^{a1} and Papavassiliou CHRISTOS^b

^a*College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, China*

^b*Department of Electrical and Electronic Engineering, Imperial College London, London, United Kingdom*

Abstract. At present, there are many shortcomings in the discontinuity of wavelet threshold function and the constant threshold of different decomposition layers and the constant error it produced. The amplitude-frequency characteristics of wavelet filters are studied and analyzed by mathematical modeling. An improved wavelet threshold function with adjustable parameters is proposed. Particle swarm optimization (PSO) algorithm is used to find the optimal parameters of the improved threshold function in a background noise environment. The improved wavelet threshold function is combined with Bayesian threshold method to obtain the threshold based on Bayesian criterion, which makes the threshold adaptive in different layers and overcomes the shortcomings of fixed threshold. Finally, the speech signal with optimal wavelet coefficients is obtained after reconstruction. Compared with the traditional threshold function, Simulation results show that the improved threshold function achieves precise notch denoising, effectively retains the singularity and eigenvalues of the signal, and reduces the signal distortion.

Keywords. Speech enhancement, wavelet transform, threshold function, bayesian threshold

1. Introduction

Noise in speech signal is mainly generated in the process of signal acquisition and transmission, and the ultimate purpose of noise elimination is to get more acceptable speech signal from noisy signal. Nowadays, many denoising algorithms have been proposed[1]. At present, wavelet technology and its improved algorithms are widely used in speech recognition and mobile communications. References [2] proposed a method to improve the threshold function in order to overcome the shortcomings of the wavelet transform soft and hard threshold method. References [3] The improved method only has a certain effect on periodic signals, and the effect scene is relatively single.

There are other algorithms, such as Spectral Subtraction (SS), Wiener Filter(WF), Least Mean Square algorithm (LMS), Subspace Method(SM), Deep Neural Networks(DNN) and so on. There are various problems in these methods, Therefore, scholars have proposed improved methods or new methods[4-7], such as wavelet

¹ Corresponding Author, Ruixuan Zhang, College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, China; Email: 13643292271@163.com

transform enhancement algorithm is one of them[8]. Aiming at the deficiencies of traditional denoising methods that the threshold is fixed, this paper proposes a method of combining wavelet threshold function with adjustable parameters and Bayesian threshold is proposed, The particle swarm optimization algorithm is used to find the optimal adjustment parameters[9], which makes it adaptive at different levels and can remove clutter more effectively, Finally, simulation experiments verify the effectiveness of the new algorithm.

2. Wavelet Threshold Function Denoising

2.1 Denoising Principle of Wavelet Threshold Function

For a noise signal $y(n)$ with length N , the model can be expressed as:

$$y(n) = x(n) + e(n) \quad (1)$$

Among them, $y(n)$ is the contaminated signal, $x(n)$ represents a pure signal and $e(n)$ is a noise signal[10].

Following are the basic steps of multi-scale wavelet denoising[11]:

- (1) For the noisy signal, first set a critical threshold, as a comparison value for comparison with wavelet coefficients, the multi-scale wavelet decomposition of the noisy signal is carried out by choosing the appropriate wavelet base and the number of decomposition layers, so that the wavelet coefficients of each layer can be obtained.
- (2) Choose appropriate threshold and threshold function for each layer of high frequency wavelet system, and get the wavelet coefficients after each processing.
- (3) The enhanced speech signal can be obtained by reconstructing the high-frequency and low-frequency wavelet coefficients of each layer.

2.2 Improved threshold function

From the analysis of the denoising steps of the wavelet threshold function, we can see that the denoising effect is related to the threshold function, the noise threshold of each layer and the reconstruction accuracy, therefore, the improvement of the threshold function is the key content of people's research. There are two common methods of traditional threshold function: soft threshold function and hard threshold function[12-13]. The hard threshold function is expressed as:

$$y_{j,k} = \begin{cases} w_{j,k} & |w_{j,k}| > \lambda \\ 0 & |w_{j,k}| \leq \lambda \end{cases} \quad (2)$$

When the wavelet coefficients are larger than the threshold value, the basic characteristics of the original initial signal are preserved to a certain extent. Since the threshold function is discontinuous at the threshold $\pm \lambda$, the pseudo-Gibbs phenomenon will easily occur when the wavelet coefficients are reconstructed after denoising.

Soft threshold function:

$$y_{j,k} = \begin{cases} sgn(w_{j,k})^*(w_{j,k} - \lambda) & |w_{j,k}| > \lambda \\ 0 & |w_{j,k}| \leq \lambda \end{cases} \quad (3)$$

Formula (3) shows that there is a constant difference between the wavelet coefficients $w_{j,k}$ of noisy signals and the output wavelet coefficients $y_{j,k}$. Compared with the hard threshold function, the soft threshold function is continuous at $\pm \lambda$, but the reconstruction accuracy of the soft threshold function is not enough, this results in the loss of some useful speech signals. From the above analysis, we can see that both hard threshold function and soft threshold function have some shortcomings. The expression of improved wavelet threshold function in reference [14] is as follows:

$$y_{j,k} = \begin{cases} \text{sgn}(w_{j,k}) * \left(w_{j,k} - \lambda * k \sqrt{w_{j,k}^2 - \lambda^2} \right) & |w_{j,k}| > \lambda \\ \text{sgn}(w_{j,k}) * k * \frac{w_{j,k}^2}{\lambda} & |w_{j,k}| \leq \lambda \end{cases} \quad (4)$$

Among them, λ is the threshold of each decomposition wavelet coefficients, k is the adjustment parameter. From the proposed threshold function formula, we can see that the change of the threshold function is between the soft and hard threshold functions, and the denoising effect is better than the traditional method to a certain extent. The modulus transformation of coefficients after wavelet transform conforms to the exponential attenuation characteristic. When the threshold value is lower than the threshold value, the threshold function does not satisfy the exponential form. This will inevitably result in unsatisfactory denoising and speech distortion, thus affecting the quality of reconstructed speech signal.

Reference [15] pointed out that the wavelet threshold function is based on an improved non-negative dead zone threshold function:

$$y_{j,k} = \begin{cases} \text{sgn}\left(w_{j,k} - \frac{\lambda^2}{2|w_{j,k}|}\right) * e^{2(\lambda - |w_{j,k}|)} & |w_{j,k}| > \lambda \\ \text{sgn}(w_{j,k}) * \frac{\lambda(e^{8|w_{j,k}|} - e^{8p})}{2(e^{8\lambda} - e^{8p})} & |w_{j,k}| \leq \lambda \end{cases} \quad (5)$$

In the formula $p \in (0, \lambda)$, although the improved threshold function solves the shortcoming that the wavelet coefficients are directly zero when the non-negative threshold function is less than the threshold value, when the wavelet decomposition coefficients are larger than the threshold value, the threshold function is basically fixed when the input noise intensity is different, which will cause distortion to the enhanced speech quality.

Considering that soft and hard threshold functions are odd functions, combining the above two ideas of threshold functions. In this paper, an improved method based on exponential function is proposed, and the exponents with different bases are set at the threshold of absolute value λ . This method can better denoise the wavelet coefficients of each layer of noisy signal and greatly improve the effect of denoising. The following is the improved expression of threshold function: the place is also continuous, and the following is the improved expression of threshold function:

$$y_{j,k} = \begin{cases} \text{sgn}(w_{j,k}) * \left(w_{j,k} - \frac{\lambda}{1+\alpha} * \gamma \sqrt{w_{j,k}^2 - \lambda^2} \right) & |w_{j,k}| > \lambda \\ \text{sgn}(w_{j,k}) * \frac{\alpha}{1+\alpha} * e^{10 * (|w_{j,k}| - \lambda)} * |w_{j,k}| & |w_{j,k}| \leq \lambda \end{cases} \quad (6)$$

In order to prove the continuity of the improved threshold function, since the improved threshold function is also an odd function, it is only necessary to prove that it is continuous in the right half of the definition domain.

When $|w_{j,k}| \rightarrow \lambda$:

$$\lim_{w_{j,k} \rightarrow \lambda^+} sgn(w_{j,k}) * \left(|w_{j,k}| - \frac{\lambda}{1+\alpha} * \lambda \sqrt{w_{j,k}^2 - \lambda^2} \right) = \frac{\alpha}{1+\alpha} * \lambda \quad (7)$$

$$\lim_{w_{j,k} \rightarrow \lambda^-} sgn(w_{j,k}) * \frac{\alpha}{1+\alpha} * e^{10(|w_{j,k}| - \lambda)} * |w_{j,k}| = \frac{\alpha}{1+\alpha} * \lambda \quad (8)$$

From Formulas (7) and (8), we can see $\lim_{w_{j,k} \rightarrow \lambda^+} = \lim_{w_{j,k} \rightarrow \lambda^-} = \frac{\alpha}{1+\alpha} * \lambda$, So we can know that the threshold function proposed at λ is continuous. Similarly, we can know that the function at $-\lambda$ is also continuous.

When $|w_{j,k}| > \lambda$, The output signal $w_{j,k}$ is between $|w_{j,k}| * sgn(w_{j,k})$. This is because we take any real number for α , $0 < \frac{1}{1+\alpha} \leq 1$, And Formula $\sqrt{w_{j,k}^2 - \lambda^2}$ knows, when $\gamma \in (0, 1)$, The value of arbitrary $w_{j,k}$, the values are between (0, 1), when $w_{j,k} \rightarrow \infty$, $\sqrt{w_{j,k}^2 - \lambda^2}$ approaches zero. therefore $y_{j,k} \rightarrow w_{j,k}$, In this way, the new threshold function does not have the problem of fixed deviation value relative to the soft threshold function.

2.3 Threshold selection

In the wavelet threshold de-noising algorithm, the acquisition of threshold value lambda also determines the effect of de-noising. If the value of λ is too large, some useful signals in speech signal will be removed in the process of de-noising, which will easily lead to speech distortion. On the contrary, if the value of λ is too small, there will be more residual noise components in de-noising, resulting in unsatisfactory de-noising effect[16]. In order to reduce the threshold of different layers, it is necessary to be able to adaptively follow the increase of wavelet decomposition scale. So the Bayesian threshold is used to acquire the noise threshold, and the following process is calculated concretely:

$$r(T) = E(x-x)^2 = E_x E_{y|x} (x - x)^2 = \iint (n(y) - x)^2 P(y|x) P(x) dy dx = \delta^2 \omega\left(\frac{\delta_x^2}{\delta^2}, \frac{T}{\delta}\right) \quad (9)$$

Among,

$$\omega(\delta_x^2, T) = \delta_x^2 + 2(T^2 + 1 - \delta_x^2) \bar{\phi}\left(\frac{T}{\sqrt{1+\delta_x^2}}\right) - 2T(1 + \delta_x^2) * \phi(T, 1 + \delta_x^2) \quad (10)$$

Density function:

$$\phi(x, \delta_x^2) = \frac{1}{\sqrt{2\pi\delta_x^2}} \exp\left(-\frac{x^2}{2\delta_x^2}\right) \quad (11)$$

$$\bar{\phi}(x) = \int_x^\infty \phi(t, 1) dt \quad (12)$$

Threshold calculation expression:

$$T_j = \frac{\delta^2}{\delta_x} \quad (13)$$

In the above formula, δ^2 is the noise variance, δ_x is the standard deviation of subband coefficients, and j represents a layer in the stratification. For formula (13), the estimates proposed by Donoho and Johnstone are used to calculate δ^2 .

$$\bar{\delta} = \frac{\text{median}(|y_{j,k}|)}{0.6745} \quad (14)$$

y (i) in the formula represents the wavelet coefficients of different layers.

$$\bar{\delta} = \frac{1}{n} \sum_{i=1}^n y_{j,k}^2 \quad (15)$$

Among them, n represents the length of each decomposition wavelet coefficients, which can be seen from $\delta_y^2 = \delta_x^2 + \delta^2$:

$$\overline{\delta_x} = \sqrt{\max(\delta_y^2 - \delta^2, 0)} \quad (16)$$

Therefore, the threshold based on Bayesian criterion can be obtained by (14) 15 (16). It can be seen that the Bayesian threshold has self-adaptability in different layers and overcomes the shortcomings of fixed threshold, can better realize the wavelet threshold function denoising.

3. Parameter acquisition of improved threshold function based on particle swarm optimization

3.1 Wavelet Denoising Fitness Function

In this paper, an improved threshold function is proposed. There are two adjusting parameters, namely, lambda and alpha. In the face of different input signal noise ratios, in order to improve the threshold function, the adjusting parameters can be changed to adapt to different input signal-to-noise ratios. Particle swarm optimization (PSO) is used as a tool to find the optimal adjustment parameters. The PSO algorithm is a simulation of the social information sharing mechanism of a group of birds foraging for food. Each particle in the PSO algorithm is a solution in the solution space. It adjusts its flight according to the flight experience of itself and its companions. Get the optimal solution in the current solution space. Taking output signal-to-noise ratio (SNR) as fitness function, the larger the output signal-to-noise ratio, the better the denoising effect[17]. The SNR criteria are defined as follows (17):

$$\text{SNR} = 10 \lg \frac{\sum_{i=1}^N x(i)^2}{\sum_{i=1}^N (y(i) - x(i))^2} \quad (17)$$

Among, x (i) denoised signal, y (i) denoised input signal and N denoised input signal.

3.2 PSO algorithm and Optimizing Threshold Function to Adjust Parameter Estimation

PSO algorithm initializes a group of particles as a feasible solution. The best position that each particle passes through is the optimal solution found by the particle itself, which is called the individual extreme value, and the best position that the entire group passes through is the optimal solution found for the entire group, which is called the global

extreme value[18]. The following formula is that particles update their own speed and position through individual and group extremes.

$$v(t+1) = v(t) + c_1 * rand() * [p_i(t) - x_i(t)] + c_2 * rand() * [zbest(t) - x_i(t)] \quad (1 \leq i \leq N) \quad (18)$$

$$p_i(t) = p_i(t) + 0.5 * v(t) \quad (19)$$

Threshold function is continuous and differentiable, which is beneficial to speech signal denoising. In this paper, particle swarm optimization algorithm is used to determine the optimal adjustment parameters of threshold function. The steps of denoising are as follows:

- (1) Four wavelet coefficients of noisy speech signals are obtained by four-layer wavelet transform. The wavelet coefficients of different layers are used as the input of particle swarm optimization, where the four groups of lengths are $N/2$, $N/4$, $N/8$ and $N/16$ respectively.
- (2) Setting some parameters of particle swarm optimization algorithm to initialize, the learning factor $C1 = 1.50$, $C2 = 1.50$, the number of particles $N = 20$, limiting the speed of particles to a range of $[-0.5, 0.5]$, the number of iterations to 300, and the value range of particles to $[0, 10]$;
- (3) The fitness values of all particles in the first iteration are calculated according to the fitness function expressed in formula (17). The maximum fitness values of individuals and groups are determined from the initial fitness values of particles, which are expressed by $gbest$ and $zbest$ values respectively.
- (4) According to formula (18) (19), the velocity and fitness values of each particle in each generation are updated and determined to be within a given range. Then, the fitness values of the next generation of particles are updated according to the fitness function SNR.
- (5) Renewal of the optimum values $gbest$ and global optimum $zbest$ in individual particles.
- (6) If the maximum number of iterations is met, the update will be stopped, otherwise the update will continue to step (4).
- (7) The optimal value required for output is $zbest$, which is obtained by the last stable iteration.
- (8) The optimal adjustment parameters are obtained, which are substituted into the improved threshold function and combined with Bayesian threshold to denoise, and then the denoised wavelet coefficients are reconstructed to obtain the enhanced speech signal.

4. Simulation results and analysis of experiments

This paper uses MATLAB to process noisy speech signals with different input SNR. The pure speech signal used in the experiment is a SP01 English speech. The content is Birch can slide on smooth boards. The length of the data is 22529. Background noise is white, dining room noise and pink noise, respectively. Voice and noise are taken from NOIZEUS database and Noisex-92 standard noise database. The input SNR is set according to the different logarithmic power ratio of the noise signal to the pure speech signal of the same time length. The selected SNR of the noisy speech is -10 dB, -5 dB, 0 dB, 5 dB, 10 dB and 15 dB, respectively. Speech and noise sampling frequencies are 8000 Hz, single channel, sampling bit is 16 bit. Here, sym6 is chosen as the wavelet basis function. The signal decomposition scale is $n=4$. Documents [7]([13],[15]) and the

improved threshold function method are used to process the noisy signal. The following two tables are speech enhancement data under different threshold functions.

Figure 1 shows the curves of output SNR. From the output SNR curves, it can be seen that the improved threshold function has better denoising effect than other thresholding functions under the same input noise conditions, especially in the low signal-to-noise ratio of -10dB 5dB, the SNR increases more than 1 dB, and it is also improved in other ranges. The curve of output RMSE shown in Figure 2 shows that the proposed threshold function has higher stability and smaller changes. These two curves show that the improved threshold function has better noise reduction effect, stronger anti-interference and less distortion of speech.

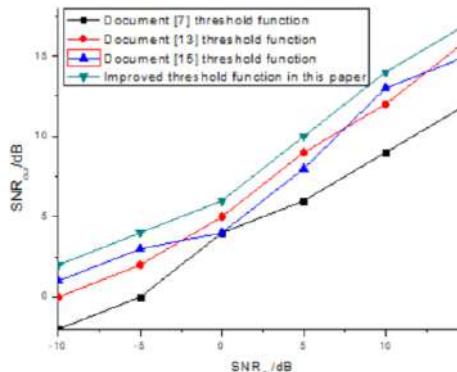


Figure 1. Output SNR Effect Contrast Curve

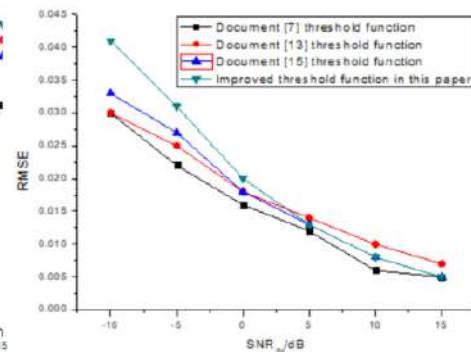


Figure 2. Output RMSE Effect Contrast Curve

Table 1 The LSD obtained with various threshold function in different backgrounds

Noise	SNR _{in} /dB	Document[7]	Document[13]	Document[15]	Improved threshold function
White	-5	2.3572	2.3533	2.4485	2.1074
	0	2.22592	2.2925	2.2248	1.992
	5	2.6561	2.9417	2.6367	2.2573
	10	2.4769	2.6519	2.4244	2.0729
	-5	2.5024	2.8985	2.8401	2.245
Pink	0	2.6802	2.9139	3.004	2.4468
	5	2.5809	2.5597	2.7417	2.2683
	10	2.3483	2.1798	2.3724	2.0895
Babble	-5	2.9137	2.8448	2.8678	2.6558
	0	2.5566	2.6244	2.5481	2.5113
	5	2.2415	2.2366	2.2284	2.1732
	10	1.9202	1.8551	1.8882	1.8487

Table.2 Comparison of output PESQ of various threshold function

Noise	SNR _{in} /dB	Document[7]	Document[13]	Document[15]	Improved threshold function
White	-5	0.7884	1.435	1.5202	1.5792
	0	1.2699	1.6698	1.6693	1.8124
	5	1.9731	1.9853	1.8357	2.071
	10	2.2386	2.2504	2.1074	2.3815
Pink	-5	1.429	1.4481	1.6207	1.6469
	0	1.7078	1.6771	1.7194	1.8562
	5	2.1105	2.0934	2.0182	2.1614
	10	2.3991	2.4492	2.3435	2.4365
Babble	-5	1.5704	1.6071	1.5944	1.5585
	0	1.9082	1.875	1.9428	1.8921
	5	2.2338	2.1911	2.2349	2.2358
	10	2.5114	2.5226	2.5121	2.5323

Table 1 and Table 2 show PESQ and LSD values of noisy speech signals with different signal-to-noise ratios under different background noise conditions. The larger the PESQ value, the better the denoising effect. On the contrary, the smaller the LSD value, the better the denoising effect. It can be seen from the table that the improved threshold function algorithm is better than other algorithms in White noise and Pink noise, especially in low input signal-to-noise ratio. The PESQ and LSD values are better than other threshold function algorithms in this paper. In the case of Babble noise, although PESQ is slightly lower than other algorithms, it is better than others in the logarithmic spectrum distance. It may be that the speech spectrum of Babble noise with similar spectrum will pollute the intelligibility of low frequency band more obviously. However, the overall effect of the proposed algorithm is better than that of other algorithms in these noisy environments. More useful speech signals are retained, and the quality and clarity of speech perception are improved to a certain extent, which provides ideal pure signals for further processing of speech signals.

5. Conclusion

The improved threshold function proposed in this paper avoids the problems of fixed deviation and discontinuity at the threshold generated in the wavelet transform, and combines with the Bayesian threshold acquisition method to solve the problem of constant thresholds for different layers. In this paper, we use pure speech signal to add Gaussian white noise with different signal-to-noise ratios and then use different threshold functions to denoise the simulation, and compare the denoising effects of different threshold functions from the signal-to-noise ratio and root mean square error. The results show that the improved threshold function in this paper has a better denoising effect than other threshold functions, which improves the overall speech quality after denoising and achieves the ideal speech denoising effect. For the problem of unobvious multi-channel coincidence noise optimization, future work can consider introducing the latest fitness dependent optimize algorithm (FDO) and Dynamic Cat Swarm Optimization algorithm to achieve better denoising effects.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China subsidization project (51579047), the Natural Science Foundation of Heilongjiang Province (QC2017048), the National Defense Technology Fundamental Research Funds (JSHS2015604C002), the Natural Science Foundation of Harbin (2016RAQXJ077), the fundamental research funds for the central universities (HEUCF180407), and the Open Project Program of State Key Laboratory of Millimeter Waves (K201707).

References

- [1] Goel P, Garg A. Developments in spectral subtraction for speech[J]. International Journal of Engineering Research and Application, 2012, 2 (1) :55-63.
- [2] Qian Quliang, Dong Baowei. Application of improved wavelet threshold denoising in bearing fault diagnosis[J]. Data Communication, 2020, 3(4): 17-19.
- [3] Zhao Qiang, Yang Yang, Wang Yu, Guo Changjiang. An improved wavelet threshold denoising method and its simulation [J]. Modern Computer, 2021(10): 126-128.
- [4] El-Fattah M A A, Dessouky M I, Abbas A M. Speech enhancement with adaptive wiener filter[J]. International Journal of Speech Technology, 2014, 17 (1) :53-64.
- [5] Ma Jiachen, Hu Jiajun, Ma Liyong, Zhang Yong, Xie Wei. Active noise control based on improved particle swarm algorithm [J]. Computer Application Research, 2015, 9: 35-39.
- [6] U Rajendra Acharya, Hamido Fujita, Muhammad Adam, et al. Automated Characterization and Classification of Coronary Artery Disease and Myocardial Infarction by Decomposition of ECG signals:A comparative Study[J].Information Sciences, 2017, 377:17-29.
- [7] Galiano G, Velasco J. Rearranged Nonlocal Filters for Signal Denoising[J].Mathematics and Computers in Simulation, 2015, 118:213-223.
- [8] Zhentao Yu, Jie Chen, Suqin Xu, Cheng Chi, Tingting Li. Wavelet-based Adaptive Enhancement Method of Aeromagnetic Anomaly Signal[J]. IOP Conference Series: Earth and Environmental Science, 2019, 237(3):032017 (6pp).
- [9] Manuel Vilares Ferro, Víctor M. Darriba Bilbao, Jesús Vilares Ferro. Adaptive scheduling for adaptive sampling in pos taggers construction[J]. Computer Speech & Language, 2020, 60:
- [10] LU Y, Huang YM, Xue W, Zhang GB. Seismic data processing is based on wavelet transform method to denoise [J]. Cluster calculation, 2019, 22 (3): 6609-6620. DOI: 10.1007/s10586-018-2355 -0.
- [11] Du Shiqiang, Song Yukun, Zhang Xuan, Zhang Duoli. An improved wavelet threshold denoising algorithm [J]. Microelectronics and Computer, 2021, 38(02): 40-46.
- [12] Liu Fengshan, Lv Zhao, Zhang Chao, Wu Xiaopei. Research on Speech Enhancement Alogrithm Based on Modified Wavelet Threshold Function [J]. Journal of Signal Pro- cessing, 2016, 32(2):203-213.
- [13] Han Wang, Wenxi Lu, Zhenbo Chang. Simultaneous identification of groundwater contamination source and aquifer parameters with a new weighted-average wavelet variable-threshold denoising method[J]. Environmental Science and Pollution Research, 2021, :1-16.
- [14] Lu Yong. The application of improved wavelet threshold function in speech enhancement[J]. Information Technology and Network Security, 2019, 38(08): 38-41.
- [15] Qin Aina, Dai Liang, Li Fei, Cao Weihua. A Speech Enhancement Algorithm Based on Improved Wavelet Threshold Function [J]. Journal of Hunan University:Nat- ural Sciences, 2015, 42 (4):136-140.
- [16] Wang Ze, Wan Feng, Wong Chiman, Zhang Liming. Adaptive Fourier Decomposition Dased ECG Denoising[J]. Computers in Biology and Medicine, 2016, 77:195-205.
- [17] Zhan Zhan, Qin Huibin. Wavelet threshold denoising algorithm based on new threshold function[J]. Computer Technology and Development, 2019, 29(11): 47-51.
- [18] Liu Tianyu, Wang Zhu. A multi-objective particle swarm optimization algorithm with diversity control[J]. Journal of Xidian University (Natural Science Edition), 2021,48(3):106-114. DOI:10.19665/j.issn1001-2400.2021.03.014.

Application of Text Mining Method for Classification of Work Order in Power Grid Production

Yadi ZHAO¹, Zhifeng WEI, Bingqiang GAO and Shuo ZHANG

State Grid Information & Telecommunication Accenture Information Technology Co., Ltd.-Beijing, China

Abstract. With the completion of the State Grid Corporation's maintenance system, the number of substations has increased dramatically, the grid structure has become increasingly complex, and there have been internal and external reasons such as the contingency of emergencies, and equipment failures have occurred from time to time. This paper aims to explore the potential value of massive data, show the laws of business data, and further give full play to the comprehensive support of data for enterprise operation and production management, and promote the realization of intelligent and lean power grid core business. This paper uses power system data to provide reliable data support for equipment defect full cycle management and equipment state analysis through ANOVA and neural network statistical analysis. At the same time, we use Term Frequency-Inverse Document Frequency(TF-IDF)Algorithm to calculate the importance of keywords and construct the power keyword library. By constructing Bayesian text classification model, we can classify the defect parts, defect categories and defect causes automatically. This method can be applied to the construction of power grid production work order text analysis system, improve the data quality and system automation level, help the business department to improve work efficiency and provide the basis for power grid business analysis. This method is applied to the data cleaning of the primary production equipment of power grid enterprises, and the accuracy of data error correction for equipment defects with voltages above 110kV is between 93% and 95%, and good results have been achieved.

Keywords. Power system data, statistical analysis, text mining

1. Introduction

Under the influence of the development of modern enterprises and the macro-economy, as the operators of the main power network business and the main trading body of the regional power market, Power grid enterprises have higher demand for the timely and effective enterprise management operation data[1]. With the continuous development of network technology and information technology, the digital degree of power grid enterprises is gradually improved, the application of equipment (asset) operation and maintenance lean management system (PMS2.0) is deepening, while the continuous increase of information system, the enterprise data is also growing greatly[2]. Extension

¹ Corresponding Author, Yadi ZHAO, State Grid Information & Telecommunication Accenture Information Technology Co., Ltd.-Beijing, China; E-mail: 15510266633@163.com

of digital information enables them to realise on an increased efficiency of the electrical grids, and further control the network dynamically to improve the security and reliability[3]. With the increasing increase of substation scale, the grid structure is increasingly complex, equipment failure occurs, the traditional equipment failure prediction lacks the systematic scientific support and effective early warning and decision-making mechanism. Meanwhile, most of the business data reporting of power grid enterprises in China still adopts manual method, which is easily affected by personnel quality and department management mode, that makes the data appear inconsistent or inaccurate in the process of creating, resulting in serious data crossover and redundancy[4].

How to better tap the potential value in these massive data, further give play to the comprehensive support of data for enterprise operation and production management, and promote the intelligent and lean realization of the core business of the power grid has become an important topic facing enterprises in the era of big data. State Grid Corporation proposes a new strategy for developing leading international energy Internet companies with Chinese characteristics, accelerates the construction of a new power systems. It requires greater use of corporate data value, big data intelligent analysis and multi-source big data governance, active enabling business management and external business services[5]. As an important foundation, big data method has been applied to many specialties in the field of electrician to improve data quality and realize reasonable operation of data[6-7]. Chen Yongjun analyzed the defect data of power network operation analysis decision support system and proposed a time series method for equipment defect prediction[8]. Tarik provides a text classification, which is a process of automatically assigning sets of documents into class labels depending on their data contents[9]. From the defect data of secondary equipment, Zhang Yanxu et al proposed a defect data mining method based on Apriori algorithm, which can effectively analyze the weak links of secondary equipment and provide reference for equipment control[10]. At the same time, we also used big data methods to carry out electricity recovery risk prediction and electricity potential sensitive customer prediction, which achieved relevant results[11-12].

For descriptive text data, text mining method can be used to organize and solidify the data based on the low efficiency and incomplete data[13-15]. This paper makes full use of the data of PMS2.0 and other systems, through text mining[16] and data analysis of equipment defects and other data to provide reliable data support for the whole cycle management of equipment defects and equipment status analysis, providing accurate and effective auxiliary decision-making suggestions and improving the level of lean transportation and accurate maintenance. Through this study, we have realized the correct prediction of the probability of defects of the main network equipment and treatment in advance, reduce the risk of power grid operation, ensure that the management personnel have more energy to study and judge the management process, making lean scientific management possible.

This paper aims to explore the potential value of massive data, show the laws of business data, and further give full play to the comprehensive support of data for enterprise operation and production management, and promote the realization of intelligent and lean power grid core business. The paper first introduces some theoretical methods, such as Pearson correlation analysis, analysis of variance, text mining, logistic regression algorithm and neural network algorithm. And then, we apply these theoretical methods for equipment fault prediction and equipment defect management. Finally, we make the conclusion and propose the content of future work.

2. Theoretical methods

2.1. Pearson correlation analysis

Correlation analysis is the process of describing the degree of close relationship between objective things with appropriate statistical indicators. The degree of correlation between the two variables is expressed by the correlation coefficient r . When two variables are positively correlated, the r values are between 0 and 1, indicating an increase as the other variable increases, and when two variables are negatively correlated, the r values are between -1 and 0, indicating that the other variable decreases as one variable increases. The closer the absolute value of r to 1, the stronger the association of the two variables and the weaker vice versa.

2.2. Analysis of variance

The ANOVA was investigated by analyzing the contribution magnitude of variants of different sources to the total variants, thus determining the magnitude of the influence of controlled factors on the study results. The rationale for ANOVA is that there are two basic sources of differences between means across treatment groups, and we used F values to infer whether each sample came from the same population.

- Experimental conditions: differences caused by different treatments are called intergroup differences, it is expressed as the sum of square of the deviation between the mean in each group.(Ms_b)
- Random error: differences caused by measurement error or differences between individuals, called within-group differences, are expressed as the sum of the square sum of the deviation of the variable mean in each group from the values of the variables within that group.(Ms_w)
- The MS_b/MS_w ratio constitutes the F distribution.

2.3. Text MINING

Text Mining refers to the classification of a text into a known text category, mainly including text preprocessing and classifier model construction. The text classification process is shown in Figure 1.

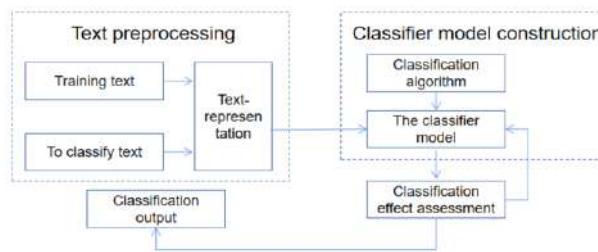


Figure 1. Text classification process

Text preprocessing refers to through Chinese text segmentation, text representation, text feature selection, etc procedures, change the text data table that expresses the natural

language into the machine language data that the computer can handle. Text representations are often represented by a vector space model: $d_i = (t_{i1} : w_{i1}, t_{i2} : w_{i2}, \dots, t_{in} : w_{in})$, d_i represents the text, t_{ij} represents the word j in the i text, w_{ij} represents the weight of the t_{ij} in the text d_i . The classifier model construction process is to generate the classifier model based on the training set data after text pretreatment through relevant algorithms, and use the classifier model to automatically classify the new text data.

In this paper, text preprocessing is implemented by jieba word segmentation and the term frequency-inverse document frequency method(TF-IDF), the classifier construction takes a Bayesian classification algorithm.

2.3.1. Text Segmentation Algorithm

Chinese word segmentation can be roughly divided into two categories: dictionary-based word segmentation and statistics-based word segmentation methods. The word segmentation method used in this paper is jieba word segmentation, which is a combination of the two before and after, that is, based on the prefix dictionary to scan the word graph to form a directed acyclic graph of all possible word segmentation results, and dynamic programming to find the path of maximum probability, thus solving the problem of double understanding word combination.

In statistics-based word segmentation method, the statistical sample content comes from some standard corpora. If there is a sentence M , it has m word segmentation options:

$$\{A_{i1}A_{i2}\dots A_{in_i}, A_{21}A_{22}\dots A_{2n_2}, \dots, A_{m1}A_{m2}\dots A_{mn_m}\} \quad (1)$$

where the subscript n_i represents the number of words in the i th participle. If we choose the best segmentation method r , the statistical distribution probability corresponding to this word segmentation method should be the largest, which is: $r = \arg \max_i P(A_{i1}, A_{i2}, \dots, A_{in_i})$. In order to simplify calculations, Markov hypothesis is usually used, that is, assuming that the probability of each participle is only related to the previous participle, it can be written as:

$$P(A_{ij} | A_{i1}, A_{i2}, \dots, A_{i(j-1)}) = P(A_{ij} | A_{i(j-1)}) \quad (2)$$

Then the joint distribution can be obtained as:

$$P(A_{i1}, A_{i2}, \dots, A_{in_i}) = P(A_{i1})P(A_{i2} | A_{i1})P(A_{i3} | A_{i2}) \dots P(A_{in_i} | A_{i(n_i-1)}) \quad (3)$$

Through the standard corpus, the binary conditional probability between any two word segments can be calculated approximately. For example, any two words a_1, a_2 , the conditional probability distribution can be approximately expressed as:

$$\begin{aligned} P(a_2 | a_1) &= \frac{P(a_1, a_2)}{P(a_1)} \approx \frac{\text{freq}(a_1, a_2)}{\text{freq}(a_1)} \\ P(a_1 | a_2) &= \frac{P(a_2, a_1)}{P(a_2)} \approx \frac{\text{freq}(a_1, a_2)}{\text{freq}(a_2)} \end{aligned} \quad (4)$$

where $\text{freq}(a_1, a_2)$ indicates the number of times two words appear next to each other in the corpus, $\text{freq}(a_1)$, $\text{freq}(a_2)$ respectively indicates the statistical times that each word appears in the corpus. Then for a new sentence, using the corpus to establish statistical probability, the word segmentation method corresponding to the maximum probability

can be found by calculating the joint distribution probability corresponding to various word segmentation methods, which is the optimal word segmentation.

2.3.2. TF-IDF algorithm

Using Jieba word segmentation can divide the document, but not every word is meaningful (the word contributes less to the content of the document.), so we use TF-IDF method to judge the importance of words to the document.

The main idea of TF-IDF is: if a word or phrase appears in a certain category with a high frequency of TF and rarely appears in other categories, it is considered that this word or phrase has good classification ability, suitable for classification, and can be marked as a keyword glossary. Suppose that d is the classification category, f is a feature word based on the content of the ticket, $TF_{f,d}$ represents the number of times a given word appears in the category.

Document frequency is represented by DF , it represents the number of all categories containing a vocabulary in a sample set, it is mapped to a smaller value range and expressed as inverse document frequency (IDF) as follows:

$$IDF_f = \log(N/(DF_f + 1)) \quad (5)$$

in which N represents the total number of categories in the sample collection, DF_f represents the number of categories that contain the vocabulary f . The significance is that if the fewer categories include the word, the larger the IDF, which means that the vocabulary has a good classification ability. The weight of the term is represented by TF-IDF, and the calculation formula is as follows:

$$TF-IDF = (TF_{f,d}) * (IDF_f) \quad (6)$$

This method can calculate the importance of a certain vocabulary in a certain classification, identify whether the vocabulary is a keyword vocabulary, realize the construction of classification rules, and accurately locate the basic content of the text description, so as to provide a basis for subsequent classification analysis.

2.3.3. Bayesian Classification Algorithm

In the naive Bayes classifier, it is determined based on each feature that the label should be assigned to the given input value. Besides, the Naive Bayes classifier determines the label prior probability by calculating the frequency of each label on the training set, and the contribution of each feature is combined with its prior probability to obtain the likelihood estimate of each label. The label with the highest likelihood estimate will be assigned to the input value. Let each data sample describe the value of the n properties with a n -dimensional eigenvector $X = \{x_1, x_2, \dots, x_n\}$, we assume m categories which are represented by $\{c_1, c_2, \dots, c_m\}$, respectively. Given an unknown data sample X (there is no class number), if we assign an unknown sample X to the class c_i , it must be $P(c_i|X) > P(c_j|X) \quad 1 \leq j \leq m, j \neq i$. Basis Bayes Theorem, because of $P(X)$ is the known constant, the maximizing posterior probability $P(c_i|X)$ can be converted to maximizing prior probability $P(X|c_i)P(c_i)$. Assuming that the values of each property are independent of each other, therefore the prior probabilities $P(x_1|c_i), P(x_2|c_i), \dots, P(x_n|c_i)$

can be obtained from the training dataset. The classification principle is shown in Figure 2.

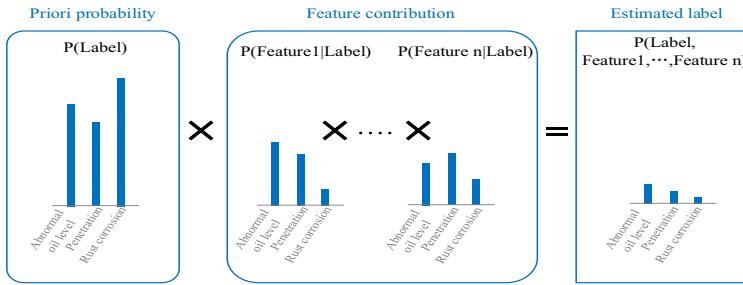


Figure 2. Classification principle description

Suppose that label is the output label, features represent the n eigenvalues of the input. The calculation process of the algorithm is as follows:

$$\text{Calculate } P(\text{features}) = \sum_{\text{label} \in \text{labels}} P(\text{features, label}) \quad (7)$$

The likelihood label can be expanded as the probability of the label multiplied by the probability of a given label feature, in the case of feature independence, we have:

$$\begin{aligned} P(\text{features, label}) &= P(\text{label}) \times P(\text{features} | \text{label}) \\ &= P(\text{label}) \times \prod_{f \in \text{features}} P(f | \text{label}) \end{aligned} \quad (8)$$

where $P(\text{label})$ is the prior probability of a given label, each $P(f | \text{label})$ is the contribution of a single feature to the likelihood of the label, it can be calculated as:

$$P(f | \text{label}) = \text{count}(f, \text{label}) / \text{count}(\text{label}) \quad (9)$$

Calculate $P(\text{label} | \text{features}) = P(\text{features, label}) / P(\text{features})$, select the label corresponding to the maximum probability value as the label result for the new input value.

When the training set has features that never appear with a given label, the $P(f | \text{label})$ calculated value is 0, which would cause the input to never be assigned to this label, resulting in a reduced classification accuracy. To avoid such situations, we typically apply the "Laplace corrections" for "smooth" processing when estimating probability values. Let N represent the number of possible categories in the training set D , N_i represents the possible number of values for the i 'th attribute, D_c represents a collection of class c samples composed in the training set D , $D_{c_i x_i}$ represents a collection of samples with value x_i on the i 'th attribute in D_c , the correction calculation is as follows:

$$\begin{aligned} \text{count}(\text{label}) &= \frac{|D_c| + 1}{|D| + N} \\ \text{count}(f, \text{label}) &= \frac{|D_{c_i x_i}| + 1}{|D_c| + N_i} \end{aligned} \quad (10)$$

2.4. Logistic Regression Algorithm

Logistic regression is a study of dichotomous variable Y and a series of influencing factors X_n , the multivariate analysis method of the relationship is further developed on the basis of the linear model. Its general form is:

$$p = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n, P = \frac{1}{1+e^{-p}}. \text{ In the formula, } P \text{ is a probability}$$

occurrence of the variable Y , between 0-1. Logistic regression model has fast calculation speed, obvious results and good fitting effect. It is widely used in big data, machine learning, economics and other fields.

2.5. Neural network algorithm

The BP neural network is a multi-layer feedforward network trained for propagation from the direction of error. The BP neural network maps the input variable to the output variable via the excitation function and the constant weight threshold for adjusting each layer connection. To align the output variables of the network with the expectation, the learning of BP neural networks generally requires repeated training multiple times so that the error value tends to zero and eventually reaches zero.

The implementation process of the BP neural network algorithm is as follows: The input matrix of the sample is assumed to be $X = (x_{ij})$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, p$. Each row of data represents a set of input samples, and each set corresponds to a set of output samples, and then the actual output samples corresponding to all input samples are $Y = [y_1, y_2, \dots, y_n]^T$. We regard each column as an indicator of the sample, and then the input sample I_1 of the input layer is X .

If the implied layer of the network contains m neurons, weights matrix $W = (w_{ij})$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, p$, threshold matrix $B = [b_1, \dots, b_m]^T$, then the input of the implied layer is:

$$I_2 = W_{m \times p} X_{p \times n} + \text{Bones}_{1 \times n} \quad [17] \quad (11)$$

In which $\text{ones}_{1 \times n}$ represents a matrix with all elements of 1. The expression of the implied layer incentive function is:

$$f(x) = \frac{1}{1+e^{-x}} \quad (12)$$

Thus we can get the output of the implied layer as $O_2 = f(I_2)$, the input for the output layer is $I_3 = X_{jk} O_2 + B_{jk} + \text{ones}_{1 \times n}$. Since the transfer function is a class of linear functions, the $O_3 = I_3$ can be considered for the output of the output layer.

The initial weight threshold selected by the BP algorithm can determine the weight threshold from the second to the n adjustment weight threshold, so the accuracy of the BP network training is determined by the selection of the initial weight threshold. If the initial weight threshold is inappropriate, it is likely to make the network slow convergence and easily fall into the local optimal solution, so the selection of the initial power threshold has a great relationship to the quality of the BP network training results.

Genetic Algorithm (GA) is a heuristic global search algorithm (Three-dimensional coordinate measurement algorithm by optimizing BP neural network based on GA) inspired by the evolutionary processes of living organisms. Genetic algorithm simulates the problems to be solved into a process of biological evolution, generating the next generation of solutions through selection, crossover and variation. By calculating the fitness of individuals, the individuals with low fitness are eliminated, the individuals with high fitness are increased, and the optimal fitness function value solution is found after n iterations. The characteristics of genetic algorithms can compensate for the defect that BP neural networks can easily fall into local optimal solutions, so usually we use genetic algorithms to optimize BP neural networks to build correlation models.

3. Scene application

3.1. Equipment fault prediction

We use Pearson correlation analysis, single-factor variance analysis and other methods to conduct correlation analysis of various types of data and determine the main factors affecting the equipment failure. Then, we establish a Logistic regression fault prediction identification model to achieve an advance warning of the device state. Finally, we establish a classified neural network model to judge the possible fault types of the early warning equipment, supporting the differentiated operation and maintenance of the substation equipment.

We get the data from SCADA system, PMS2.0 system, intelligent inspection robot system, equipment online monitoring system, weather forecast system, etc. We take data from a prefecture-level company from January 2019 to August 2020, obtaining data from five dimensions, including operating conditions, equipment attributes, system mode, external environment, and seasonal characteristics, with a total of 135 fields and 179311 data. The specific data sources are listed in the Table 1.

Table 1. Data required for equipment failure prediction

Data category	Data source	Data field
operating conditions	SCADA system	Equipment current, Voltage, Oil temperature, Switch knife gate position, Load rate, Load value,etc.(Class 34 data)
equipment attributes	PMS2.0 system	Equipment name, Equipment coding, Equipment type, Equipment model, Manufacturer, Leave the factory date, Date of delivery, Voltage grade,etc.(Class 39 data)
system mode	intelligent inspection robot system	Appearance of equipment, Equipment separate/merged status, Electric energy meter measure instruction, Infrared temperature measurement, Noise,etc.(Class 28 data)
external environment	equipment online monitoring system	Transformer, Online monitoring system for the lightning arrester, Online temperature and humidity monitoring system, Infrared imaging online analysis system,etc.(Class 22 data)
seasonal characteristics	weather forecast system	Regional ID, Maximum forecast temperature, actual highest temperature, Minimum forecast temperature, Actual minimum temperature, Temperature and humidity, Thunderstorm, Hail,etc.(Class 12 data)

Equipment fault influencing data can be divided into continuous, classification and unstructured types. According to different data types, Pearson correlation analysis and single-factor variance analysis etc. are used to analyze the fault influencing factors.

The operating influence factors of equipment failure mainly extracted "the average monthly load rate" as the characteristic variables. We calculate the correlation coefficient value of 0.9317 through the correlation analysis, so there is a strong positive correlation between the device failure rate and the load rate. At the same time, we analyzed a strong positive correlation between the equipment failure rate and the operation life, temperature and humidity, with the correlation coefficient values of 0.7484, 0.617 and 0.602, respectively. The experimental results are shown in Figure 3, Figure 4 and Figure 5.

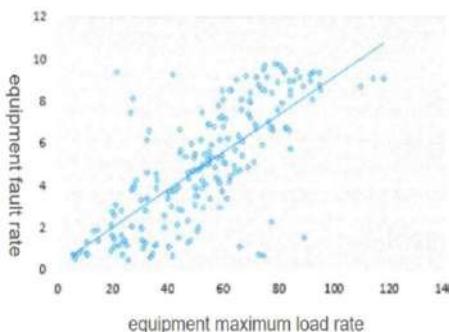


Figure 3. Monthly average load rate and fault rate scatter plot

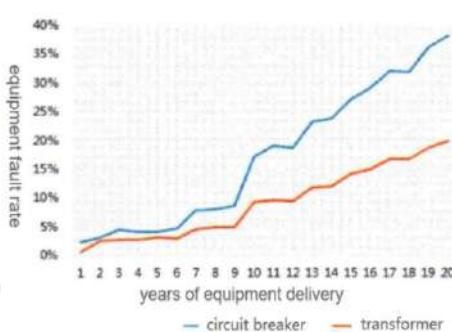


Figure 4. Correlation between equipment years of operation and failure rate

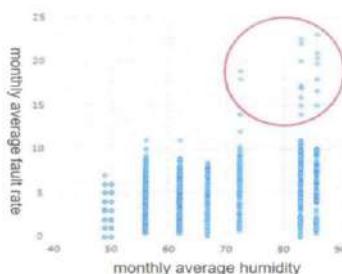


Figure 5. Relationship between equipment failure rate and temperature/humidity

We perform a variance analysis for the "region where the device belongs" variables. According to the variance analysis table, calculating $P=0.0387 < 0.05$, we can assume that the "region where the device belongs" has a significant impact on the equipment failure rate. Using the same method, we analyzed that the equipment model number has a significant influence on the equipment failure rate as shown in Table 2.

Table 2. ANOVA of the area where the equipment belongs

Error source	Quadratic sum	Free degree	Mean square error	F-value	P-value	F-critical value
Intergroup	5826.434784	12	485.536	4.0938	0.03876	3.9705
Within the group	20497992.24	143819	142.526	-	-	-
Total	20503818.68	143831	-	-	-	-

Based on the "time (month)", "manufacturer", "precipitation", "thunderstorm", "gale" and other classification or unstructural factors, we use the statistical comparison method, the examples are as follows. It can be seen in Figure 6 and Figure 7 that different manufacturers and weather conditions will affect the improvement of equipment failure rate.

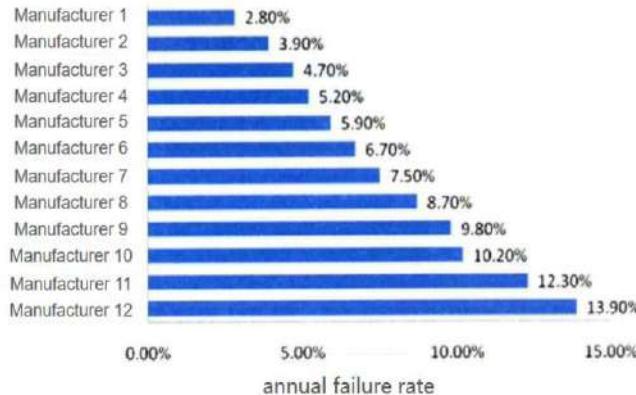


Figure 6. Equipment failure rate of different equipment manufacturers

We included 143,832 records for full year 2019 as training samples and the first quarter 2020 as test data. We used the Logistic regression model to obtain the training results with different probabilities. From the Table 3, it works best when the classification critical probability is 0.5.

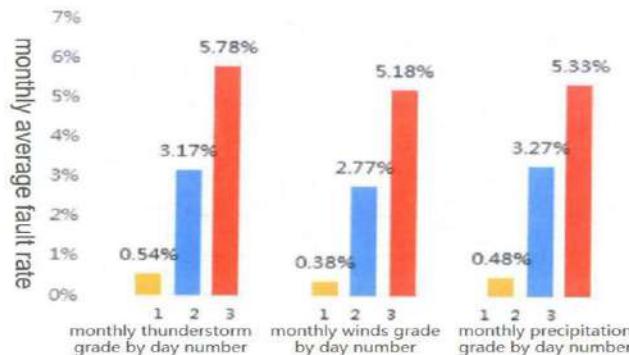


Figure 7. The relationship between substation failure and thunderstorm gale, precipitation, etc.

Table 3. Training results at different probabilities p

Classification critical probability-p	0.3	0.4	0.45	0.5	0.55	0.6	0.7	0.8	0.9
Identification accuracy	77.86%	82.39%	90.11%	94.23%	89.55%	82.17%	79.40%	77.22%	69.93%
Fault shooting rate	64.31%	73.42%	76.36%	82.53%	84.21%	85.39%	87.28%	88.01%	87.37%

Based on the regression model constructed above, the samples from April-May 2020 were predicted, as shown in the Table 4. The total prediction accuracy was approximately 90.45% with high accuracy.

Table 4. Analysis of the fault prediction error of the substation

Forecast month	Actual number of no failure has occurred	Actual number of failure has occurred	0-0	0-1	1-0	1-1	Predictive accuracy
April, 2020	9813	549	9054	759	52	497	92.17%
May, 2020	10519	609	9351	1168	96	513	88.64%

We performed fault identification through a neural network method with an identification accuracy of 93.16% and the results shown in the Table 5. At the same time, we can calculate that variables such as load rate, operation years of equipment, temperature and humidity, equipment model play an important role in the prediction process.

Table 5. Identify results

Transformer No.	Forecast Fault	Actual Failure	Whether identification is accurate
Transformer 1	The oil temperature is too high	The oil temperature is too high	Yes
Transformer 2	overload	overload	Yes
Transformer 3	The gas content of the oil exceeds the standard	The gas content of the oil exceeds the standard	Yes
Transformer 4	The oil level is abnormal	The oil level is abnormal	Yes
...

In internal management, we often use the influencing factors of equipment failure prediction to carry out equipment defects management, and the purpose is to reduce the time cost of labor input and increase work efficiency through intelligent means.

3.2. Equipment defect management

The main classification algorithm is used in the construction of the power grid production work order text analysis system, and its functional framework is shown in Figure 8. The keyword library management module extracts production keywords based on defect grading standards, realizes the original keyword mining function, and provides keyword query at the same time. The basic data classification standard library and the original work order library are based on the keyword library, and the classification rules are composed of keywords, and training is performed based on the rules and tags, forming an automatic classification after the work order is entered. The classification knowledge base and statistical analysis report function are based on the keyword extraction and classification functions to establish classification rules and various production business relationships to support more business applications. The classified work orders are used as basic data according to different businesses. Perform report statistics to realize different functional applications based on grid business.

As one of the important contents of power equipment operation and maintenance, defect management has always been a work of great concern for power enterprise managers and production and maintenance personnel. In actual work, there are irregularities in defect management, especially the non-standard description and handling of defects, which affect the level of defect management. The description content

of the equipment defect is the closest to the site situation and can reflect the specific situation of the site. Therefore, the word segmentation processing of the defect description text can obtain the keywords of the specific situation of the site. Perform word frequency analysis based on the processed keywords and classification criteria, use the TF-IDF algorithm to calculate the frequency of words in this category and the frequency of words in other categories, generate work order keywords, and build a keyword library. The principle is shown in Figure 9. Based on the keyword database, a sample data set of classification rules is established, and the classification rules are obtained by means of machine learning and Bayesian model training.

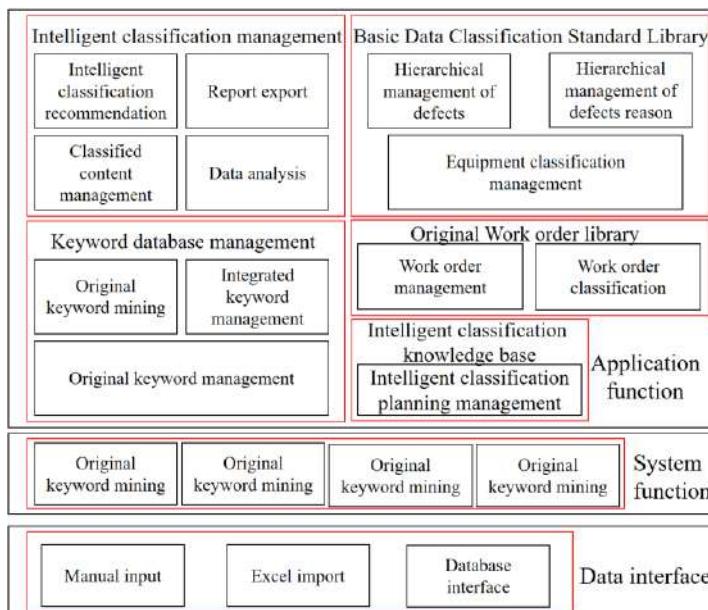


Figure 8. The functional framework of the text analysis system for power grid production work orders

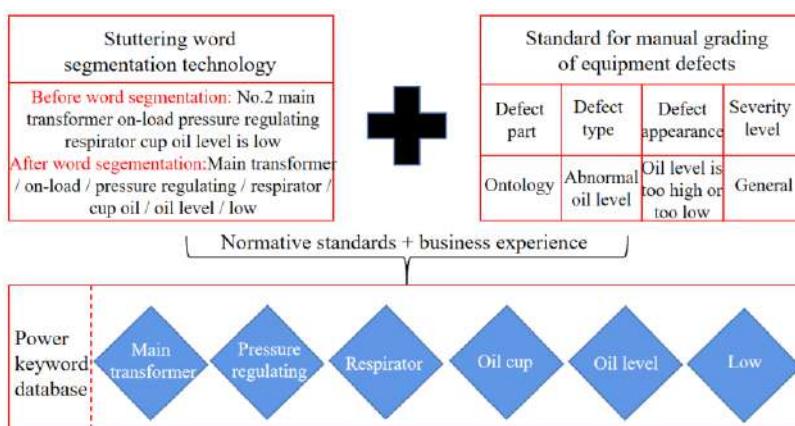


Figure 9. Keyword library construction principle

Based on the defect data, the defect location, defect category, and defect cause are classified respectively, where the defect locations are divided into first, second, and third levels, the defect causes are divided into first and second levels, and algorithms can be used for iterative calculations. Figure 10 shows an example of the whole process of automatic classification. The defect type is used as the target classification content. The input value is "The oil level of the No. 2 main transformer on-load pressure regulator respirator is low", and the output type value is "Abnormal oil level".

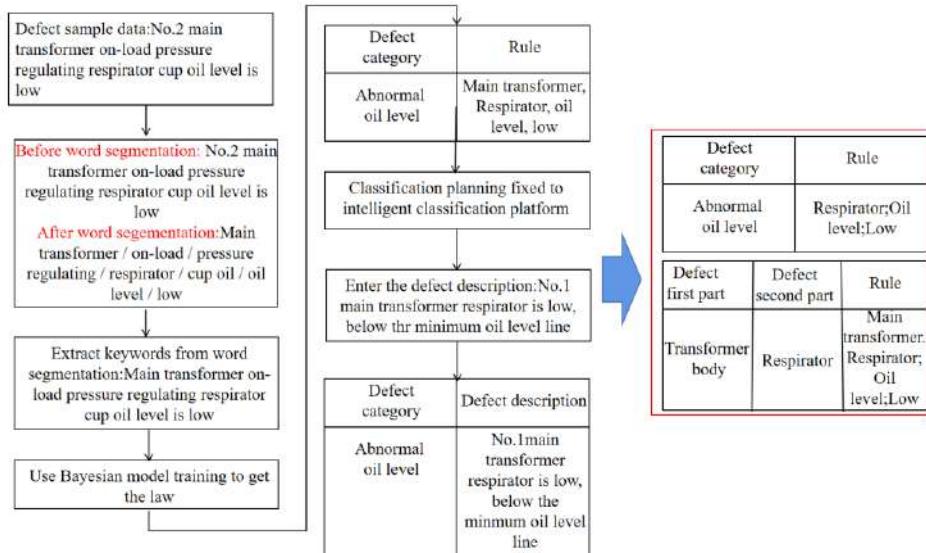


Figure 10. Example of the whole process of classification

4. Conclusion

The power grid production work order text analysis system applies text word segmentation, keyword mining, and automatic classification technology to each module to solidify the keyword database and cleaning rules for cleaning the production data of the primary equipment of the power grid. This method was applied to the data cleaning of primary production equipment for a power grid enterprise in the past seven years, and the total number of defects of equipment with a voltage of 110kV or above was 36255. The classification results are shown in Table 6, where there are 21,220 pieces of error correction data, accounting for 58.52%; 31,857 pieces of error correction data, accounting for 87.87%; data correction accuracy rates are between 93%-95%, achieving good results.

Table 6. Intelligent classification effect

	Number before correction/piece	Accuracy	Number after correction/piece	Accuracy
Defects	4398	12.13%	31857	93.53%
Defect category	15035	41.48%	21220	94.61%

The text analysis system of power grid production work orders ranges from a general data quality management platform that can only find problems to the system from

discovery to problem solving. It combines classification rules and standard formulation to provide data problem solutions.

At present, the power industry has not established a standardized power word database, which is in its infancy. Therefore, the establishment of this system provides a basis and development direction for grid companies to establish a set of standardized power word databases in the future. On the basis of realizing the automatic data cleaning function, the system has improved data quality and clear efficiency. It has functions such as intelligent classification of work orders, construction and optimization of keyword databases, and statistical analysis assistance. It can be used for system application construction departments, equipment management departments and various business research departments.

In the future, it is planned to apply more algorithms to the system to optimize the functions of each module, and at the same time increase the business expansion of the application part, such as increasing the research on the distribution network data, the intelligent classification of equipment accident event data, etc., to support more and higher power grid business analysis to promote the high-quality development of ubiquitous power Internet of Things business.

References

- [1] Bird, S. & Klein, E. & Loper,E. Python natural language processing. POSTS & TELECOM PRESS, 2009.
- [2] Chen, Y. J. Prediction of equipment defects in Power System. Zhejiang University, Hang Zhou, Zhejiang Province, China, 2003.
- [3] Kaneriya, S., Tanwar, S., Nayyar, A., Verma, J. P., Tyagi, S., Kumar, N., ... & Rodrigues, J. J.. Data consumption-aware load forecasting scheme for smart grid systems. IEEE Globecom Workshops (GC Wkshps) , 2018, 1-6.
- [4] Dang, F. F. Research on business data quality control technology of Power Grid Enterprises. North China Electric Power University, Beijing, China, 2014.
- [5] Du, X. M. & Qin, J. F. & Guo, S. Y. Text Mining of typical fault cases of power equipment. High Voltage Engineering, 2018, 44(4): 1078-1084.
- [6] Ju, X. L. Power marketing analysis system based on data warehouse technology.Digital Technology and Application, 2012(3): 68-69.
- [7] Li, Y. H. & Wang, J. X. & Wang, X. L. Power system network loss evaluation method based on hybrid cluster analysis.Automation of Electric Power Systems, 2016(1): 60-65.
- [8] Ma, R. & Zhou, X. & Peng, Z. Data Mining of correlation chatacteristics of load characteristics statistical indexes considering temperature factors .Proceedings of The Chinese Society for Electrical Engineering, 2015(1).
- [9] Tarik A. Rashid , Arazo M. Mustafa and Ari M. Saeed. A Robust Categorization System for Kurdish Sorani Text Documents. Information technology Journal, 2017(1): 27-34.
- [10] Wang, X. Research on Power Internet of things Architecture for Smart Grid Construction . Power Systems and Big Data, 2018, 21(10): 34-37.
- [11] Zhao Y. D., Wu Z., Chen X. F. et.al. Application of the big data method for electricity bill recovery risk prediction.Telcomunication Science, 2019, 35(02):125-133.
- [12] Chen X. F., Zhao Y.D., Zhang L. P. et.al. Application of big data methods for power potential sensitive customers.Telcomunication Science, 2020,35(11):117-124.
- [13] Wu, G. Y. & Zhang, Q. B. & Wu, H. C. Text Mining Analysis of power customer complaints based on natural language processing technology.Power Systems and Big Data, 2018(10).
- [14] Zhang, Y. X. & Hu, C. C. & Hang, S. Data Mining and Analysis method of secondary equipment defects based on Apriori Algorithm. Automation of Electric Power Systems, 2017(19).
- [15] Zou, B. P. & Zhou, L. Research on data service of Power Grid Enterprises. Electric Power Information and Communication Technology, 2010(9): 30-32.
- [16] Zou, Y. F. & He, W. M. & Zhao, H. Y. Application of Text Mining Technology in power work order data analysis . Modern Electronics Technique, 2016,39(17): 149-152.
- [17] Hang, L. Statistical learning methods. TSINGHUA PRESS, 2012.

Modeling Ebola Virus Dynamics by Colored Petri Nets

Dmitry A. ZAITSEV^{a,1} and Peyman GHAFFARI^b and Virginia Sanz SANCHEZ^c

^a*University of Information Technology and Management in Rzeszów, Poland*

^b*Department of Mathematics, University of Aveiro, Portugal*

^c*Faculty of Pharmaceutical Sciences, University of Iceland, Iceland*

Abstract. We develop techniques to translate deterministic cellular automata models into colored Petri nets based on an example of known cellular automata for modeling and mimicking Ebola virus dynamics. Cellular automata for Ebola virus dynamics use parametric specifications of rules that is peculiarity of the present study. The simulation results completely coincide with known results obtained via dedicated simulator. Having uniform language for models specification brings in advantages for models mutual transformations and simulation.

Keywords. Ebola virus, immune response, cellular automata, colored Petri net, simulation

1. Introduction

Ebola is one of the most lethal infectious diseases known to mankind. The Ebola virus causes an acute, often fatal hemorrhagic illness. The incubation time is something between two days and three weeks after becoming infected with the virus [1].

Cellular automata (CA) are widely applied for modeling processes of virus spreading both on micro-level of cells of some organ [2] to macro-level of geo-spatial representation [3].

A Petri net is a well-known formalism for specification of concurrent processes [4]. A family of Petri net classes offers a uniform language for modeling in various domains with possibility of applying both model-checking and simulation techniques of analysis.

In the present paper, we develop a technique for transformation of cellular automaton models for virus spreading into colored Petri net (CPN) models using a known cellular automaton that specifies Ebola virus dynamics [2].

2. Cellular automata and Petri nets

A cellular automaton [4] represents a homogenous structure of elements (cells) having rather simple behavior such that their collective behavior represents a mass parallel

¹ Corresponding Author, Dmitry A. Zaitsev, Information Technology Department, University of Information Technology and Management in Rzeszów, ul. Sucharskiego 2, 35-225 Rzeszow, Poland; E-mail: daze@acm.org.

process capable of modeling processes in many application domains. A state of cell in the next tact of time depends on the cell current state and states of its neighbors. Usually, an infinite square lattice of cells in multidimensional space is considered in theoretical research though for practical applications, a finite part of the lattice is simulated. Even one dimensional elementary (having binary state) cellular automata are Turing-complete [5] that illustrates power of the concept. Traditionally, von Neumann (sparse) and Moore (dense) neighborhoods of a cell are considered though recently a generalized neighborhood, that fills the gap between two mentioned neighborhoods within multidimensional lattices, has been introduced and studied [6], [7].

A Petri net [8] represents a bipartite directed graph with a dynamic process introduced on it. The first part of vertices, depicted as circles or ovals and called places, represents conditions. The second part of vertices, depicted as rectangles or bars and called transitions, represents events. Dynamic elements, called tokens and depicted as dots, are situated within places; they are consumed and produced as the results of firing transitions. Within a conventional Petri net, tokens are elementary; the net state is specified by the place marking represented by a vector which components contain the number of tokens in the corresponding places. A conventional Petri net is applied for modeling concurrent processes within a wide variety of application domains. Comparable simplicity of Petri net concept allows us to apply analytical methods for model checking [6] though its descriptive power is not too high for detailed specification of systems and processes.

A colored Petri net [9], [10] represents a combination of a Petri net graph and a functional programming language ML. Tokens are considered as elements of data types traditionally called color sets. Arcs and transitions are inscribed with ML constructs. Besides, time characteristics are introduced for the processes simulation. Supplied with a range of random distribution functions, CPN Tools [10] allows us to provide statistical analysis of timed and other characteristics. For composition of complex models, hierarchical design is provided via the transition substitution when a transition is substituted by a subnet. CPNs find wide application for the performance evaluation of systems, especially networks, grids, and clouds [11]. For this purpose, classical simulation technique is applied refined with special measuring components of models to compute statistical characteristics on-the-fly directly in the process of simulation [9]. For comparably small models, application of state space technique for the model properties analysis is possible.

3. Cellular automata models of Ebola virus dynamics

We use Ebola virus dynamics model [2]. A two dimensional CA with the cell states within the set $\{0,1,2,3\}$ is considered where 0 represents a healthy state, 1 and 2 represent an infected cell, and 3 represents a depleted (dead) cell. Here we present a model for one of three rules studied in [2] that explains basic principles of the transformation.

Let us consider the mentioned rules description where asterisk symbol “*” denotes any state and a matrix represents states of a cell (in the center) and its neighbors according to Moore neighborhood [6]. Before a matrix, a rule symbol is written; the rule application results in a new state of the central cell that is written after the equality symbol.

CA Q with rule q specification follows:

$$q \begin{pmatrix} * & * & * \\ * & 0 & * \\ * & * & * \end{pmatrix} = \begin{cases} 1, & \text{at least one } * \text{ is 1 or 2} \\ 0, & \text{otherwise} \end{cases}$$

$$q \begin{pmatrix} * & * & * \\ * & a & * \\ * & * & * \end{pmatrix} = 3, \text{ for } 1 \leq a \leq 2$$

$$q \begin{pmatrix} * & * & * \\ * & 3 & * \\ * & * & * \end{pmatrix} = 0.$$

The three above lines specify all possible four states of a cell within the set $\{0,1,2,3\}$: first line – 0, second line – 1 or 2, and third line – 3. Note that, in second and third line (for the current cell state 1, 2, or 3), the next state does not depend on neighbors. The first line can be considered as a variant of totalistic rule with check that the sum of all neighbor cell states is equal to or greater than 1.

CA Q specifies fast propagation of decease when, according to the first line of its description, at least one deceased cell within Moore neighborhood causes the current cell got infected. Together with other rules studied in [2], it serves for the subsequent composition of stochastic automata that is beyond the scope of the present paper.

4. CPN model of a cell

We compose colored Petri net model of a cell functioning under rule q shown in Fig. 1. In colored Petri net (CPN) of modeling system CPN Tools [10], [11] places are depicted as ovals; each place has its name, written inside a place, data type of tokens (color set), and initial marking. A token is not elementary in CPN; in the general case, it is an element of abstract data types such as cortege, record, list, or union composed of elementary types, such as integer or real numbers, strings and others. In the present work, we use only color sets: an integer number (INT), and a timed integer number (tint). Basic declarations look rather simple:

```
colset tint = INT timed;
var i,ii,j1,j2,j3,j4,j5,j6,j7,j8:INT;
var c:tint;
```

The reserved word “colset” serves for defining new sets of colors while the reserved word “var” serves for declaration of variables. A color set is specified using the equality sign “=”, a new color set tint is defined on basic integer color set INT using modifier “timed”. Definition of variables consists of a list of names separated by semicolon “;” symbol with the name of color set.

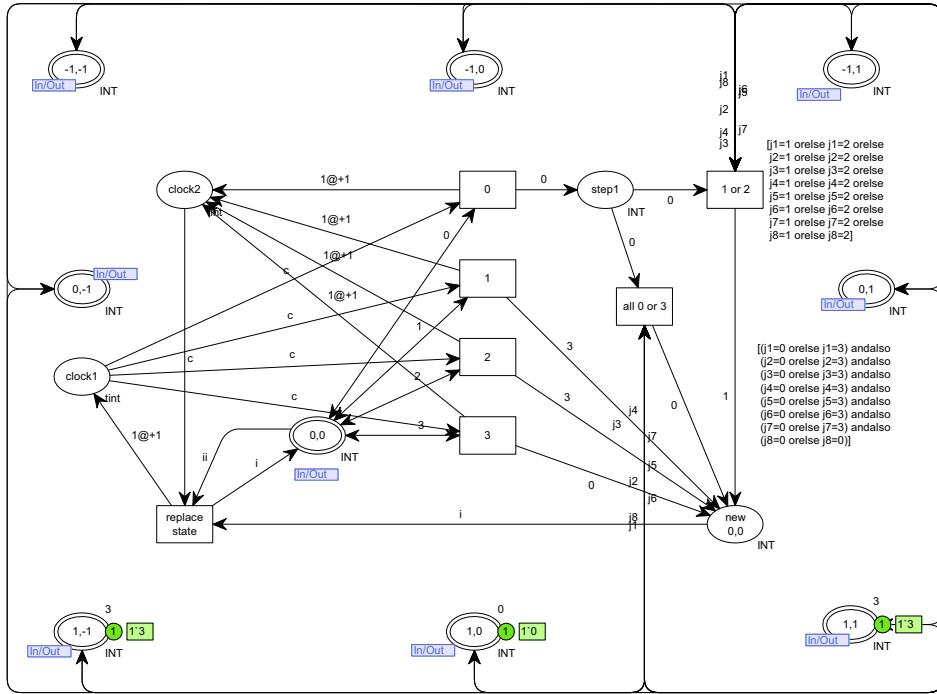
Within our model, we have only no more than one token inside each place though in the general case the place marking is represented by a multiset composed using double plus sign “++”. Recall that, each element belongs to a multiset in definite number of copies (multiplicity), the multiset element is represented as $k'e$, where k denotes multiplicity and e denotes an element. On the initial marking CPN Tools

creates current marking of simulation system which is displayed in green colors and changes as result of transitions firing.

Timed delays are associated either with transitions or with separate arcs. For timed multisets, an element is represented as $k`e@t$, where t is the token activation time. A transition fires instantaneously, increasing the step counter Step value, a delayed token spends time before its activation in the corresponding output place of transition in passive state not possible for enabling any transition. The timed delays are specified as $k`e@+d$, where d species the delay. Here we use integer numbers to specify delay though in the general case, a random distribution function can be applied, CPN Tools providing a series of random distribution functions including Gauss, Poisson and others.

In our models, we use a timed CPN to specify synchronous work of all the cells. The marking change within a lattice is implemented in two acts. At the first act, all the neighbors and the current cell itself can read the current “0,0” cell state using two directed arcs without change of the marking. During this act, a new state of the current cell is computed and stored within additional place “new 0,0”. Then, at the second act, the current state of all cells is instantaneously replaced by their new states. For this purpose, we have a clock token $1@t$ moving within a circle created by places $clock1$ and $clock2$ and transitions which connect the mentioned places. The clock extracted by variable c inscribed on the arcs from place $clock1$ activates one of the four alternative transitions having names “0”, “1”, “2”, and “3”, corresponding to the current “0,0” cell state. The corresponding alternative values of the current state are inscribed on the transitions incoming arcs. We use double directed arcs though one of the mentioned transition fire only once because of the clock arc; after firing a transition, the clock token is delayed by “ $1@+1$ ” to give time for all cell to compute their new state.

The automaton Q specification in the Section 2, contains simple subrules for the cell states equal to 1, 2, and 3: transform 1 or 2 into 3 and transform 3 into 0 without taking into consideration state of the neighboring cells. Each of these transformations is specified by a corresponding transition, putting the new current state value into place “new 0,0”. For the current cell state equal to 0, the automaton Q , according to its first subrule, checks states of its neighbors. There are two alternatives of this check represented by transitions “1 or 2” and “all 0 or 3”. Now let us pay attention to the naming of cells within the cell model. Because the model should suit to each cell of the lattice, we do not use the cell index; instead of it we use neighbor offsets with regard to the current state index. Thus, the current cell is named as “0,0”, the left upper cell is named as “-1,-1” and so on because we use a matrix-like orientation of coordinate axes. To extract values of the neighboring cells, we use a set of integer variables $j1, j2, j3, j4, j5, j6, j7, j8$ corresponding to the clockwise enumeration of neighbors starting from the upper neighbor. Note that, to avoid a tangled picture with manifold crossed lines, we implement connections of either of transitions “1 or 2” and “all 0 or 3” with neighboring cells as a bus merging all arcs in a single line; the bus encircles the cell shape with endings of arcs coming to the neighbor cells states.

Figure 1. Model of cell with rule q .

To check the state of neighbors, we employ a guard function of a transition. The guard function represents a logic expression that combines check of characteristics for tokens extracted by various incoming arcs of a transition. The logic expression is written in square brackets using the following notations: “orelse” denotes disjunction and “andalso” denotes conjunction. Thus we interpret the first variant of the first subrule for automaton Q , implemented by transition “1 or 2” as

$$(j_1 == 1) \vee (j_1 == 2) \vee (j_2 == 1) \vee (j_2 == 2) \vee (j_3 == 1) \vee (j_3 == 2) \vee (j_4 == 1) \vee (j_4 == 2) \vee (j_5 == 1) \vee (j_5 == 2) \vee (j_6 == 1) \vee (j_6 == 2) \vee (j_7 == 1) \vee (j_7 == 2) \vee (j_8 == 1) \vee (j_8 == 2).$$

We interpret the second variant of the first subrule for automaton Q as the above expression negation with regard to the state value range, implemented by transition “all 0 or 3” as

$$\begin{aligned} & ((j_1 == 0) \vee (j_1 == 3)) \wedge ((j_2 == 0) \vee (j_2 == 3)) \wedge ((j_3 == 0) \vee (j_3 == 3)) \\ & \quad \wedge ((j_4 == 0) \vee (j_4 == 3)) \wedge ((j_5 == 0) \vee (j_5 == 3)) \\ & \quad \wedge ((j_6 == 0) \vee (j_6 == 3)) \wedge ((j_7 == 0) \vee (j_7 == 3)) \\ & \quad \wedge ((j_8 == 0) \vee (j_8 == 3)). \end{aligned}$$

For the further composition of a lattice using the designed cell model, places corresponding to the current and neighboring cell states are labeled by the input/output

(In/Out) port tag drawn in blue color. They represent parameters of the model and will be studied in detail when composing a lattice.

5. Composing CPN model of a lattice

To compose a lattice of CA Q , we create a separate place to store the corresponding cell state, we use cell indices as names, for instance “0,0”, “0,1” etc. With each cell we associate a cell model represented by a separate transition, to which name prefix “c” is added, to obtain the main page of CA Q model, shown in Fig. 2. The arcs reflect connections with neighboring cells; for Moore neighborhood applied in the present model connections look rather dense. Note that, because the current cell state is not actually localized within the cell model, we have actually 9 connections for each cell represented by bidirected arcs. Neighbors only read the current cell state values while the current cell model changes it.

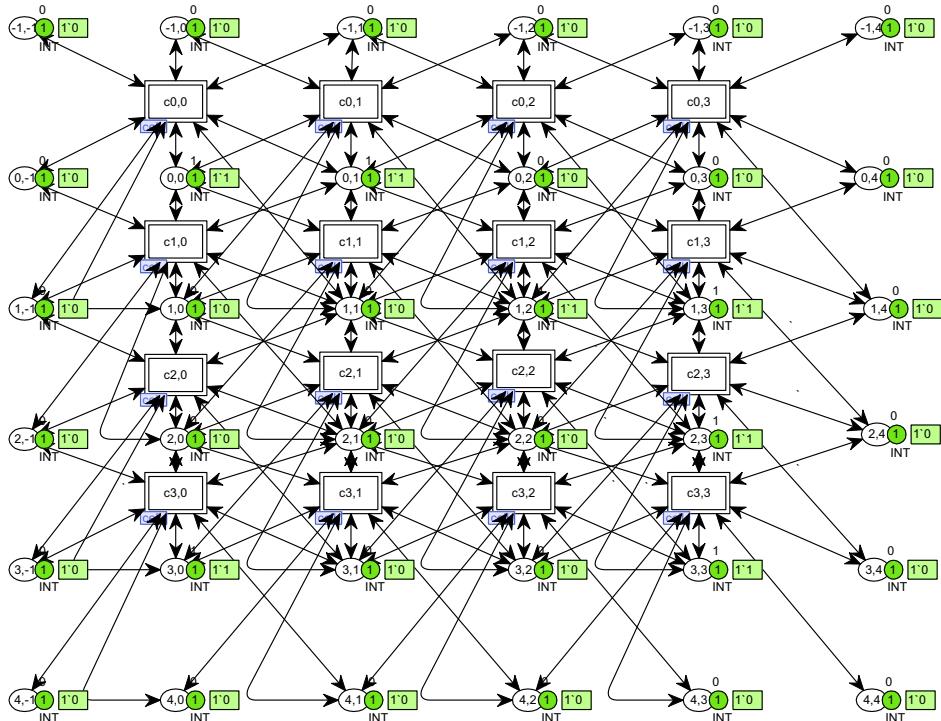


Figure 2. Model of CA Q (with cell model in Fig. 1), finite 4x4 lattice with constant border conditions.

As a mathematical model, CAs are thought of as infinite lattices and this definition is applied for deriving analytical results with regard to CA properties. For computer simulations, we always use a finite part of lattice obtained according to given size of CA, usually square for a two dimensional automaton. Thus, a question regarding the borders of finite lattice arise because if we simply cut a square finite lattice, cells on the borders will use indices of nonexistent cells to implement their rules. The simplest

solution, implemented in our model of lattice (Fig. 2), consists in adding places storing states of neighbors for one more layer of cells on each border; it corresponds to constant border conditions, though individual for each border cell. Thus in Fig. 2, actually cell state of 5x5 lattice is stored while cell state of 4x4 lattice is modified according to the rule q. Note that we can implement other border conditions, for instance, connect opposite edges to obtain a torus [12] or use additional subnets which implement varying border conditions.

Now let us consider in detail how hierarchical models are composed within CPN Tools. The basic operation is substitution of a transition by a subnet. In Fig. 2, each transition, which implements modification of the corresponding cell state, is tagged by the transition substitution tag “cell”. It means that the subnet shown in Fig. 1 works inside each transition with specific data according to the cell neighborhood state. To use parameters while substituting a transition, socket-port mapping is provided by CPN Tools [11]. Places of a submodel, which map on the corresponding places of a model, are called ports and indicated by port tags, for instance In/Out tags in Fig. 1. Ports are mapped to places of the model, which are called sockets, using special tools of Hierarchy palette of CPN Tools. It is supposed that transition substitution and port-socket mapping are implemented manually.

When the CA Q model is composed, we can simulate CA dynamics for various input state of lattice and border conditions. For this purpose CPN Tools offers step-by-step mode of simulation for model debugging. Then simulation of specified number of steps can be launched. It is convenient to use time break-points to check the lattice state when all the cells implemented change of state at the current timed tact.

6. Analysis of simulation results

Specifying here Ebola virus dynamics model in CPN language, we apply mainly simulation to study the model behavior and properties. Besides, though it is feasible for rather small models without using high performance resources, we can build complete state space of model, obtain automatic report on it containing conclusions of such basic properties as boundedness, liveness, and fairness [7], [11]. There is an additional possibility to compose requests to state space for finding specific properties which can be interesting, for instance, when modeling viruses. For 1x1 CA Q shown in Fig. 3, CPN Tools built state space drawn in Fig. 4.

In Fig. 3, we renamed vertices because state space palette of tools has more strict requirements to syntax: we replaced zero by “O” and one by “I” and sign minis by “m” to have names starting from a letter and containing no special characters. Thus, we have a cell with state 1 that has 7 neighbors with state 0 and one neighbor with state 1. Built state space of CA Q 1x1 (Fig. 3) represented in Fig. 4 contains 9 states; after an initial part of first two transitions, the models falls in a loop of states from 3 to 9. Initially drawn state space does not reflect detailed descriptions of states and transitions; firing a transition is represented by an arc connecting state, a state is specified by its number and the number of incoming and outgoing arcs. Fig. 4 contains attributes of all transitions expanded, as well as expanded state description for state 7. Transition specification indicates the transition name and values of tokens which were taken at transition firing (in curly brackets). Compound names of vertices are written including a prefix with the subnet name followed by, separated by quotation symbol, the node name. State specification lists marking of all transitions. The obtained report

concludes that the net is bounded, its initial marking is a home marking, dead markings are absent though some instances of transitions are dead, for example “cell'all0or3”, and classifies transitions into impartial, fair, just and unfair, no unfair transitions found.

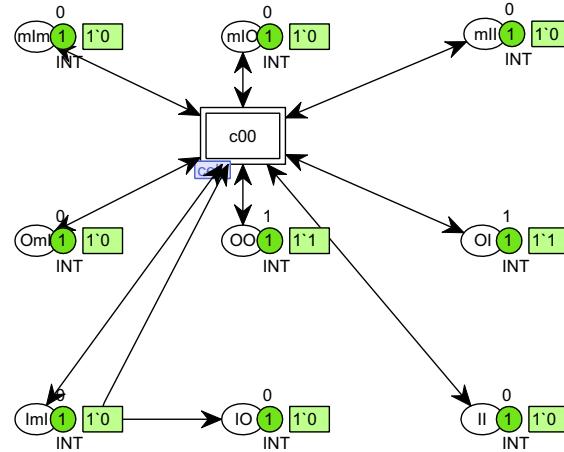


Figure 3. CA Q 1x1.

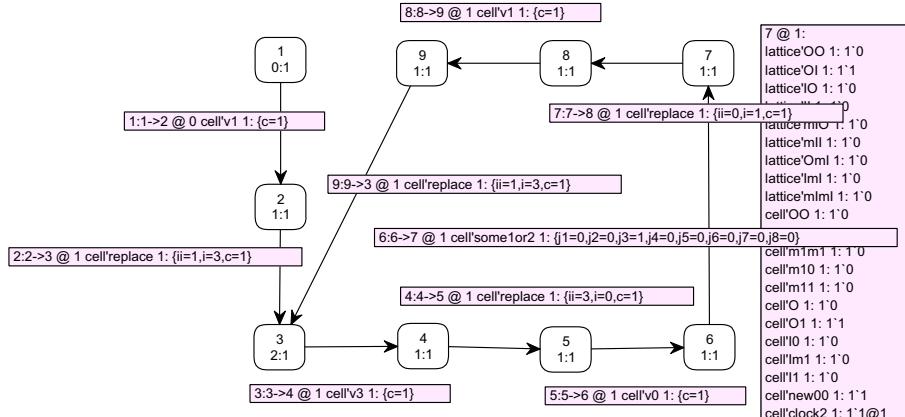


Figure 4. State space of CA Q 1x1 (Fig. 3).

Thus, the CA Q 1x1 model with the specified initial state, coming from CA state 1, falls into a loop of CA states 3->0->1->3 stored in place “cell’OO”. With our model, we observed other variants of cyclic behavior of CA studied in [2] for 3x3 and 7x7 lattices that indirectly acknowledges correctness of our constructs.

For statistical analysis of composed models having big size of the cell matrix, simulation on prolonged intervals of time is applied that also acknowledges conclusions of [2] with regard to parameters choice to provide recovery after the decease spreading. Combining state space and simulation techniques allows us to study the constructed models comprehensively that is an advantage of applying a colored Petri net and modeling system CPN Tools.

7. Conclusions

In this paper we present a technique to transform so called SEIR models used in mathematical Epidemiology into colored Petri net models. Applying CPN models for specifying both behavior patterns of individual cells and virus dynamics within real-life scale organs results in a uniform concept for model-checking and simulation.

Acknowledgement

We would like to thank the EU COST Association for bringing the authors together through funding events and STSM-grants within the framework of IMAAC (COST Action CA 16227).

References

- [1] Rachah, Amira; Torres, Delfim F. M. Dynamics and Optimal Control of Ebola Transmission, *Mathematics in Computer Science*: Vol. 10, pp. 331-342 (2016).
- [2] Emily Burkhead, Jane Hawkins, A cellular automata model of Ebola virus dynamics, *Physica A: Statistical Mechanics and its Applications*, Volume 438, 2015, Pages 424-435.
- [3] Gerardo Ortiz, Fred Brauer, Iris Neri Modelling and simulating Chikungunya spread with an unstructured triangular cellular automata, *Infectious Disease Modelling* 5 (2020) 197-220.
- [4] Jarkko Kari, Theory of cellular automata: A survey, *Theoretical Computer Science*, 334(1–3), 2005, 3-33.
- [5] Matthew Cook, Universality in Elementary Cellular Automata, *Complex Systems*, 15(1), 2004, 1-40.
- [6] Zaitsev D.A. A generalized neighborhood for cellular automata, *Theoretical Computer Science*, 666 (2017), 21-35.
- [7] D. A. Zaitsev, T. R. Shmeleva and P. Ghaffari, "Modeling Multidimensional Communication Lattices with Moore Neighborhood by Infinite Petri Nets," 2021 International Conference on Information and Digital Technologies (IDT), 2021, pp. 171-181, doi: 10.1109/IDT52577.2021.9497552
- [8] Reisig, W., *Understanding Petri nets. Modeling techniques, analysis methods, case studies*, Springer, 2013.
- [9] Zaitsev D.A. and Shmeleva T.R. *Modeling With Colored Petri Nets: Specification, Verification, and Performance Evaluation of Systems* (pp. 378-404) Chapter 14 in T. Shmelova, N. Rizun, D. Kucherov and K. Dergachov (Ed.) *Automated Systems in the Aviation and Aerospace Industries*. IGI-Global: USA, 2019.
- [10] Jensen, Kurt, Kristensen, Lars M. *Coloured Petri Nets. Modelling and Validation of Concurrent Systems*, Springer, 2009.
- [11] Zaitsev D.A. *Clans of Petri Nets: Verification of protocols and performance evaluation of networks*, LAP LAMBERT Academic Publishing, 2013, 292 p.
- [12] Dmitry A. Zaitsev, Tatiana R. Shmeleva & Birgit Proll, Spatial specification of hypertorus interconnect by infinite and reenterable coloured Petri nets, *International Journal of Parallel, Emergent and Distributed Systems*, 2021, DOI: 10.1080/17445760.2021.1952580

Image Segmentation by Pairwise Nearest Neighbor Using Mumford-Shah Model

Nilima SHAH^{a,b}, Dhanesh PATEL^a and Pasi FRÄNTI^{b,1}

^a*The M.S. University of Baroda, Vadodara, Gujarat, India*

^b*University of Eastern Finland (UEF), Joensuu, Finland*

Abstract. Mumford-Shah model has been used for image segmentation by considering both homogeneity and the shape of the segments jointly. It has been previously optimized by complex mathematical optimization methods like Douglas-Rachford, and a faster but sub-optimal k-means. However, they both suffer from fragmentation caused by non-convex segments. In this paper, we present hierarchical algorithm called Pairwise nearest neighbor (PNN) to optimize the Mumford-Shah model. The merge-based strategy utilizing the connectivity of the pixels prevents isolated fragments to be formed, and in this way, reaches better quality in case of images containing complex shapes.

Keywords. Image segmentation, Mumford-Shah, agglomerative clustering, PNN

1. Introduction

Image segmentation is a fundamental pre-processing task applied in more complex machine vision applications. The goal is to find regions of interest in an image that possess homogeneity and spatial connectivity. Let function $h(x, y) \in R^3$ represent RGB color of a pixel in the image at location (x, y) . We aim to find k segments $\Omega_1, \Omega_2, \dots, \Omega_k$ so that the pixel colors or some other feature would have smooth variation within the segments but discontinuous over the segment borders. We characterize the problem in terms of approximation theory, as the problem of finding an approximation f of h so that each f_i represents the segment Ω_i via a piecewise smooth function.

Existing approaches to image segmentation [1] include thresholding, clustering and classification methods [2], [3], [4], region-based methods [5], and edge based active contour methods [6], [7], [8]. The segmentation problem has also been dealt as an energy minimization problem. Obtaining a generalized solution that works without prior knowledge about the objects, their characteristics like shape, color, texture, appearance of shadows and overlapping of objects, is still an open problem.

Clustering has also been applied to image segmentation via grouping the pixels by minimizing intra segment similarity using k-means algorithm [9]. K-means is known to be sensitive to initialization, but it can be improved significantly by better initialization technique and by repeating the algorithm 100 times [10]; or by using random swap algorithm [11] which practically never gets stuck to an inferior local minimum.

¹ Corresponding Author, Pasi FRÄNTI, University of Eastern Finland (UEF), Joensuu, Finland; E-mail: pasi.franti@uef.fi

K-means and random swap minimize sum-of-squared error (*sse*) between the pixels and centroids:

$$sse = \frac{1}{n} \sum_i \|f - h\|^2 \quad (1)$$

However, minimizing *sse* leads to color quantization where every pixel is mapped to its nearest color regardless of its neighbors resulting in a fragmented segmentation. This is typically solved by adding spatial constraint to the equation or by using explicit convexity criterion in the algorithm.

Mumford-Shah is an energy minimization model that has been applied for image segmentation [12]. Its simplified version is defined as follows:

$$E(f, \tau) = \int \int_{\Omega} \|f - h\|^2 dx dy + \lambda L(\tau) \quad (2)$$

where the first term equals to *sse*, $L(\tau)$ denotes the total length of the segment boundaries, and λ is a control parameter. Large values of λ will produce shorter boundaries whereas smaller values of λ will put more emphasis on the homogeneity of the segments. The Mumford-Shah model has been extensively used in the areas of image denoising, segmentation and restoration [13].

Algorithmic techniques to minimize the Mumford-Shah includes Douglas-Rachford algorithm [14], implicit methods [15], and k-means [16]. The last one is virtually as fast as standard k-means and significantly faster than the other optimization techniques using Mumford-Shah model [15], [17], [18], [19], [20], [21]. However, the algorithm itself does not control the convexity and it can still lead to fragmentation with isolated sub-segments.

In this paper, we propose hierarchical merge-based clustering algorithm for optimizing Mumford-Shah model. It allows only neighbor segments to be merged, and in this way, prevents isolated sub-segments to be formed. Classical agglomerative clustering itself is a slow algorithm, but we show that utilizing the 2-D neighborhood we can reach $O(n^2)$ time complexity by rather straightforward implementation.

2. Pairwise nearest neighbor using Mumford-Shah

Agglomerative clustering [22], [23] is a popular alternative to k-means but a simple implementation can be several orders of magnitude slower. Initially, each pixel forms its own segment s_i . Two segments are then repeatedly merged until the desired number of segments is reached. The segment pair to be merged are selected using the Ward's method [24] so that the increase in *sse* is the least among all possible merges. Thus, it generates clusters by a sequence of locally optimal merge operations.

2.1. Pairwise nearest neighbor method

In the context of vector quantization, agglomerative clustering is known as the pairwise nearest neighbor (PNN) method due to Equitz [25]. Many variants of PNN exist. Most of them try to speed up the algorithm while some also try to improve the clustering quality. We next give a brief literature review of the agglomerative clustering.

A straightforward implementation of the PNN algorithm takes $O(n^3)$ time because there can be n steps in total, and every step takes $O(n^2)$ time. Even if the merge costs were stored in a matrix, the search for the best merge would still require exhaustive

search [26]. In [27] an alternative variant called fast exact PNN was proposed to reduce the time complexity to $O(\tau n^2)$, where τ is significantly smaller than n with all realistic data sets. The main idea is to maintain a nearest neighbor to store the lowest merge cost for each segment to avoid repeating unnecessary calculations [25]. The method is exact in a sense that it does not compromise the accuracy of the PNN.

In [28], an improved variant called Lazy-PNN was developed. The running time of the algorithm is dominated by $O(nk)$ time operations [25], [29]. However, some of the merge cost calculations can be delayed and a remarkable number of calculations can be avoided completely. Practical tests indicated that the lazy-PNN is about 35% times faster than the fast exact PNN [30], and 200-500 times faster than the original PNN [23], [30].

In [31], the fast exact PNN was also applied to multilevel nonparametric image thresholding, achieving $O(n \log n)$ but this does not generalize to segmentation in RGB space. In [32], the algorithm uses a heap data structure in which all pairwise merge costs are stored and the smallest merge cost is always found as the element of the root node of the heap. The computation time of the algorithm is at most $O(n^2 \log n)$ at the cost of $O(n^2)$ memory consumption, which is not suitable for segmenting large images.

In [33], a more general approach called iterative shrinking is proposed. Instead of merging, the number of clusters is reduced by a sequence of cluster removal operations. Clusters are removed one at a time by reassigning the objects from the removed cluster to the nearby clusters. The PNN method can be considered as a special case of the iterative shrinking, in which the objects of the removed cluster are all forced to move to the same cluster selected for the merge [25].

In [34], an approximate k-nearest neighbor graph is used for reducing the number of merge cost calculations. The graph is utilized so that the search for the cluster with the smallest merge cost is limited only to the clusters that are directly connected by the graph. This reduces the time complexity from $O(n)$ to $O(k)$ for a single node. The graph construction becomes then the bottleneck but fast approximate variants using k-d tree, divide-and-conquer or projection-based search were considered in [35]. The time complexity of the algorithm can be improved accordingly from $O(\tau n^2)$ to $O(\tau n \log n)$ at the cost of slight increase in *sse* [22], [35]. All the above variants of agglomerative clustering aim at faster speed except the iterative shrinking which aims at better quality.

2.2. Adopting Mumford-Shah to PNN

In this paper, we adopt the PNN variant from [34] to image segmentation using Mumford-Shah model. The main difference is that we take spatial connectivity of the pixels into account instead of merely minimizing *sse*.

The basic structure of the PNN method is shown in Algorithm 1. Given a set of n pixels $X = \{x_1, x_2, \dots, x_n\}$, the method starts by assigning each pixel x_i to its own segment represented by centroid c_i . In each step, the number of segments is reduced by merging two segments a and b . The merge cost is calculated as follows:

$$cost_{i,nb} = \frac{n_i n_{nb}}{n_i + n_{nb}} \|c_i - c_{nb}\|^2 - \lambda \cdot 2 \cdot common_{i,nb} \quad (3)$$

The first term calculates how much *sse* is increased by the merge, and second term how much the boundary length is reduced. It can be calculated as follows:

$$common_{i,nb} = (blen_i + blen_{nb} - blen_{i+nb})/2 \quad (4)$$

where b_{len_i} is the boundary length of the segment i . The PNN method applies a local optimization strategy: all possible cluster pairs are considered and the one increasing the Mumford-Shah model according to (3) is always chosen:

$$a, b = \arg \min_{\substack{i, j \in [1, k] \\ i \neq j}} cost_{i,j} \quad (5)$$

Algorithm 1: Pairwise nearest neighbor

```

PNN( $X, k$ )  $\rightarrow S$ 
FOR  $i \leftarrow 1$  to  $n$  DO
   $S_i \rightarrow \{x_i\}$  ;
REPEAT
   $(S_a, S_b) \leftarrow \text{SearchNearestClusters}(S)$ 
  Merge( $S_a, S_b$ )
UNTIL  $|S| = k$ 

```

2.3. Implementation of Mumford-Shah PNN

The proposed algorithm follows the basic structure of the PNN method in [34] with two main differences. First, we use merge cost derived from Mumford-Shah model. Second, since we operate on image pixels, we can implement the search for the nearest neighbors more efficiently by utilizing the connectivity of the segments.

In the beginning, each pixel forms its own segment. The centroid of the segment is the pixel value as such, $c_i = x_i$, and the segment size is set to $csize_i = 1$. An array of its four neighbors is recorded (nb_i). For each segment, we also maintain its nearest neighbor (nn_i) and the corresponding merge cost value (nnc_i). These are found by calculating the merge cost (3) of the segment and with its all neighbors. We explain next how these data structures are updated during the merge process.

The next merge is always found by linear search among all the nearest neighbors. In case the merge cost of two or more neighbors are equal, we select the neighbor as per the row-major order. While the merging process, the algorithm considers only spatially connected neighbors. This prevents fragmentation to appear which happens with the k-means variants that only minimizes (3) without explicit connectivity constraint.

We maintain a heap structure of all the segments according to the descending order of nnc . After the merge, the centroid of the merged cluster is updated as:

$$c_i \leftarrow \frac{c_i \cdot csize_i + c_{nb_i} \cdot csize_{nn_i}}{csize_i + csize_{nn_i}} \quad (6)$$

The boundary length is also updated using (4), and the list of neighbors of cluster i is updated by selecting the unique neighbors of the two merged clusters avoiding duplicates. The nearest neighbor of the merged cluster i is then resolved by recalculating the merge cost to all its neighbors.

We reach slightly better time complexity than in [27], $O(n^2) < O(\tau n^2)$ although the difference is not huge because τ is data dependent variable which is usually small. In comparison to [34], our algorithm is still slower as $O(n \log n) < O(n^2)$. The main reason is that the original PNN does not need to maintain the partition whereas we need to update the pixel labels in order to calculate the boundary length. The Mumford-Shah PNN is slower than the Mumford-Shah k-means. This is the price we need to pay for the better segmentation quality.

3. Experiments

We compare the proposed method (MS-PNN) with two other methods: *regularized k-means* (reg-KM) [36], and *Mumford-Shah k-means* (MS-KM) [16]. The experiments were conducted on desktop with Intel Core i5 processor with 2.50GHz speed, 8 GB RAM, 64-bit Windows 10 operating system. To evaluate the quality of the segmentations we used two measures: *bi-directional consistency error (BCE*)* [37] and *structural similarity index (SSIM)* [38]. As test set, we use 11 images from the Weizmann dataset [37] which has human segmented ground truths.

The results are summarized in Table 1. They show that the proposed method outperforms the other methods by clear margin. The average BCE* values are 0.33 (MS-PNN), 0.38 (MS-KM), 0.46 (KM) and 0.47 Reg-KM. The SSIM values are best among 9 out 11 images. The better quality comes at the cost of higher processing time; 476s for MS-PM compared to, 166s (Reg-KM) and <1s (KM and MS-KM).

The visual quality in Fig. 1 shows the main benefit of MS-PNN; it is the only method without isolated fragments which is observed in all k-means variants. While k-means is reasonably good in optimizing the Mumford-Shah model, these isolated sub-segments are persistent and only the merge-based algorithm can avoid them completely.

The main limitation of the algorithm is the λ parameter of the Mumford-Shah model which needs to be manually tuned. Fig. 2 shows its effects on few sample images. Here we have used the values of $\lambda=0.25$ and $\lambda=0.50$.

Table 1. BCE* (the lower the better) and SSIM (the higher the better) for the three methods.

Image	BCE*				SSIM			
	KM	Reg-KM	MS-KM	MS-PNN	KM	Reg-KM	MS-KM	MS-PNN
1	0.22	0.65	0.54	0.05	0.77	0.57	0.82	0.96
2	0.73	0.72	0.44	0.46	0.43	0.43	0.54	0.74
3	0.53	0.57	0.42	0.51	0.52	0.52	0.63	0.72
4	0.05	0.03	0.31	0.03	0.95	0.95	0.81	0.95
5	0.18	0.23	0.24	0.08	0.80	0.80	0.82	0.94
6	0.76	0.65	0.52	0.62	0.59	0.59	0.70	0.56
7	0.44	0.20	0.20	0.16	0.61	0.77	0.78	0.89
8	0.57	0.58	0.23	0.21	0.42	0.42	0.83	0.82
9	0.44	0.58	0.41	0.33	0.57	0.57	0.72	0.80
10	0.61	0.53	0.50	0.79	0.46	0.46	0.46	0.60
11	0.54	-	0.35	0.38	0.41	-	0.50	0.66
Average	0.46	0.47	0.38	0.33	0.59	0.61	0.69	0.79

4. Conclusions

We have introduced a new approach by embedding the well-known Mumford-Shah model into the merge-based hierarchical PNN algorithm. The results of the proposed MS-PNN are much better than the k-means variants in terms of segmentation quality (BCE*) and reconstruction quality (SSIM) of the images. The resulting segment boundaries of MS-PNN are much smoother without fragmentation. The drawback of the algorithm is slower than the corresponding k-means variant using Mumford-Shah but this can be tolerated as much better segmentation quality is obtained. Furthermore, it is quite possible to speed-up the method further from $O(n^2)$ to $O(n \log n)$ by utilizing better data structures. This is our future work.

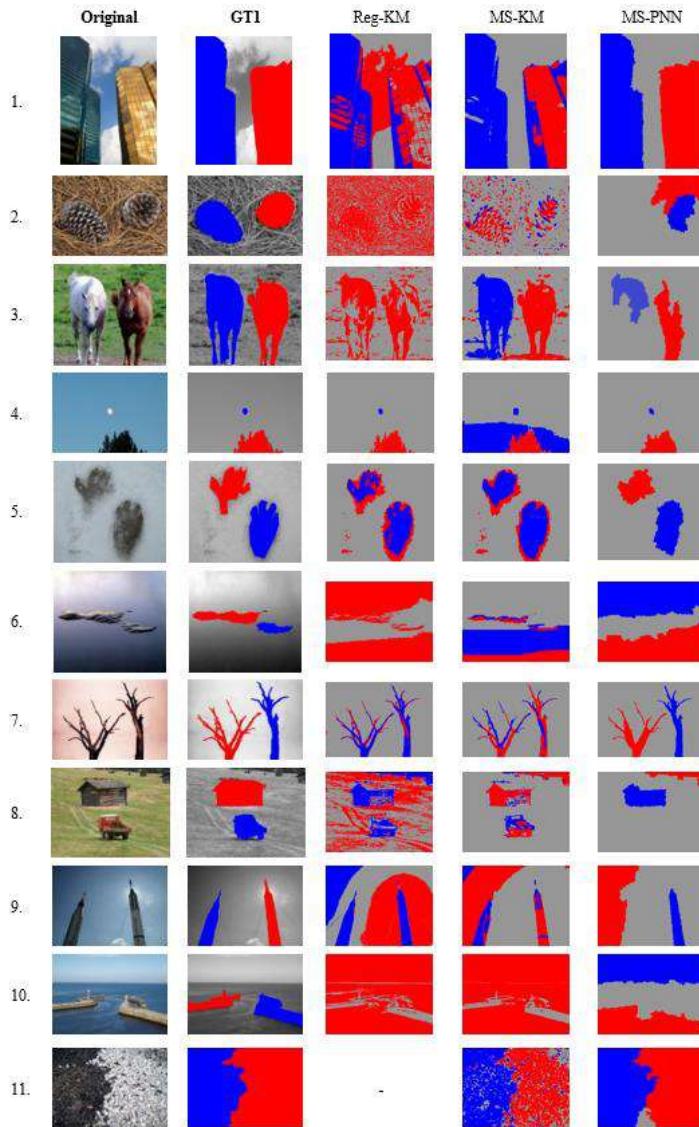
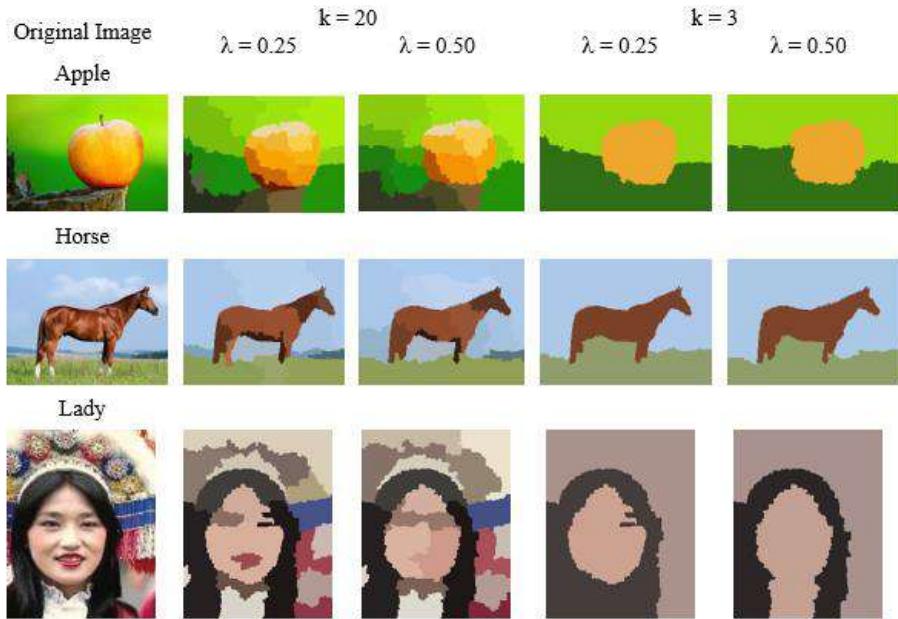


Figure 1. Visual comparison of the segmentation results using three methods.

Figure 2. Examples of the effect of λ on the segmentation result.

References

- [1] N. Shah, D. Patel and A. Jivani, "Review on image segmentation, clustering and boundary encoding," International Journal of Innovative Research in Science, Engineering and Technology, vol. 2, pp. 6309-6314, 2013.
- [2] Y. Yang and S. Huang, "Image segmentation by fuzzy c-means clustering algorithm with a novel penalty term," Computing and Informatics, vol. 26, pp. 17-31, 2007.
- [3] Y. Tolias and S. Panas, "Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions." IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 28, pp. 359-369, 1998.
- [4] F. Gibou and R. Fedkiw, "A Fast Hybrid k-Means Level Set Algorithm For Segmentation," 4th Annual Hawaii International Conference on Statistics and Mathematics, 8 2003.
- [5] J. Martin, "Implementing the region growing method Part 1: The piecewise constant case," 2002.
- [6] T. Chan and L. Vese, "An active contour model without edges," in Proc. Scale-space theories in computer vision. Scale-Space 1999, Lecture Notes in Computer Science, Springer, 1999.
- [7] T. Chan and L. Vese, "Active contours without edges," IEEE Transactions on Image Processing, vol. 10, pp. 266-277, 2 2001.
- [8] T. Chan and L. Vese, "A level set algorithm for minimizing the Mumford-Shah functional in image processing," in Proc. Variational and Level Set Methods in Computer Vision Workshop, 2001.
- [9] V. K. Dehuriya, S. K. Shrivastava and R. C. Jain, "Clustering of Image Data Set Using K-Means and Fuzzy K-Means Algorithms," in 2010 International Conference on Computational Intelligence and Communication Networks, 2010.
- [10] P. Fränti and S. Sieranoja, "How much can k-means be improved by using better initialization and repeats?," Pattern Recognition, vol. 93, pp. 95-112, 2019.
- [11] P. Fränti and S. Sieranoja, "K-means Properties on Six Clustering Benchmark Datasets," Applied Intelligence, vol. 48, pp. 4743-4759, 12 2018.
- [12] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," Communications on Pure and Applied Mathematics, vol. 42, pp. 577-685, 1989.
- [13] L. Bar, T. Chan, G. Chung, M. Jung, N. Kiryati, R. Mohieddine, N. Sochen and L. Vese, Mumford and Shah Model and its applications to image segmentation and image restoration, 2 ed., Springer Science+Business Media LLC, 2011, pp. 1097-1154.

- [14] N. Shah, D. Patel and P. Fränti, "Fast Mumford-Shah two phase image segmentation using proximal splitting scheme," *Journal of Applied Mathematics*, vol. 2021, 6618505, 13 pages, 2021.
- [15] L. Ambrosio and V. Tortorelli, "Approximation of functional depending on jumps by elliptic functional via t-convergence," *Communications on Pure and Applied Mathematics*, vol. 43, pp. 999-1036, 12 1990.
- [16] N. Shah, D. Patel and P. Fränti, "K-means image segmentation using Mumford-Shah model," *Journal of Electronic Imaging*, 2021. (submitted)
- [17] E. Brown, T. Chan and X. Bresson, "Convex formulation and exact global solutions for multi-phase piecewise constant Mumford-Shah image segmentation," 2009.
- [18] X. Cai, R. Chan and T. Zeng, "Image segmentation by convex approximation of the Mumford-Shah model," UCLA CAM Report, 2012.
- [19] W. Fleming and R. Rishel, "An integral formula for total gradient variation," *Arch. Math.*, vol. 11, pp. 218-222, 1960.
- [20] G. Koepfler, C. Lopez and J. Morel, "A multiscale algorithm for image segmentation by variational method," *SIAM Journal of Numerical Analysis*, vol. 31, pp. 282-299, 1994.
- [21] T. Pock, A. Chambolle, D. Cremers and H. Bischof, "A convex relaxation approach for computing minimal partitions," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 810-817, 2009.
- [22] S. W. Ra and J.-K. Kim, "A fast mean-distance-ordered partial codebook search algorithm for image vector quantization," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, pp. 576-579, 1993.
- [23] G. N. Lance and W. T. Williams, "A general theory of classificatory sorting strategies: 1. Hierarchical systems," *The computer journal*, vol. 9, pp. 373-380, 1967.
- [24] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, pp. 236-244, 1963.
- [25] W. H. Equitz, "A new vector quantization clustering algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 1568-1575, 1989.
- [26] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Springer, Boston, MA, 1992.
- [27] P. Fränti, T. Kaukoranta, D.-F. Shen and K.-S. Chang, "Fast and Memory Efficient Implementation of the Exact PNN," *IEEE Transactions on Image Processing*, vol. 9, pp. 773-777, 2000.
- [28] T. Kaukoranta, P. Fränti and O. Nevalainen, "Vector quantization by lazy pairwise nearest neighbor method," *Optical Engineering*, vol. 38, pp. 1862-1868, 1999.
- [29] D.-F. Shen and K.-S. Chang, "Fast PNN algorithm for design of VQ initial codebook," in *Visual Communications and Image Processing'98*, 1998.
- [30] P. Fränti and T. Kaukoranta, "Fast implementation of the optimal PNN method," in *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269)*, 1998.
- [31] O. Virmajoki and P. Fränti, "Fast pairwise nearest neighbor based algorithm for multilevel thresholding," *Journal of Electronic Imaging*, vol. 12, pp. 648-660, 2003.
- [32] T. Kurita, "An efficient agglomerative clustering algorithm using a heap," *Pattern Recognition*, vol. 24, pp. 205-209, 1991.
- [33] P. Fränti and O. Virmajoki, "Iterative shrinking method for clustering problems," *Pattern Recognition*, vol. 39, pp. 761-775, 2006.
- [34] P. Fränti, O. Virmajoki and V. Hautamaki, "Fast agglomerative clustering using a k-nearest neighbor graph," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, pp. 1875-1881, 2006.
- [35] O. Virmajoki and P. Fränti, "Divide-and-conquer algorithm for creating neighborhood graph for clustering," in *Proceedings of the 17th International Conference on Pattern Recognition. ICPR 2004.*, 2004.
- [36] S. Kang, B. Sandberg and A. Yip, "A regularized k-means and multiphase scale segmentation," 2011.
- [37] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A Database of Human Segmented Natural Images and its Application to evaluating seg algo and measuring ecological statistics," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*.
- [38] A. C. Brooks, X. Zhao and T. N. Pappas, "Structural Similarity Quality Metrics in a Coding Context: Exploring the Space of Realistic Distortions," *IEEE Transactions on Image Processing*, vol. 17, pp. 1261-1273, 2008.

I Want to Control Your Move: Human-Human Interface (HHI) Neuromuscular Electrical Stimulator (NMES)

Ching Yee YONG¹ and Terence Tien Lok SIA

School of Engineering and Technology, University College of Technology Sarawak, Sibu, Sarawak, Malaysia

Abstract. Neuromuscular electrical stimulation (NMES) has been widely used in rehabilitation hubs to restore or replace the motor function of individuals who have upper neuron damage such as stroke and spinal cord injury. However, the utilization of sensors in NMES is limited and results in the lack of data for upper limb movement analysis. The proposed system implemented NMES integrated with human-to-human interface (HHI) in the rehabilitation process for stroke patients. The therapist (controller) can coach the motion of patients (subject) by injecting his own signal for patients to follow. Ten (10) subjects were tested with five (5) repeating trials. The EMG value was extracted from the finger flexion and extension at the controller side, then injected into the control unit for further stimulation of the subject. In order to evaluate the repeating motion by the subject, an accelerometer was attached to the finger. Performance evaluation of the subject was executed by comparing the flexion angle with the controller side. The result showed that the error of the system was less than 10.29 % for the first trial and gradually reduced to 1 % after 5 trials.

Keywords. Accelerometer, EMG, human-human interface (HHI), motion, neuromuscular electrical stimulation (NMES), home-based rehabilitation

1. Introduction

The rehabilitation for stroke patients can help them relearn skills that were lost when a stroke affected a part of their brains. Normally, the patients will undergo rehabilitation exercises through physical therapy to strengthen the muscles [1]. Unlike the patients with significant locomotors impairment, the gait analysis was introduced as the therapist's clinical tool for analyzing the walking pattern of the patients. Normally, the physician carries out an observational gait assessment to identify the specific gait dysfunctions of patients through the practice routine [2]. The physician provides useful information and guides the patients in locomotors training by using video images or the naked eye in observing the whole gait analysis [3]. However, the observational gait assessment has lacked in studying the psychometric properties. This study [4] shows that physical therapists can give an accurate rating of gait deviation when the task is clearly defined and with a single parameter to judge. The observational gait assessment has limited study in psychometric properties. The physical therapists can give an

¹ Corresponding Author, Ching Yee YONG, School of Engineering and Technology, University College of Technology Sarawak, Sibu, Sarawak, Malaysia; E-mail: ching.yee@ucts.edu.my.

accurate rating of gait deviation when the task is clearly defined and with a single parameter to judge [5]. The home-based rehabilitation program allows the patients to perform NMES at home, thus help them to regain independence in performing daily activities. The result shows that the performance of the patient can be dependent on the physiotherapist by providing the performance feedback to correct the patients' error [6]. In sports science, the process of the motion assessment can be done by utilizing the diagnostic tools or sensors to provide measurable information in coaching the athletes in performing exercise programs [7]. The data has been an important factor in this era which may help to analyze, create new strategies and thus boost a player's performance. The real-time performance is targeted for monitoring the performance of an athlete which may avoid muscle fatigue [8].

NMES can be used as a motor relearning tool to let the patients participate in goal-oriented repetitive movement therapy [9]. This kind of improvement in upper limb function using the NMES device can be called a therapeutic effect. Moreover, the patients who are in the chronic phase of the stroke where the motor relearning strategies do not work against it [10]. NMES can be used as a neuroprosthesis with the primary intent to let the patients execute functional tasks with the affected upper limb while using the device as part of routine daily living. This improvement on the affected upper limb while using the NMES device is called the neuro-prosthetic effect [11].

Knutson et al. [12] proposed a method of electrical stimulation which is contralaterally controlled functional electrical stimulation (CCFES) to help improve recovery of volitional hand function for patients with hemiplegic stroke. This method provides the patient with the opportunity to control the stimulation of the paretic hand in real-time by opening and closing the non-paretic hand. On the non-paretic hand side, it is worn with an instrumented glove that is able to modulate the stimulation intensity to the paretic hand by detecting the degree of hand opening. Ambrosini et al. [13] proposed the system with a multi-joint robot and NMES hybrid system for the coordinated upper limb physical practice at the elbow, wrist, and fingers. The proposed system is a wearable device that supports the patients to perform sequencing limb movements such as elbow extension, wrist extension associated with hand open, wrist flexion, and elbow flexion.

Rong et al. [14] proposed the closed-loop control system which is an EMG-controlled neuroprosthesis to exploit maximally the user contribution during hand-to-mouth movement with the help of a passive exoskeleton for weight support. The intensity of the volitional muscle activity from EMG data is then measured at the stimulation period of 50 ms. The EMG signal is generated where further compared to the preset threshold to trigger the electrical stimulation. Cunningham proposed the use of contra-laterally controlled functional electrical stimulation (CCFES) to help to improve the recovery of volitional hand function for patients with hemiplegic stroke [15]. Cunningham proposed a closed-loop control system which is an EMG-controlled neuroprosthesis to exploit maximally the user's contribution during hand-to-mouth movement with the help of passive technology.

The stimulation intensity refers to the current where its amplitude can vary from 0 to 140 mA for some ES devices [6]. Muscle contraction required high intensity during ES where the maximal contraction is not focusing in most ES applications. This is because the ES only activates the small portion of motor units underneath the surface electrode which does not contribute much in producing maximal contraction. At this point, the level of discomfort will increase as increasing the intensity. Furthermore, the intensity required to achieve the desired contraction is varied among the patients

according to the individual sensation, skin resistance, size of muscle and size of electrodes. The contact of electrodes with the skin in different sessions will result in variation of the intensity provided. Nevertheless, the intensity may need to be increased during the session to adjust for accommodation.

The existing system for NMES only has a preset value of the stimulation parameter at which the patients cannot control the stimulation intensity. The proposed system integrated HHI as the communication channel to control the upper limb movement of patients. Furthermore, sensors are deployed in detecting flexion angle to control the stimulation intensity. Since the NMES is applied to the patient's skin, the safety of the patient is an essential consideration in designing the output circuit of NMES. Galvanic isolation was done for the H-bridge topology and transformer.

2. Materials and Methods

Figure 1 depicted the flow of the study. A pushbutton was included in the design in order to ensure continuous data uploading to the cloud after the WiFi module was activated. After this step was taken, the system will automatically perform the following such as getting the input from the controller and trigger the same strength to the subject's muscle. All the data was uploaded into the cloud for storage and further analysis by a professional practitioner. The subject can monitor his/her progress through the LED indicators.

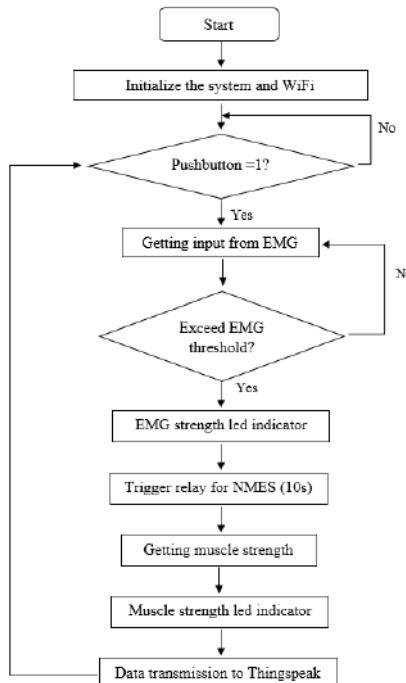


Figure 1. Methodology flow chart.

Ten (10) individuals (male and female) with the ages of 21-25 years were recruited in the study. The subjects are healthy and without joint issues in the upper limb and do

not encounter any significant soreness in the flexor carpi ulnaris during the participation.

The HHI required both the controller (physiotherapist) and the subject (patient) to complete the NMES process. On the controller side, two of the surface electrodes which are red and black for the EMG circuit were placed at the forearm and dorsum of the hand respectively. On the other hand, the surface electrodes were positioned at the forearm and attached at the top of the ulnar nerve which is directly connected to the little finger and the adjacent half of the ring finger. The end of the surface electrodes is connected to a signal stimulator to produce electrical stimulation for fingers' movement observation. Two (2) electrodes were used to deliver the pulses of electrical current. The accelerometer sensor was attached to the subject's finger with elastic bands to obtain the movement values. Figure 2 shows the proposed prototype unit.

The frequency and pulse width of the stimulation signals were set for constant at 120Hz and 110 μ s. Meanwhile, the level of intensity (amplitude of signal release to the subject) was set to 2 for the initial condition. The subject's hand was placed parallel to the table to obtain a 0° at the z-axis initially. 12 LEDs located at the top of the prototype were used to indicate muscle strength. 6 LEDs represent different strengths for the controller and subject respectively. The EMG value represents the muscle strength produced from the controller side while the accelerometer sensor values represent the muscle strength produced from the finger movement by the subject. The result will be displayed through 6 LEDs where each led represents a specific range of values.

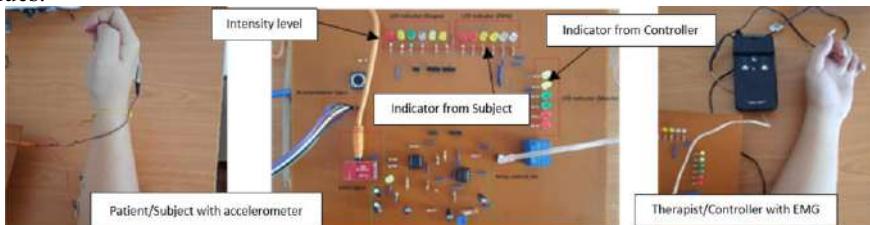


Figure 2. Controller set-up (left), HHI based NMES prototype (center), Subject set-up (right).

The EMG detects the electrical signals transmitted from the motor neuron when there are muscle contraction activities. This electrical signal is referred to as the analog signal, whereas it needs to be converted to a digital signal by using the analog to digital converter (ADC) for further analysis. The collected data is further used to determine signal strength from the controller's elbow movement and angle of elbow flexion produced during NMES. The proposed system examined the EMG signal from the controller elbow flexion and checked it with the EMG threshold. When the EMG threshold is exceeded, the relay is triggered ON to perform NMES with a duration of 10 seconds at the controlled person's hand. The muscle movement is detected using an accelerometer sensor with reading. The EMG signal level and muscle movement level for a controlled person hand are both shown by the led indicator. The proposed system collects four types of data which are EMG signal level, muscle movement level, EMG trigger value, and the muscle movement reading at the z-axis. The data is transmitted to the Thingspeak through a WiFi shield with a time of approximately 15 seconds per data.

3. Results and Discussion

Table 1 shows the experimental results of 10 subjects for the first trial. The process was repeated 5 times for every subject. When the intensity level was increased (stimulation signals), the error was reduced. The percentage error of each subject was computed by the formula as follow:

$$\text{error (\%)} = \frac{\text{EMG controller} - \text{EMG subject}}{\text{EMG controller}} \times 100\% \quad (1)$$

Table 1. Data collection and experimental results for the first trial

Subject	Intensity Level	Indicator (Controller)	Indicator (Subject)	Error (%)
		LED	EMG	
1	2	5	349	3.725
2	3	5	340	1.765
3	3	5	330	0.909
4	3	5	342	1.170
5	3	5	340	2.353
6	3	6	350	1.143
7	2	6	350	10.286
8	2	6	350	3.429
9	2	6	360	7.778
10	2	6	350	5.714
⋮	⋮	⋮	⋮	⋮

Figure 3 shows the mechanism of the proposed stimulator. The design was divided into LED and sensors parts. The patient can monitor their progress at home by observing the illumination of the LED lights. The progress can be categorized as satisfactory if the number of LED lit up by the subject is the same as the controller's.

The collected data from the sensors are purposely for a physiotherapist for progress analysis. The professional therapist can monitor and analyze the patient/user progress by analyzing the numerical values. The user may perform the rehab activity at home since all the data will be uploaded online in real-time automatically. The therapist can retrieve the data at anytime, anywhere through a PC or a mobile phone. The main goal of this project is to design a portable IoT gadget for rehabilitation purposes. The error was between 0.91% and 10.29% at the first trial. This error can be related to the skin impedance, fat tissue of an individual or misplacement of an electrode at the stimulation area. The error was gradually reduced after several trials.

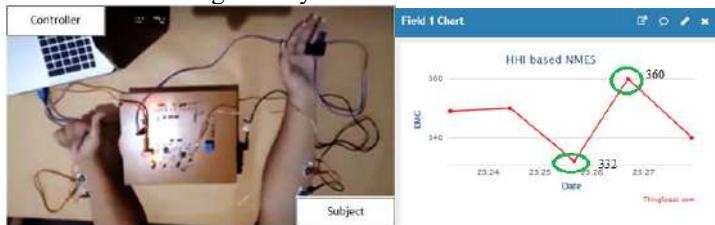


Figure 3. The controller is injecting the motion signal to the NMES (left), the IoT platform.

4. Conclusion and Future Work

The proposed system can be used in the rehabilitation hub for physiotherapy sessions. The HHI required the physiotherapist to take the role of the controller while the patients are being coached for motion. Furthermore, the therapist can control the

stimulation timing through the push button on the board. The therapist and the patients can also monitor the stimulation result by observing the LED indicators and sensors values. The LED indicators will show the result based on the capability of the patients to perform the same action as the therapist does. Besides, the data of the stimulation progress is uploaded and stored in the cloud for further analysis. Thingspeak serves as the IoT platform where the data of each patient can be visualized and analysed in the cloud. The therapist can easily access this app through mobile or website to monitor the data.

Future work will cover the implementation of fuzzy logic using MATLAB to achieve digital control of triggering the electrical stimulation based on the EMG value change at the controller side. The EMG value, angle of finger flexion and extension are used as the input linguistic variable while the stimulation intensity is used as the output linguistic variable. Through this fuzzy logic model, the stimulation intensity is varied based on the rule set to achieve automation control for both controller and subject sides.

Ethical Clearance and Conflict of Interest

This study was reviewed ethically according to the code of ethical principles and standards. Written informed consent forms were sent out to the participants and parents for consent approval. The collected information has complied with Data Protection Regulations, and not disclosed to an external party. Standard operating procedures were followed in promoting the best possible ethical research practice. The authors declare no conflict of interest. Gratitude to the University College of Technology Sarawak.

References

- [1] Arip, Eza SM, Waidah I, Md JN, Abduljalil R. Virtual reality rehabilitation for stroke patients: recent review and research issues. In AIP Conference Proceedings. 1905(1): p. 050007. 2017.
- [2] Ridao-Fernández C, Pinero-Pinto E, Chamorro-Moriana G. Observational gait assessment scales in patients with walking disorders: systematic review. BioMed research international. 2019.
- [3] Yong CY, Sudirman R, Chew KM. Motion detection and analysis with four different detectors. Third International Conference on Computational Intelligence, Modelling & Simulation; 2011. p. 46-50.
- [4] Nadeau S, Martina B, Francois B. Gait analysis for poststroke rehabilitation: the relevance of biomechanical analysis and the impact of gait speed. Physical Medicine and Rehabilitation Clinics. 2013; 24(2): 265-276.
- [5] Naili JE, Anna-Clara E, Maura DI, Michael HS, Margaret H, Charlotte KH, Eva WB. The impact of symptomatic knee osteoarthritis on overall gait pattern deviations and its association with performance-based measures and patient-reported outcomes. The Knee. 2017; 24(3):536-546.
- [6] Stevens-Lapsley JE, Jaclyn EB, Pamela W, Donald GE, Robert SS, Margaret S, Wendy MK. Relationship between intensity of quadriceps muscle neuromuscular electrical stimulation and strength recovery after total knee arthroplasty. Physical therapy. 2012; 92(9):1187-1196.
- [7] Seshadri DR, Magliato S, Voss JE, Drummond C. Clinical translation of biomedical sensors for sports medicine. Journal of medical engineering & technology. 2019; 43(1):66-81.
- [8] Kim H, Jongho L, Jaehyo K. Electromyography-signal-based muscle fatigue assessment for knee rehabilitation monitoring systems. Biomedical engineering letters. 2018; 8(4):345-353.
- [9] Cano-De-La-Cuerda R, Molero-Sánchez A, Carratalá-Tejada M, Alguacil-Diego IM, Molina-Rueda F, Miangolarra-Page JC, Torricelli D. Theories and control models and motor learning: clinical applications in neurorehabilitation. Neurologia. 2015; 30(1):32-41.
- [10] Li S. Spasticity, motor recovery, and neural plasticity after stroke. Frontiers in neurology. 2017; 8(120).
- [11] Knutson J., Mary YH, Terri ZH, John C. Improving hand function in stroke survivors: a pilot study of contralaterally controlled functional electric stimulation in chronic hemiplegia. Archives of physical medicine and rehabilitation. 2007; 88(4):513-520.

- [12] Knutson JS, Michael JF, Lynne RS, John C. Neuromuscular electrical stimulation for motor restoration in hemiplegia. *Physical medicine and rehabilitation clinics of North America*. 2015; 26(4):729.
- [13] Ambrosini E, Simona F, Marta T, Thomas S, Christian K, Giancarlo F, Alessandra P. An EMG-controlled neuroprosthesis for daily upper limb support: a preliminary study. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2011 p. 4259-4262.
- [14] Wei R, Li W, Pang M, Hu J, Wei X, Yang B, Wai H, Zheng X, and Hu X. A Neuromuscular Electrical Stimulation (NMES) and robot hybrid system for multi-joint coordinated upper limb rehabilitation after stroke." *Journal of neuroengineering and rehabilitation*. 2017; 14(1):1-13.
- [15] Cunningham DA, Jayme SK, Vishwanath S, Kelsey AP, Andre GM, Ela BP. Bilateral Contralaterally Controlled Functional Electrical Stimulation Reveals New Insights Into the Interhemispheric Competition Model in Chronic Stroke. *Neurorehabilitation and neural repair*. 2019; 33(9): 707-717.

Analysis of Heart Rate Asymmetry During Sleep Stages

Chang YAN ^{a,1}, Peng LI ^b, Yang LI ^a, Jianqing LI ^a and Chengyu LIU ^{a,1}

^a*School of Instrument Science and Engineering, Southeast University, Nanjing, China*

^b*Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA*

Abstract. It is one of the hot spots in recent years to explore changes in the sleep stage by assessing autonomic nervous activity. In recent years, heart rate asymmetry (HRA) is often used to measure the activity of autonomic nerves. However, the relationship between HRA and sleep stage is not clear. We performed Porta's index (PI), Guzik's index (GI), slope index (SI) and area index (AI) analyses on RR intervals per 30-s for understanding the HRA during sleep. Two measurement protocols were set: 1) the HRA values were calculated; 2) the degrees of heart rate deviation from symmetry were estimated. Results showed that PI significantly decreased from N1 and N2 to N3 ($p<0.01$), and it is increased the highest in REM than other stages ($p<0.05$). The asymmetry of HRA were significantly lower in N3 (PI and AI p 's <0.05 ; GI and SI p 's <0.01), and it increased in REM (PI $p<0.05$; GI, SI and AI p 's <0.0001). The results suggested that HRA has the potential to be used in sleep stage monitoring.

Keywords. sleep, heart rate asymmetry, autonomic nervous system

1. Introduction

About one-third of a person's life is asleep, so quantifying the quality of sleep plays a vital role in human health. Sleep can be divided into 5 stages with 30s as the step size which named wake (W), N1, N2, N3 and rapid eye movement (REM) according to American Academy of Sleep Medicine.

A number of studies indicated that during the onset of sleep and transitions to different sleep states, the autonomic nervous system (ANS) has specific performance in different stages of sleep [1], [2]. During non-REM (NREM) sleep, that is combined with N1, N2 and N3, parasympathetic activity causes a decrease in heart rate, while sympathetic nerve activity results in heart rate acceleration in REM sleep [3]. Therefore, understanding the changes of ANS in different stages of sleep is helpful for accurately assess individual sleep quality. Heart rate variability (HRV), a non-invasive method, has been widely applied to detect the ANS modulation [4], [5], which describes the fluctuation of RR intervals from an electrocardiogram (ECG), including to understand ANS changes during different sleep stages.

Poincaré plot was used as a visualized tool for evaluating the increase and decrease of HRV based on the scatter diagram drawn from the current RR interval as abscissa

¹ Corresponding Authors, School of Instrument Science and Engineering, Southeast University, Nanjing, People's Republic of China; E-mail: cyan@seu.edu.cn (C. Y.); chengyu@seu.edu.cn (C. L.).

against the preceding one as ordinate in Cartesian coordinate [6]. Several metrics had been proposed to estimate the asymmetry of points besides the line of identity of Ponicare plot which were also named as heart rate asymmetry (HRA), such as Porta's index (PI) [7] assesses the asymmetry by points number, Guzik's index (GI) [8] estimates the asymmetry through points distance, slope index (SI) [9] quantify the asymmetry via phase angle and area index (AI) [10] calculated the asymmetry by the way of area of sector. It was reported that the severity of obstructive sleep apnea has a certain correlation with the oscillation pattern of heart rate acceleration and deceleration [11]. However, there is no relevant research on the relationship between HRA and sleep stage.

The purpose of the present study is to show the performance of heart rate asymmetry to analyze differences of HRV or ANS in different sleep stages. Besides, the values of HRA further deviates from 50 indicates the RR intervals are more asymmetry. Hence, the asymmetry of HRA in different sleep stages were also explored.

2. Data and methods

2.1. Data

The MIT-BIH Polysomnographic Database [12], [13] was used in this study, which consists of 18 recordings of polysomnographic signals during sleep collected from 16 male subjects (age:32-56, weight: 89-152 kg). The sampling frequency of all recordings including an ECG channel were 250 Hz. The R waves are annotated and the sleep stages were scored already. Each recording comes from a separate participant except that the files “slp01a” and “slp01b”, “slp02a” and “slp02b” were collected from the same person, respectively.

2.2. Algorithm of heart rate asymmetry

For a RR interval series $(RR_1, RR_2, \dots, RR_i, RR_{i+1}, \dots, RR_{n-1}, RR_n)$, figure 1 summarized the algorithm of the four HRA measures which can be achieved by the following formulas:

- Porta's index

PI measures the proportional to the number of points below LI [7]. The formula to calculate PI is

$$PI = \frac{b}{m} \times 100 \quad (1)$$

wherein, b and m present the number of points below LI and not located on LI, respectively. Similarly hereinafter.

- Guzik's index (GI)

GI was the ratio of the cumulative distance corresponding the points [8], which can be calculated by

$$GI = \frac{\sum_{i=1}^l D_i}{\sum_{i=1}^m D_i} \times 100 \quad (2)$$

wherein, D_i can be calculated as

$$D_i = \frac{|RR_{i+1} - RR_i|}{\sqrt{2}} \quad (3)$$

- Slope index (SI)

SI assesses the percentage of cumulative phase angle above LI [9]. The way to

calculate SI is

$$SI = \frac{\sum_{i=1}^l |R\theta_i|}{\sum_{i=1}^m |R\theta_i|} \times 100 \quad (4)$$

wherein, $R\theta_i = \theta_{LI} - \theta_i$, where θ_{LI} is the angle between LI and abscissa, and $\theta_i = \text{atan}(RR_{i+1}/RR_i)$.

- Area index (AI)

The area of the i th sector is denoted by

$$S_i = \frac{1}{2} \times R\theta_i \times r^2 \quad (5)$$

Then AI was defined as the ratio of S_i above LI [10]. The formula of AI is,

$$AI = \frac{\sum_{i=1}^l |S_i|}{\sum_{i=1}^m |S_i|} \times 100 \quad (6)$$

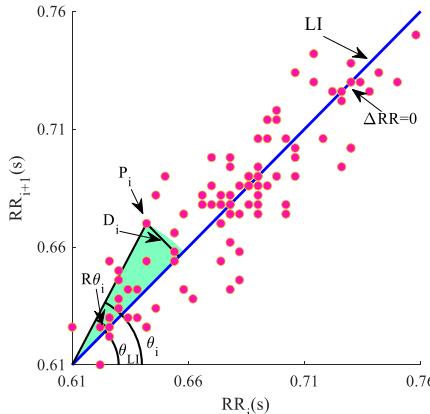


Figure 1. Schematic diagram of HRA algorithm in a Poincaré plot. Red dots show points in the Poincaré plot. LI is the line of identity. P_i is the i th point which can be denoted by (RR_i, RR_{i+1}) . $\Delta RR = RR_{i+1} - RR_i$. D_i represents the distance of P_i respect to LI; The angle between the line between P_i and the origin of the coordinate and the abscissa (and LI) is denoted by θ_i (and $R\theta_i$); θ_{LI} is the phase angle between the abscissa and LI.

2.3. Asymmetry of HRA metrics

HRA greater than or less than 50 indicates asymmetry. The farther away from the specific value, the higher the asymmetry. Therefore, the results of subtracting 50 from the HRA index value and then taking the absolute value were named HRA asymmetry, i.e.,

$$\text{Asymmetry of PI} = |PI - 50| \quad (7)$$

$$\text{Asymmetry of GI} = |GI - 50| \quad (8)$$

$$\text{Asymmetry of SI} = |SI - 50| \quad (9)$$

$$\text{Asymmetry of AI} = |AI - 50| \quad (10)$$

2.4. HRA analysis of RR intervals

The RR intervals with a duration of 30s were constructed with the annotated R waves according to the sleep stage label in the database. A RR interval was excluded if it was

less than 0.7 times the mean of the RR intervals or larger than 1.5 times mean of the RR intervals. In addition, SI and AI were impacted by the reference lines and reference points [14]. The minimum of every RR interval series was applied as the reference point. Two procedures were performed on the preprocessed RR intervals.

- 1) the HRA values were estimated from the RR intervals in different sleep stage;
- 2) the degrees of heart rate deviation from symmetry during sleep stage changes were assessed.

2.5. Statistical analysis

The values of different HRA belonging to the same sleep stage were averaged within each participant. Linear mixed-effect models were performed to examine the differences in the eight HRA measures across different sleep stages. When the alpha level is 0.05, statistical significance is accepted. Matlab software (Ver. R2021a, The MathWorks Inc., Natick, MA, USA) were used for all statistical analyses.

3. Results

The HRA results of RR intervals corresponding to different sleep stage were summarized as mean + 95% confidence interval in Figure 2. It was found that the asymmetry of heart rate variability was significantly decreased in N3 stage than the other two NREM sleep stages (Fig 2(a), $p<0.01$). Then PI in REM sleep was significantly difference with other sleep stage ($p<0.05$). GI, SI and AI showed no significant difference among the different sleep state ($p>0.05$).

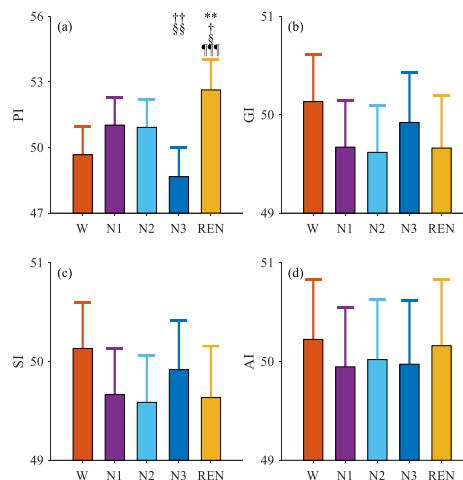


Figure 2. HRA results of RR intervals corresponding to different sleep stage. (a) PI; (b) GI; (c) SI; (d) AI. The error bar indicates the 95% confidence interval in each stage. *: W vs. the four sleep stages; †: N1 vs. N2, N3 and REM; §: N2 vs. N3 and REM; ¶: N3 vs. REM. The number of symbols means the degree of significant difference, e.g., †, §, ¶: $p<0.05$, **, ††, §§: $p<0.01$; ¶¶: $p<0.0001$.

The results of asymmetry of HRA assessed by PI, GI, SI, and AI calculated were summarized as mean + 95% confidence interval in Figure 3. There was a significant tendency of decreased in the asymmetry of RR intervals in the N3 sleep stage compared

with W, N1 and N2 (asymmetry of GI, SI and AI, p 's <0.05 , figure 3 b, c, d). Asymmetry of PI showed N3 stage showed no significant difference with N2 stage. All of the four measures suggested that the asymmetry of REM was higher than N3 sleep (PI, $p<0.05$; GI, SI and AI, $p<0.0001$).

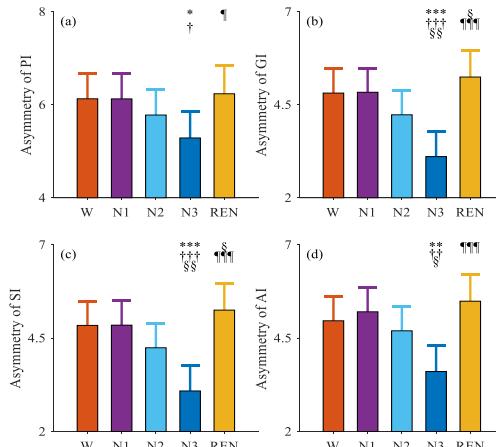


Figure 3. Asymmetry of HRA results of RR intervals corresponding to different sleep stage. (a)-(d) Asymmetry of PI, GI, SI and AI, respectively. The error bar indicates the 95% confidence interval in each stage. *: W vs. the other four stages; †: N1 vs. N2, N3 and REM; §: N2 vs. N3 and REM; ¶: N3 vs. REM. The number of symbols means the degree of significant difference, e.g., *, †, §, ¶: $p<0.05$, **, ††, §§: $p<0.01$; ***, †††, ¶¶¶: $p<0.0001$.

4. Discussion

Due to the regulation of autonomic nervous system (ANS), the speed of heart beats accelerates or decelerates during sleep. The activation of sympathetic or parasympathetic speeds up or slows down the heartbeat, respectively [15]. The pattern of heart rate increases or decreases, or the activity state of the ANS in different sleep stages were analyzed by four heart rate asymmetry measures that is PI, GI, SI and AI.

It is found that the lowest HRA and asymmetry of HRA were detected in N3 stage, and then they increased significantly in REM sleep by PI. This result is consistent with previous findings that heart rate decreases during sleep stage N1, N2 and N3, and increases in REM sleep [16]. PI is a method to estimate the asymmetry of number of points aside the LI which is most intuitive way to express heart rate changes. The number of points below LI decreased in N3 stage, and it increased in REM sleep, so PI could distinguish N3 and REM from other sleep state. However, GI, SI and AI, estimated the heart rate asymmetry by distance, angle and area, respectively, cannot find the difference. It is speculated that the accumulation of the product of the distance, angle or area and quantity buffers the difference among different sleep states.

All of the four asymmetries of HRA indices showed N3 was much lower than other stages. It may be that the heart rate in the N3 stage becomes slower and more stable, and the asymmetry of the heart rate does not deviate from 50 much. Whereas, compared with N3, the heart rate accelerates very fast in the REM sleep. Our study suggests that the HRA indices and asymmetry of HRA metrics could be promising indicators for the clinical N3 and REM sleep monitoring.

Study limitations. The 30s RR intervals were analyzed by HRA measures in this

paper. However, the statistical stability and consistency were not systematic investigated. The physiological significance of asymmetry of HRA during sleep was not entirely understood. It remains to be further studied. Only 16 people's ECG data were used in this study. We will use more data to study the HRA during sleep in future.

5. Conclusion

In this study, PI, GI, SI and AI were used to describe the heart rate asymmetry features during sleep. Both two different analyzing protocols suggested that the asymmetry of N3 stage was the lowest and it increased in REM sleep. Besides, the results obtained by 30s RR series. Our study suggests that the HRA analysis may be applicable in recognizing different sleep stages, especially in differentiating N3 and REM.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (62101129).

References

- [1] Cf G, Mh K. Sleep and control of heart rate. *Clin Chest Med.* 1985 Dec;6(4):595–601.
- [2] Boudreau P, Yeh W-H, Dumont GA, Boivin DB. Circadian Variation of Heart Rate Variability Across Sleep Stages. *Sleep.* 2013 Dec;36(12):1919–28.
- [3] Somers VK, Dyken ME, Mark AL, Abboud FM. Abboud. Sympathetic-Nerve Activity during Sleep in Normal Subjects. *New Engl J Med.* 1993 Feb;328(5):303–7.
- [4] Toscani L, Gangemi PF, Parigi A, Silipo R, Raggianti P, Sirabella E, et al. Human heart rate variability and sleep stages. *Ital J Neuro Sci.* 1996 Dec;17(6):437–9.
- [5] J S. Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss Med Wkly.* 2004 Sep;134(35-36):514–22.
- [6] P. W. Kamen and A. M. Tonkin. Application of the Poincaré plot to heart rate variability: a new measure of functional status in heart failure. *Aust N Z J Med.* 1995 Feb;25(1):18–26.
- [7] Porta A, Casali KR, Casali AG, Gnechi-Ruscone T, Tobaldini E, Montano N, et al. Temporal asymmetries of short-term heart period variability are linked to autonomic regulation. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology.* 2008 Aug;295(2):R550–7.
- [8] Guzik P, Piskorski J, Krauze T, Wykretowicz A, Wysocki H. Heart rate asymmetry by Poincaré plots of RR intervals. *Biomed Tech (Berl).* 2006 Oct;51(4):272–5.
- [9] Karmakar CK, Khandoker AH, Palaniswami M. Phase asymmetry of heart rate variability signal. *Physiol Meas.* 2015;36(2):303.
- [10] Yan C, Li P, Ji L, Yao L, Karmakar C, Liu C. Area asymmetry of heart rate variability signal. *BioMed Eng OnLine.* 2017 Sep;16(1):112.
- [11] Jiang J, Chen X, Zhang C, Wang G, Fang J, Ma J, et al. Heart rate acceleration runs and deceleration runs in patients with obstructive sleep apnea syndrome. *Sleep Breath.* 2017 May;21(2):443–51.
- [12] Goldberger Ary L., Amaral Luis A. N., Glass Leon, Hausdorff Jeffrey M., Ivanov Plamen Ch., Mark Roger G., et al. PhysioBank, PhysioToolkit, and PhysioNet. *Circulation.* 2000 Jun;101(23):e215–20.
- [13] Ichimaru Y, Moody GB. Development of the polysomnographic database on CD-ROM. *Psychiatry Clin Neurosci.* 1999;53(2):175–7.
- [14] Yan C, Li P, Yao L, Karmakar C, Liu C. Impacts of reference points and reference lines on the slope- and area-based heart rate asymmetry analysis. *Measurement,* vol. 137, pp. 515–526, Apr. 2019.
- [15] Piskorski J, Guzik P. The structure of heart rate asymmetry: deceleration and acceleration runs. *Physiol Meas.* 2011;32(8):1011.
- [16] ŽEmaitytė D, Varoneckas G, Sokolov E. Heart Rhythm Control During Sleep. *Psychophysiology.* 1984;21(3):279–89.

Hierarchical Optimization Model Based on V2G Technology

Wei LUO, Xiaohu ZHU¹, Liansong YU

State Grid Electric Power Research Institute, Wuhan Nari Limited Liability Company, Wuhan 430074, China

Abstract. Large-scale electric vehicle (EV) random access to the power grid, the load peak-valley difference will become larger, seriously affect the stable operation of the power grid. In this paper, a V2G based bi-level optimal scheduling model for EV charging and discharging is proposed. The upper model takes the minimum variance of the total load as the objective function, and the lower model takes the increase of user participation and the maximization of user revenue as the objective function. The multi-population genetic algorithm is used to analyze the model, and the results show that the model can not only smooth the load fluctuation, effectively reduce the load peak-valley difference, but also maximize the economic benefits of users participating in V2G service.

Keywords. Power grid load, double-layer model, load peak-valley, genetic algorithm, user income.

1. Introduction

Compared with traditional fuel vehicles, electric vehicles have great advantages in energy saving and emission reduction, and in achieving green travel, it is an inevitable trend that electric vehicles will be chosen as the way of travel in the future[1,2]. Electric vehicles are energy storage devices that can be plugged into the grid as a load or as a distributed power device, and can also be used to enhance the efficient energy use through the two way interaction of the grid. The reasonable dispatch of electric vehicle load can improve the operation state of the power grid, ease the power load in the power grid, effectively manage the charging load, realize load peak shaving and valley filling, at the same time, cooperate with the access of new energy sources such as photovoltaic and wind energy, we will accelerate the integration of wind, light and storage of new energy sources[3,4]. Therefore, it is very important to optimize the charging and discharging strategy of EV.

In order to solve the problem of power load mutation and multi-dimension treatment of centralized dispatching mode caused by large-scale electric vehicles (evs) accessing to power grid, regional hierarchical control is an effective means to guide large-scale electric vehicles (evs) to control[5-7]. The model established in the above-mentioned literatures is mainly aimed at the benefit of power grid measurement, and the ultimate goal is to smooth the load fluctuation, but it does not take into account the response of users and the ability of regional electric vehicles to participate in dispatching. Reference

¹ Corresponding Author; Email: jzzxh18@163.com

[8-11] by defining the loss cost and the average discharge index in the charging process, consummating the user side demand, establishing the grid side and the user side charging and discharging strategy. The model in reference [12] mentioned that the upper layer aims to minimize the fluctuation of the load curve, and the lower layer aims to minimize the distribution loss of the power plant and the cost of materials such as coal, power grid measurement and customer demand imbalance. In reference [13-16], a particle swarm optimization (PSO-RRB) algorithm is proposed to model the charging power limit without considering the different demands of users.

According to the above research situation, this paper takes a residential area as the research object, comprehensively considers the relationship among dispatch center, regional agent, and users, a dual-layer charge-discharge scheduling model for electric vehicles (evs) is proposed, which takes into account the willingness and capability of the users. The upper model minimizes the total load variance of the grid system by determining the charging load and discharging power in each period, and the lower model is based on whether the users have the willingness to participate and the dispatching ability of the users, under the premise of the cooperation of the regional agent with the dispatch center plan, it is important to increase the user participation and maximize the user income. In this paper, the ordered charging model is analyzed by using multi-population genetic optimization algorithm, and the optimal power distribution based on the user's scheduling intention and ability is obtained, it provides a basis for electric vehicle charging and discharging scheduling strategy.

2. Layered and optimized architecture for electric vehicles

2.1. Hierarchical Scheduling architecture design

The idea of hierarchical dispatch is to stratify the dispatch area according to the voltage level, and then divide the dispatch area according to the region. The user communicates with the regional agent, reports his dispatch intention and ability, the agent collects the regional power demand and evaluates the regional dispatch ability, and feeds back the information to the system dispatch center. The hierarchical scheduling architecture constructed in this article is shown in Figure 1.

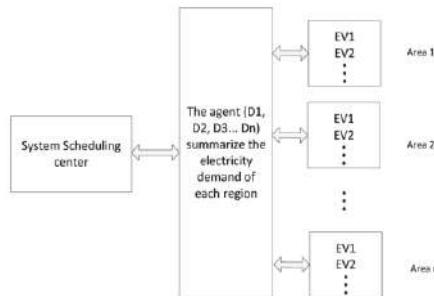


Figure 1. Hierarchical Scheduling Architecture Model for EVs

2.2. Charging and discharging load analysis of electric vehicles

The research on the trip law of electric vehicle is the basis of the implementation of the dispatching strategy. When the residents in the urban community use the fuel car or electric vehicle, it has little influence on the trip characteristics. This article is based on

the 2009 U.S. household car travel survey [17-20] and the following is the probability density function of when electric cars go home and leave home. Assuming that the owner participates in the dispatching plan immediately after returning home, that is, the ev is immediately connected to the power grid, the time when EV i starts to connect to the power grid is $t_{g,i} \sim N(17.9, 3.42)$, and the time when EV i is expected to leave the power grid is $t_{l,i} \sim N(9.24, 3.162)$.

$$f_d(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad (1)$$

In this paper, the hierarchical architecture is simulated based on the above probability distribution model, and the electric vehicle (EV) is investigated and sampled by Monte Carlo sampling method [21], thus, the charging load of EV is modeled to optimize the parameters of the model.

3. Electric vehicle response capability based on V2G

For electric vehicles in the V2G program, the battery capacity can be divided into the battery capacity $B1$ for the V2G program, the battery capacity $B2$ for daily driving, and the battery capacity $B3$ for standby driving, as shown in Figure 2.

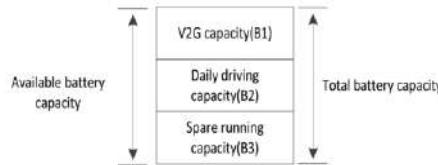


Figure 2. Electric vehicle battery capacity division

The V2G response capacity of a single vehicle can be reported according to the capacity of the $B1$ when the battery capacity of the $B1$ and $B2$ are satisfied when the EV is recharged, the responsiveness of electric vehicles should also be related to the start and end of charging and the minimum time for the maximum power to reach the desired battery state, while the responsiveness of users is also a key factor in the responsiveness.

If the value of response capability is K , Where, $T_{l,i}$ represents the time when the electric vehicle leaves the grid, $T_{g,i}$ represents the time when the electric vehicle is connected to the grid, and $T_{\min,i}$ represents the shortest time when the maximum power reaches the expected battery state. When $K > 1$, the electric vehicle has the response capability. The response capacity of ev is also related to its battery capacity Q , so the formula is as follows:

$$k = \frac{(T_{l,i} - T_{g,i}) Q}{T_{\min,i}} \quad (2)$$

4. Charging and discharging scheduling strategy of electric vehicles based on double-layer optimization model

4.1. Two-level optimization model

In this paper, combined with the travel characteristics of residential areas, the overall power consumption strategy in each area is formulated by the upper dispatch center, and the regional agent makes the overall power consumption strategy in the area under the overall power consumption strategy, so that the total fluctuation curve of the system during the charging and discharging period is minimized, at the same time, the agent's actual dispatch and the dispatch center's power consumption strategy deviation sum is minimum. In the underlying model, the regional agent manages the pre-scheduled charging of the electric vehicles in the region, and maximizes the revenue of the users by combining the users'scheduling willingness and ability.

4.2. Upper objective function

The upper objective function is:

$$\min F = \frac{1}{T-1} \sum_{t=1}^T \left(P_{d,t} + \sum_{n=1}^{N_0} P_{n,t} - P_{DE} - \bar{P}_a \right)^2 + \alpha \sum_{N=1}^{N_0} \sum_{t=1}^T |PD_{n,t} - P_{n,t}| \quad (3)$$

Where T represents 24 time periods in a day; N is the number of agents; $P_{d,t}$ is the load of non-electric vehicle in t period; $P_{n,t}$ represents the charge and discharge load of electric vehicles, and N_0 represents the total number of agents; P_{DE} represents the output of other distributed energy sources in the region; \bar{P}_a represents the average load of the system within T periods, $\bar{P}_a = \frac{1}{T} \sum_{t=1}^T (P_{d,t} + \sum_{n=1}^{N_0} P_{n,t} - P_{DE})$, The scheduling deviation function between the dispatching center and regional agents is: $F = \alpha \sum_{N=1}^{N_0} \sum_{t=1}^T |PD_{n,t} - P_{n,t}|$, Where α represents the penalty coefficient of violation and represents the scheduling deviation and constraints of regional agents; $PD_{n,t}$ represents the scheduling plan made by the regional agent.

4.3. Lower level objective function

For users, the purpose of participating in the V2G scheduling plan is to reduce charging costs and travel costs, and users can obtain maximum benefits according to their response ability.

The lower level objective function with the minimum charging cost is as follows:

$$\min f = \left(\sum_{T_0}^{T_l} \bar{P}_{c,i} C_{c,i} a - \sum_{T_0}^{T_l} \bar{P}_{d,i} C_{d,i} b \right) (T_l - T_0) \bar{K}_i \quad (4)$$

T_0 is the starting time of charging and discharging, T_l is the ending time of charging and discharging, $\bar{P}_{c,i}$ and $\bar{P}_{d,i}$ are the charging power and discharging power respectively, $C_{c,i}$ and $C_{d,i}$ are the charging price and discharging price respectively,

and a , b means charge parameter and discharge parameter respectively, when charging, a is 1, b is 0, when discharging, a is 0, b is 1. It means the average response ability of \bar{K}_i in the area under the jurisdiction of the agent, the average response ability of different area under the jurisdiction of the agent is different.

5. Model solving method

There are many constraints in the bi-level optimization model of electric vehicle charging and discharging, but due to the evolution of genetic algorithm (SGA), it is not necessary to pay much attention to whether the constraints are linear or nonlinear, but directly to the objective function as the search information, with a very high degree of flexibility. When there are super-strong individuals in the group or the group size is small, the convergence boundary will move forward [22,23]. In this paper, a multi-population genetic optimization algorithm (MPGA) is used, and different parameters are set for different populations, and immigration operator is introduced to achieve the simultaneous evolution of different populations. Finally, the elite population is selected [24-26]. Based on the bi-level optimization model of electric vehicle charging and discharging, the optimal charging power of each period can be calculated by multi-population genetic optimization algorithm.

The implementation process of multi-population genetic optimization algorithm is to select the best individuals and put them into elite population each time the population evolves, and the elite population is selected separately and does not participate in the evolution until the elite population reaches the limit and then terminates the program.

6. Case analysis

In this paper, 5 large-scale communities with 5 district agents as an example, agents in coordination with the General Scheduling Center Plan, responsible for their own district charge and discharge scheduling. Set the jurisdiction of the electric vehicles are ordinary household electric vehicles, and electric vehicles are generally charged at night. Let's say an EV in one area responds to 90% of V2G. Select the base power of 200kVA, the scheduling cycle will be 0 to 24 points, to hours as time interval units, a total of 24 time periods. That's 24 hours. In the current period, assume that the electric vehicle charge-discharge load and other basic load does not change, and the owner of the average monthly charge times is 10 times.

The number of EVs and the average V2G capability of each agent in the jurisdiction are set as shown in Table 1. EV parameters in Table 2.

Table 1. The number and average response capacity of electric vehicles in each agent's jurisdiction area

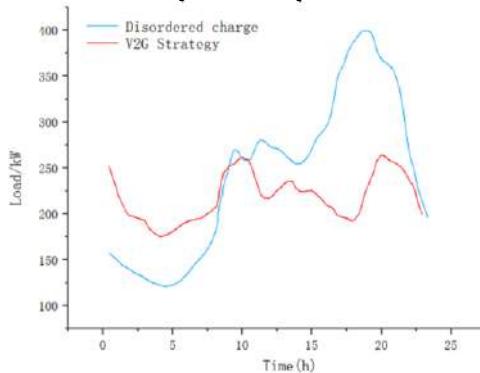
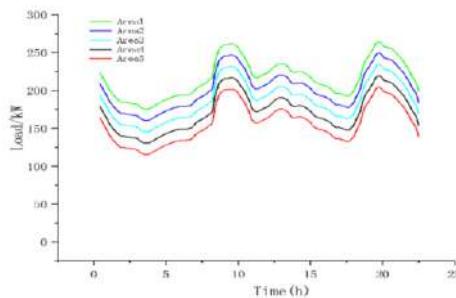
Area	Number of EV	Average response capacity
Area1	2800	1.26
Area2	3300	1.34
Area3	3800	1.41
Area4	4300	1.49
Area5	4800	1.53

Table 2. Electric vehicle parameter list

Parameter	Value
Average charging efficiency/%	90
Average discharge efficiency/%	90
Total energy stored in battery/kWh	120
The user wants the amount of charging/%	90
Power consumption for 100 kilometers/kWh/100km	13

The V2G response optimization results of the EV are shown in Figure 3., the change of power load can be seen under the condition of disordered charging and under the optimal strategy. The peak-to-peak load difference is 272.33 kw between 9:00 ~ 10:00,11:00 ~ 12:00 and 18:00 ~ 21:00. After the optimized scheduling strategy adopted in this paper, the peak load appears at 9:00 ~ 11:00,20:00 ~ 21:00. The difference between peak load and valley load is 90.21 kw, which can reduce the fluctuation of load compared with the state of disordered charging.

As can be seen from Figure 4, the optimization strategy is related to the number of regional electric vehicles and increases with the number of regional electric vehicles. The maximum peak-valley difference is 46KW in area 1 and 47.1 kw in area 5. The trend of electric load is the same, and the average difference between peak and peak is 11KW, and the average difference between valley and valley is 19.8 KW.

**Figure 3.** V2G Response Optimization Curve of EVs**Figure 4.** Load curves of each region under the optimization strategy

7. Conclusion

In this paper, an optimal scheduling strategy for EV charging and discharging in V2G mode is proposed. The objective of the upper model is to minimize the variance of the total load fluctuation of the power grid system, and the lower model is based on whether the users have the willingness to participate and the dispatching ability of the users, under the premise that the regional agents cooperate with the dispatching center planning, pay attention to the balance of regional power load. The model is analyzed by using group genetic algorithm. The results show that the model can smooth the load fluctuation, reduce the load peak-valley difference and stabilize the regional load.

References

- [1] Cai Jiayu, Zhao zhigang. A Review of The Interaction Between Electric Vehicles and Power Grid in V2G Mode [J]. *Electronics World*,2021(07):19-20.
- [2] Li Yirang, Zhangshu,Xiao Xianyong,et al.The charging and discharging scheduling strategy for electric vehicles in V2G mode considering demand from both sides [J]. *Electric Power Automation Equipment*, 2021,41(03):129-135+143.
- [3] Ruifeng Shi, Shaopeng Li, Penghui Zhang, Kwang Y. Lee. Integration of renewable energy sources and electric vehicles in V2G network with adjustable robust optimization[J]. *Renewable Energy*,2020,153.
- [4] Schetinger Annelys Machado, Dias Daniel Henrique Nogueira, Borba Bruno Soares Moreira Cé sar,Pimentel da Silva Gardenio Diogo. Techno - economic feasibility study on electric vehicle and renewable energy integration: A case study[J]. *Energy Storage*,2020,2(6).
- [5] Popoola A. I, Akinpelu E. O, Ewetumo T. Development and Performance Evaluation of an Intelligent Electric Power Switching System[J]. *Journal of Trend in Scientific Research and Development*,2021,5(3).
- [6] Yu. M. In'kov,E. N. Rozenberg,A. I. Maron. Simulation of the Process of Implementation of an Intelligent Electric Power Metering System[J]. *Russian Electrical Engineering*,2020,91(1).
- [7] Evgeny Tretyakov. Demand management by active consumers in intelligent electric power systems[J]. *E3S Web of Conferences*,2020,157.
- [8] Huang Minfa, Xiang Baolin, Peng Juanjuan, Zhang Lu. Optimal Configuration Method for EV Charging Station in Distribution Network Considering User Adjustment under V2G mode[J]. *IOP Conference Series: Earth and Environmental Science*,2021,784(1).
- [9] Heilmann C., Friedl G.. Factors influencing the economic success of grid-to-vehicle and vehicle-to-grid applications–A review and meta-analysis[J]. *Renewable and Sustainable Energy Reviews*,2021,145.
- [10] Tan Kang Miao, Ramachandaramurthy Vigna K.,Yong Jia Ying,Tariq Mohd. Experimental verification of a flexible vehicle-to-grid charger for power grid load variance reduction[J]. *Energy*,2021,228.
- [11] Popoola A. I, Akinpelu E. O, Ewetumo T. Development and Performance Evaluation of an Intelligent Electric Power Switching System[J]. *Journal of Trend in Scientific Research and Development*,2021,5(3).
- [12] Wu Wei, Lin Boqiang. Benefits of electric vehicles integrating into power grid[J]. *Energy*,2021,224.
- [13] Zheng Xue qin, Yao Yi ping. Multi-objective capacity allocation optimization method of photovoltaic EV charging station considering V2G[J]. *Journal of Central South University*,2021,28(2).
- [14] Liu Lijun, Xie Feng, Huang Zonglong, Wang Mengqi. Multi-Objective Coordinated Optimal Allocation of DG and EVCSSs Based on the V2G Mode[J]. *Processes*,2020,9(1).
- [15] Wang Dian, Sechilaru Manuela,Locment Fabrice. PV-Powered Charging Station for Electric Vehicles: Power Management with Integrated V2G[J]. *Applied Sciences*,2020,10(18).
- [16] Shigeru Tamura. A V2G strategy to increase the cost-benefit of primary frequency regulation considering EV battery degradation[J]. *Electrical Engineering in Japan*,2020,212(1-4).
- [17] Hao Tang, Chang Liu,Yonglun Cao,Kai Lv,Qianli Zhang. Hierarchical scheduling learning optimisation of two-area active distribution system considering peak shaving demand of power grid[J]. *Discrete Event Dynamic Systems*,2021(prepublish).
- [18] Ruifeng Shi, Shaopeng Li, Penghui Zhang, Kwang Y. Lee. Integration of renewable energy sources and electric vehicles in V2G network with adjustable robust optimization[J]. *Renewable Energy*,2020,153.
- [19] Yu. M. In'kov, E. N. Rozenberg,A. I. Maron. Simulation of the Process of Implementation of an Intelligent Electric Power Metering System[J]. *Russian Electrical Engineering*,2020,91(1).
- [20] Evgeny Tretyakov. Demand management by active consumers in intelligent electric power systems[J]. *E3S Web of Conferences*,2020,157.

- [21] He Wei, Chen Fan, Miao Han, et al. Cross-entropy Monte Carlo Method for Power System Abundance Evaluation [J]. *Acta electronica sinica*, 2020, 35(03): 235-245.
- [22] Li Zening, Su Su, Jin Xiaolong, Chen Houhe, Li Yujing, Zhang Renzun. A hierarchical scheduling method of active distribution network considering flexible loads in office buildings [J]. *International Journal of Electrical Power and Energy Systems*, 2021, 131.
- [23] Shen Xiaoyang. A hierarchical task scheduling strategy in mobile edge computing [J]. *Internet Technology Letters*, 2020, 4(5).
- [24] Li Jiale, Liu Zhenbo, Wang Xuefei. Public charging station location determination for electric ride-hailing vehicles based on an improved genetic algorithm [J]. *Sustainable Cities and Society*, 2021, 74.
- [25] Jordán Jaume, Palanca Javier, del Val Elena, Julian Vicente, Botti Vicent. Localization of charging stations for electric vehicles using genetic algorithms [J]. *Neurocomputing*, 2021, 452.
- [26] Wang Jianjun, Pan Junnan, Huo Jikun, Wang Ran, Li Li, Nian Tiantian. Research on Optimization of Multi-AGV Path Based on Genetic Algorithm Considering Charge Utilization [J]. *Journal of Physics: Conference Series*, 2021, 1769(1).

Research of Power Distribution System Dynamic Simulation Based on Grid's Harmonic Current

Wei QIN^{a,1}, Tiansong GU^b and Hongliang LI^b

^a Wuhan Nari Limited Library Company of State Grid Electric Power Research Institute, China

^b State Grid Jibei Electric Power Company Limited, China

Abstract. Harmonic current in power grid will cause extra power consumption of electrical equipment and affect the stable operation of power grid. Based on the dynamic model experiment system of 10kV distribution network, a working platform for parallel operation of charging pile was built, which studied the harmonic current components and analyzed the simulation waveform. The results show that with the increase of number of charging piles, the content of harmonic current will decrease correspondingly, the main components are low frequency odd harmonics; When 16 charging piles are connected in parallel, three harmonics are the main components; the waveform of the simulation calculation curve and the actual measurement curve are basically consistent, and the current amplitude of each frequency harmonic wave is basically the same, so it is proved that the dynamic model test platform is reasonable and feasible to test grid's harmonic current. The dynamic model can provide technical support for harmonic suppression research.

Keywords. Charging pile, LCL, harmonic current, dynamic simulation, artificial load.

1. Introduction

With the introduction of document “The 13th Five-Year Plan for the Development of National Strategic Emerging Industries”, new energy vehicles will occupy the main market of road traffic. Charging pile(CP), as the power supply device of new energy vehicles, is everywhere in each station. Due to the existence of power electronic devices in CP, capacitance and inductance will cause the instability of the power grid. Therefore, scholars at home and abroad have carried out studies on the suppression of parallel resonance characteristics, and have proposed a virtual resistance damping method based on capacitance and voltage feedback, so as to realize the stable operation of the system [1]. The dynamic model experiment system is essentially the operation platform of the distribution network equipment that simulates full voltage and full current. According to the principle of similarity and the requirements of the equipment, an experiment platform can be established to simulate the real scene. Reference [2] uses equipment such as fault indicator and traveling wave tester to carry out static detection, thus improving the overall performance. In this reference [3], the overhead and cable mixed line distribution

¹ Corresponding Author; E-mail: qwqc1012@qq.com.

network for new energy system access is built, which can carry out grounding fault analysis of different types and related tests of other automatic equipment.

At present, CP test research is limited to simulation analysis or power network simplification research, and has not carried out the construction of the whole distribution network. Its performance test has a certain gap with the practical operation results, which is not convincing. Dynamic simulation experiment system has studied the electrical characteristics of different power equipment, such as capacitors, switches, etc., but lack of CP test research. Therefore, CP is taken as the test object and a dynamic model experiment platform is built to carry out the harmonic current test of CP.

2. CP parallel model

CP is equipped with LCL active power filter, which adopts modern power electronics technology to automatically track harmonic current generated in the power grid. CP can compensate or cancel harmonic current to ensure the stability and reliability of power grid voltage. The single-phase CP system structure is shown in Fig.1.

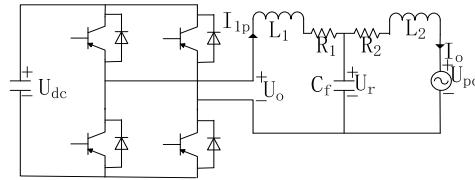


Figure 1. Structural diagram of single-phase CP system

In Fig.1, U_{dc} is the DC input voltage; I_{lp} is the closed-loop current control value at the converter side; U_r is the capacitor voltage; U_o is the output voltage; I_o is the output current of the converter network side, and U_{pc} is the voltage of the converter access point.

2.1. Single-phase CP closed-loop model

Single-phase CP is controlled by a double closed-loop network of closed-loop current I_{lp} and capacitor voltage U_r , so as to suppress the resonance phenomenon of the system. The established single-phase CP current and voltage transfer function is shown in Fig.2.

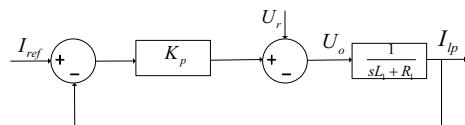


Figure 2. Current and voltage transfer function of single CP

The output voltage can be calculated as:

$$U_o(s) = U_r(s) + K_p(s)[I_{ref}(s) - I_{lp}(s)] \quad (1)$$

The transfer function of important parameters in the system structure is given by:

$$\begin{aligned} I_{lp}(s) &= H_{l1}(s)U_o(s) + H_{l2}(s)U_i(s) \\ I_o(s) &= H_{o1}(s)U_o(s) + H_{o2}(s)U_i(s) \\ U_r(s) &= H_{r1}(s)U_o(s) + H_{r2}(s)U_i(s) \end{aligned} \quad (2)$$

Among them:

$$\begin{aligned} H_{l1}(s) &= \frac{L_2 C_j s^2 + R_2 C_j s + 1}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \\ H_{l2}(s) &= \frac{-1}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \\ H_{o1}(s) &= \frac{1}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \\ H_{o2}(s) &= \frac{-(L_2 C_j s^2 + R_2 C_j s + 1)}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \\ H_{r1}(s) &= \frac{L_2 s + R_2}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \\ H_{r2}(s) &= \frac{L_1 s + R_1}{L_1 L_2 C_i s^3 + (R_2 L_1 + R_1 L_2) C_j s^2 + (L_1 + L_2 + C_j R_1 R_2) s + R_1 + R_2} \end{aligned}$$

The closed-loop model of a single CP is as follows:

$$I_o(s) = G_T(s)I_{ref}(s) + Y_{eq}(s)U_{pc}(s) \quad (3)$$

Where the current coefficient and equivalent admittance can be respectively given by:

$$G_T(s) = \frac{K_p(s)H_{o1}(s)}{1 - H_{r1}(s) + K_p(s)H_{l1}(s)} \quad (4)$$

$$Z_{eq}(s) = \frac{K_p(s)H_{l2}(s)H_{o1}(s) - H_{o1}(s)H_{r2}(s)}{1 - H_{r1}(s) + K_p(s)H_{l1}(s)} \quad (5)$$

According to NORTON theorem, CP internal circuit can be equivalent to parallel current and admittance, as shown in Fig. 3.

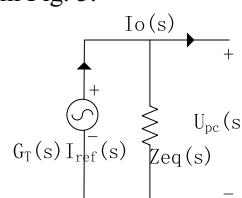


Figure 3. Norton equivalent model of CP circuit

2.2. Multi-CP parallel model

With the development of smart power grid and the construction of smart community, CP is more and more widely used. Its connection mode is generally in parallel mode. Fig.4 is a topology diagram of 7V CP connected in parallel, which is composed of NORTON equivalent circuit calculated by Equations (4) and (5). Among them, $I_t(s)$ is the total current after CP parallel connection, that is, the sum of all output currents; $U_t(s)$, $Z_t(s)$ are power grid voltage and power grid line admittance respectively.

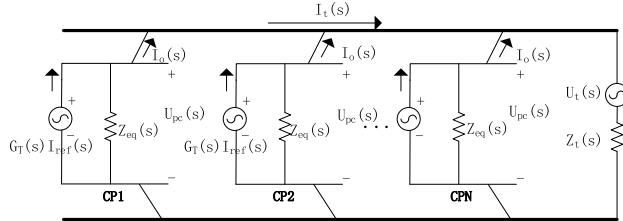


Figure 4. Network topology diagram of multi-EV chargers

To ensure the reliability of the CP network, if the parallel CP-general specifications and models are the same and the parameters are the same, the interface voltage can be given by:

$$U_{pc}(s) = \frac{\sum_{i=1}^N G_T^i(s) I_{ref}^i(s) + U_t(s) Z_t(s)}{\sum_{i=1}^N Z_{eq}^i(s) + Z_t(s)} \quad (6)$$

Where $G_T^i(s)$, $I_{ref}^i(s)$, $Z_{eq}^i(s)$ are the current coefficient, closed-loop current control value and equivalent admittance of the i^{th} cp respectively.

For the output current of the first CP, then:

$$I_g^1(s) = G_T^1(s) I_{ref}^1(s) + Z_{eq}^1(s) U_{pc}(s) \quad (7)$$

It can be obtained simultaneously:

$$I_g^1(s) = R_1(s) I_{ref}^1(s) + \sum_{k=1}^N P_1^k(s) I_{ref}^k(s) - S_G^1(s) U_t(s) \quad (8)$$

Among them:

$$P_1^k(s) = \frac{-Z_{eq}^1(s) G_T^1(s)}{\sum_{i=1}^N Z_{eq}^i(s) + Z_t(s)} \quad (9)$$

$$S_G^1(s) = \frac{Z_{eq}^1(s)Z_t(s)}{\sum_{i=1}^N Z_{eq}^i(s) + Z_t(s)} \quad (10)$$

It can be seen that $S_G^1(s)U_t(s)$ is the representation of the influence degree of charging pile when the grid voltage is distorted.

3. Experimental scheme

3.1. Dynamic model experimental system design

With the popularity of electric vehicles, the capacity and number of CPS have increased significantly. The capacity of DC CP has reached 250KW, so a 10KV distribution network dynamic mode experiment system is built, as shown in Fig.5, which can ensure the energy supply of CP network. The whole system consists of substation, 10KV distribution line, 0.4kV distribution equipment, voltage and current boosting equipment, etc. LH, R are used for grounding arc suppression coil and resistance to achieve non-metallic neutral grounding, so as to limit the grounding fault current. For arc suppression coil and resistance used for grounding, non-metallic grounding of neutral point is realized, so as to limit grounding fault current; TM main transformer, capacity of 5MW, L,C are 10KV reactor and 10KV capacitor respectively; QF is a busbar switch, which realizes different power supply forms through different switching combinations.

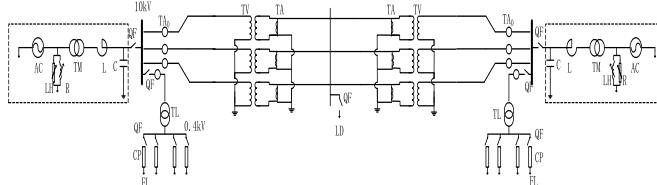


Figure 5. Topology of dynamic model for 10kV distribution experiment system

TA₀, TA and LD are the loads of zero sequence current transformer, upcurrent transformer, upvoltage transformer and 10 kV distribution network respectively. TL is the step-down transformer, realizing the transform of 10 kV to 0.4 kV output, and provides the input voltage for the CP network; FL is the analog load.

3.2. Dynamic model experimental system design

Since the neutral point of power supply adopts parallel grounding of grounding resistance and arc suppression coil, and the dynamic mode test system of distribution network operates in full working mode, the zero sequence current caused by single-phase grounding fault is very small, and the traditional current transformer cannot achieve the actual measurement accuracy, so it cannot realize the action of starting relay protection device.

As shown in Fig.6, in order to solve the accuracy problem, the primary circuit and secondary winding are wound together on the same ring core, and the secondary current value is improved through the principle of magnetic balance, so as to improve the measurement accuracy and finally meet the measurement requirements.

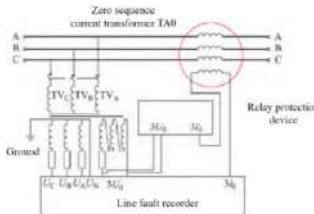


Figure 6. Structure of zero sequence current transformer

3.3. CP network

CP network of simulated intelligent cell is shown in Fig.7. The whole network is composed of power distribution cabinet, converter cabinet, charger, etc. Converter cabinet realizes the rectification or inverter of power grid current, so as to realize the AC and DC charging of the charger [4]. For DC charger, constant current and constant voltage are generally used for subsection charging. For example, charge the current to about 250A first and keep it stable, and then boost the voltage. When the voltage reaches a predetermined value of 1000V, the analog load can be charged.

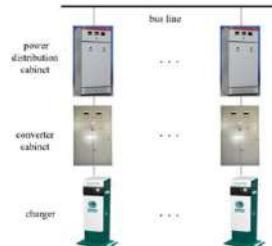


Figure 7. Network structure of CP

3.4. Simulated load

The internal structure of analog load is shown in Fig.8, which is composed of battery management system (BMS) simulator, charger interface, resistance load and controller according to different interfaces. BMS simulator can not only improve battery efficiency, but also achieve battery protection. Resistance load, according to the specification and requirements of the national standard "GB/T 18487.1-2015", the nominal value is set as 1000Ω which meets between 970Ω and 1030Ω .

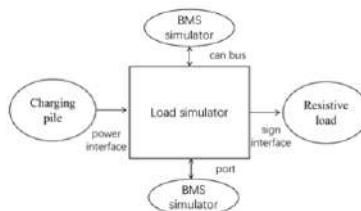


Figure 8. Internal structure of artificial load

4. Experimental scheme

Due to the existence of power distribution cabinet and converter cabinet, the influence of harmonic current on power grid current cannot be ignored. In order to realize the study on harmonic current of CP network in intelligent residential area [5-7], a simulation experiment platform is built as shown in Fig.5. The capacity of a single CP is 250 kW. In the experiment, the maximum value of CP in parallel is 16, and the total capacity is 4 MW, lower than the total capacity of the main transformer.

4.1. Harmonic current measurement

Under the normal operation of dynamic simulation experiment system, different numbers of cp are randomly connected in parallel to measure the total CP network current [8]. The summary of harmonic current data is shown in Table 1.

Table 1. Harmonic current content of different EV chargers

number	CP number	total harmonic distortion rate%	the main harmonic
1	1	4. 5	H9
2	4	3. 1	H7
3	8	2. 9	H5
4	10	2. 7	H5
5	16	1. 8	H3

4.2. Comparative analysis

When simulating 16 CPs in parallel, the network current value was saved and waveform was drawn, which was compared with the actual measured waveform, as shown in Fig.9.

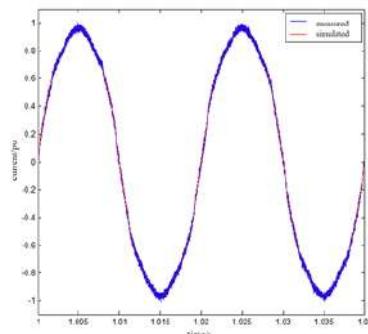


Figure 9. Comparative analysis of harmonic current in different EV chargers

It can be seen that the simulation calculation of harmonic current is basically consistent with the actual measurement curve, and there is some noise in the actual measurement curve, which may be caused by the test environment. It indicates that the dynamic mode experimental system is feasible to measure harmonic current.

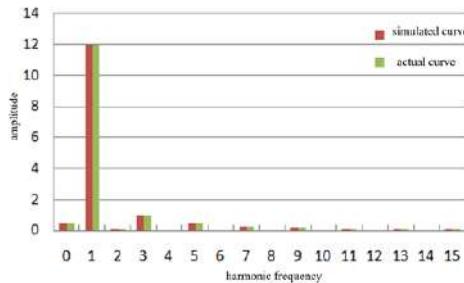


Figure 10. Comparative analysis of different frequency harmonic current amplitude

In Fig.10, the 3rd harmonic is the main harmonic component in the simulation curve and the actual curve, so the 3rd harmonic needs to be the key suppression component.

5. Conclusion

In this paper, the dynamic model experiment platform of charging pile network distribution network was built, and the parameters of power supply network, charging pile and simulated load were designed.

The research conclusions are as follows:(1) With the increase of CP number, harmonic current content will decrease correspondingly, and the main component is low frequency odd harmonics; (2)When 16 CP are connected in parallel, the third harmonic is the main component. The waveform of the simulated curve is consistent with that of the measured curve, and the harmonic component is basically consistent. The dynamic model can be used for CP harmonic current analysis, which provides an effective basis for domestic and foreign research institutes to carry out relevant tests.

References

- [1] Serpa L A, Ponnaluri S. A Modified Direct Power Control Strategy Allowing the Connection of Three-phase Inverters to the Grid Through LCL Filters [J]. IEEE Transactions on Industry Applications, 2007, 43 (5) : 1388-1400.
- [2] B. Jiang, X. Z. Dong, S. X. Shi. Experimental Research on Typical Method of Single Phase Grounding Fault Line Selection in Distribution Network [J]. Electric power automation equipment, 2015,35(11):67-74.
- [3] F. G. Zhang, T. Wu, G. S. Yi, et al. Research on dynamic model of distribution network based on fault location test [J]. Water and power energy science,2018,36(09):167-171.
- [4] X. F. Tao. Design and Implementation of Automatic Test Platform for ELECTRIC Vehicle DC Charging Pile [D]. University of Electronic Science and Technology of China,2018.
- [5] W. Chen, J. Zhou, Z. Xu, et al. Electric measurement & instrumentation,2019,56(06):26-31+49.
- [6] J. Zhou, G. Y. Ren, C. Wei, et al. Electric power system protection and control,2017,45(05):18-25.
- [7] B. B. Zhong. Harmonic Analysis of Electric Vehicle Charging Station [D]. Huazhong University of Science and Technology, 2015.
- [8] W. S. Song, X. Y. FENG. An Internal Relation between Single-phase three-level SVPWM modulation and Carrier SPWM [J]. Transactions of China Electrotechnical Society, 2012, 27(6): 131-138.

Study on Inter-Turn Short Circuit Test for Distribution Reactor

Wei QIN¹, Tiansong GU, Hongliang LI

State Grid Electric Power Research Institute, Wuhan Nari Limited Liability Company, Wuhan 430074, China

Abstract. Inter-turn short circuit is a common fault in reactor. In view of the current situation of insufficient condition based maintenance of reactor, an on-line monitoring method suitable for Inter-turn short circuit fault of distribution reactor was proposed in this paper. Firstly, the equivalent circuit model of reactor was built, and the calculation method of electrical parameters of reactor was obtained; Then, a set of on-line monitoring system for reactor short-circuit fault was built. Finally, the short-circuit test of reactor under different fault location and fault degree was carried out. The results show that the equivalent reactance and resistance change rate of inter-turn short circuit of reactor have the symmetry of short circuit position, the absolute value of change rate in the middle is the largest, and the absolute value of change rate at both ends is the smallest. With the deepening of short-circuit fault, the equivalent reactance decreases and the decreasing range increases. The equivalent resistance shows an increasing trend, and the increasing range increases. This research technology can realize the effective monitoring of inter-turn short circuit fault of distribution reactor.

Keywords. Distribution reactor, equivalent reactance, equivalent resistance, position of short circuit, degree of short circuit

1. Introduction

With the development of smart grid, the scale of distribution network is becoming larger, the power load is becoming higher, and the harmonic content in distribution network is becoming higher due to large-scale applications such as nonlinear load charging pile. In order to solve this problem, reactors, as inductive equipment, cooperate with capacitors to form passive power filters, which can improve the power quality of power grid [1-2]. Dry-type reactor is widely used in distribution network. Air is the heat sink mode. The short-circuit fault of turn insulation can easily cause the burning of windings [3-5], which may cause serious power accidents.

Electric workers all over the world have carried out a lot of research on the inter-turn short-circuit monitoring of reactor, which is generally divided into electrical and non-electrical monitoring methods [6-8]. Relevant scholars in colleges and universities [9-11] determine the location of temperature sensor based on the simulation results of temperature field, and monitor the operating status of reactor by measuring the winding temperature. The advantage of this method is simple installation and wide coverage. However, it is not possible to directly measure and calculate the temperature inside the

¹ Corresponding Author; Email: qwqc1012@qq.com

reactor. The magnetic induction detection coil method [12-13] uses the principle of electromagnetic induction to detect serious inter-turn short-circuit faults quickly with high sensitivity, but it cannot detect inter-turn short-circuit faults at different positions of the reactor. The zero-sequence amplitude comparison monitoring method [14-15] compares the bus voltage fault component with the compensating voltage fault component after inter-turn faults. Based on the above analysis, this paper presents an on-line monitoring method for inter-turn short circuit faults of distribution reactors in view of the status of insufficient maintenance of reactor.

2. Short Circuit Calculation Model

The original reactor winding is composed of n windings in parallel. Among them, some of the windings of the i -th layer have inter-turn short circuits. The closed loop formed by the short-circuit loop is recorded as the $n+1$ winding. The corresponding equivalent short-circuit model is shown in Figure 1.

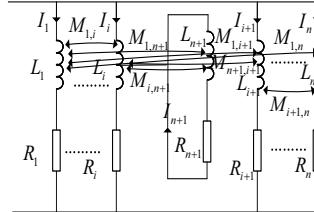


Figure 1. Equivalent short circuit model of reactor.

Second paragraph. Carrying out circuit analysis, each normal parallel winding can be equivalent to the inductance L_i and the resistance R_i in series, the voltage is equal, denoted as U , and the current is I_i . Due to the effect of inductance, there is mutual inductance $M_{i,j}$ between the windings. The short-circuit loop voltage is 0. According to electromagnetic induction, there is still induced current in the short-circuit loop, which is recorded as I_{n+1} . After a short-circuit fault occurs, the voltage equations of $n+1$ branches are

$$\begin{bmatrix} R_1 + j\omega L_1 & j\omega M_{1,2} & \cdots & j\omega M_{1,i+1} & \cdots & j\omega M_{1,n} & j\omega M_{1,n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ j\omega M_{i,1} & \cdots & R_i + j\omega L_i & \cdots & j\omega M_{i,n-1} & j\omega M_{i,n} & j\omega M_{i,n+1} \\ j\omega M_{i+1,1} & j\omega M_{i+1,2} & \cdots & R_{i+1} + j\omega L_{i+1} & \cdots & j\omega M_{i+1,n} & j\omega M_{i+1,n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ j\omega M_{n,1} & j\omega M_{n,2} & \cdots & j\omega M_{n,i+1} & \cdots & R_n + j\omega L_n & j\omega M_{n,n+1} \\ j\omega M_{n+1,1} & j\omega M_{n+1,2} & \cdots & j\omega M_{n+1,i+1} & \cdots & j\omega M_{n+1,n} & R_{n+1} + j\omega L_{n+1} \end{bmatrix} \begin{bmatrix} I_1 \\ I_i \\ I_{i+1} \\ I_n \\ I_{n+1} \end{bmatrix} = \begin{bmatrix} U \\ U \\ U \\ U \\ 0 \end{bmatrix} \quad (1)$$

3. Design of Online Monitoring System

The on-line monitoring system for short-circuit faults of distribution reactor mainly consists of analog and digital circuits. First, the bus-side windings of reactor are collected by high-precision voltage transformer and current transformer. After the signal conditioning module filters, amplifies, removes noise, then outputs the digital signal to

the control unit ARM through the AD module, completes a series of data processing and data storage work, and finally transmits the reactor voltage and current signals to the data center through wireless communication to achieve real-time monitoring of the reactor failure status, see Figure 2. Among them, the digital circuit can store the data information of the mutual inductor, can analyze in real time, and improves the timeliness of fault identification. At the same time, the data analysis uses a quasi-synchronous algorithm to ensure the accuracy of measurement results.

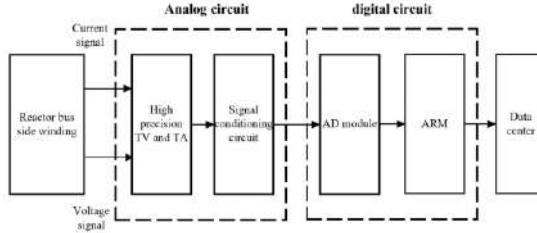


Figure 2. On-line monitoring system.

3.1. Calculation of equivalent reactance and resistance

The calculation of the total voltage and current signals of the reactor is shown in the following formulas (2) and (3), respectively.

$$u(t) = a_{ud} + \sum_{k=1}^{\infty} [a_{uk} \cos(2k\pi f_0 t) + b_{uk} \sin(2k\pi f_0 t)] \quad (2)$$

$$i(t) = a_{id} + \sum_{k=1}^{\infty} [a_{ik} \cos(2k\pi f_0 t) + b_{ik} \sin(2k\pi f_0 t)] \quad (3)$$

In the formula: f is the base frequency; a_{ud} , a_{id} are the direct component of $u(t)$, $I(t)$; a_{uk} , a_{ik} are the imaginary part of the k -th harmonic component of $u(t)$, $I(t)$; b_{uk} , b_{ik} are the imaginary part of the k -th harmonic component of $u(t)$, $I(t)$.

The real and imaginary parts of the voltage and current fundamental signals are calculated as (4):

$$\begin{aligned} a_{u1} &= \frac{2}{T} \int_0^T u(t) \cos(2\pi f_0 t) dt \\ b_{u1} &= \frac{2}{T} \int_0^T u(t) \sin(2\pi f_0 t) dt \\ a_{i1} &= \frac{2}{T} \int_0^T i(t) \cos(2\pi f_0 t) dt \\ b_{i1} &= \frac{2}{T} \int_0^T i(t) \sin(2\pi f_0 t) dt \end{aligned} \quad (4)$$

The amplitude of the base wave signal is calculated as (5)

$$\begin{aligned} u_m &= \sqrt{a_{u1}^2 + b_{u1}^2} \\ i_m &= \sqrt{a_{i1}^2 + b_{i1}^2} \end{aligned} \quad (5)$$

The phase of the fundamental signal is calculated as

$$\begin{aligned} \varphi_{u1} &= \arctan \frac{a_{u1}}{b_{u1}} \\ \varphi_{i1} &= \arctan \frac{a_{i1}}{b_{i1}} \end{aligned} \quad (6)$$

Further, the equivalent resistance and reactance values of the reactor at the base wave can be obtained by the following formulas:

$$R = \frac{a_{ul}a_{rl} + b_{ul}b_{rl}}{a_{rl}^2 + b_{rl}^2} \quad (7)$$

$$X = \frac{a_{ul}b_{rl} - a_{rl}b_{ul}}{a_{rl}^2 + b_{rl}^2} \quad (8)$$

3.2. Quasi-synchronization algorithm

In actual measurement, due to the influence of power quality, the frequency of power network may fluctuate, so the sampling period is not necessarily an integer multiple of the period, which may lead to measurement error, so the same standard is introduced. The core of the quasi-synchronization algorithm is to improve the accuracy of the mean value of periodically sampled signals based on adequate sampling data and appropriate data processing methods.

According to the Fourier expansion, it can be seen that the periodic signal $f(t)$ of voltage or current is decomposed into

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \sin(2\pi k f_0 t + \varphi_k) \quad (9)$$

The effective value is calculated as (10)

$$\overline{f(t)} = \sqrt{\frac{1}{T} \int_{T_0}^{T_0+T} f^2(t) dt} \quad (10)$$

Where: T is the voltage or current signal period; T_0 is the starting point of integration. Make $x = wt$, Then there

$$\overline{f(x)} = \frac{1}{2\pi} \int_{x_0}^{x_0+2\pi} f(x) dx \quad (11)$$

Let the number of sampling periods be N and the period deviation be Δ , The number of sampling points in each cycle is n , in the integral interval $[x_0, x_0 + n \times (2\pi + \Delta)]$ Upper equal interval sampling $n \times N + 1$ data, make

$$F^1 = \frac{1}{\sum_{i_0=0}^{N+i_0} \rho_{i_0}} \sum_{i_0=0}^{N+i_0} \rho_{i_0} f(x_{i_0})$$

The recurrence formula is (12)

$$F^n = \frac{1}{\sum_{i_0=0}^{N+i_0} \rho_{i_0}} \sum_{i_0=0}^{N+i_0} \rho_{i_0} F^{n-1}, n = 2, 3, \dots \quad (12)$$

Among them ρ_{i_0} is the weight coefficient, which can be further obtained

$$\lim_{n \rightarrow \infty} F_n = a_0 = \overline{f(x)} \quad (13)$$

The sum of the product of the sampling value and the corresponding weight coefficient is the average value.

4. Data analysis

In the test, the reactor is dry-type air core type, the voltage level is 35 kV, and the equivalent resistance and reactance are 0.182Ω and 1.24Ω respectively.

4.1. Influence of frequency fluctuation

Due to the phenomenon of frequency fluctuation in power grid voltage, the whole cycle sampling can not be guaranteed. Considering that under abnormal conditions of power system, the allowable deviation of power supply frequency shall not exceed ± 1 Hz and the frequency fluctuation range is [49 Hz ~ 51 Hz], the equivalent resistance and equivalent reactance of reactors with different frequencies under normal conditions are measured. The calculation formulas of equivalent resistance change rate dR/df and equivalent reactance change rate dX/df are shown in equation (14). The values of quasi synchronous algorithm and untreated case are compared and analyzed, as shown in Figure 3 and Figure 4 respectively.

$$\begin{aligned} dR/df &= \frac{\Delta R}{\Delta f} = \frac{R_1 - R_2}{f_1 - f_2} \\ dX/df &= \frac{\Delta X}{\Delta f} = \frac{X_1 - X_2}{f_1 - f_2} \end{aligned} \quad (14)$$

Where: R_1 and X_1 are the equivalent resistance value and reactance value of frequency f_1 respectively; R_2 and X_2 are the equivalent resistance value and reactance value of frequency f_2 respectively; ΔR 、 ΔX 、 Δf is the change of equivalent resistance, equivalent reactance and frequency respectively.

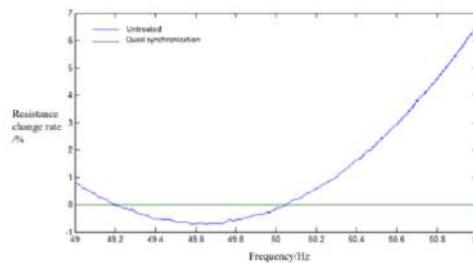


Figure 3. Rate of change for resistance at different frequencies.

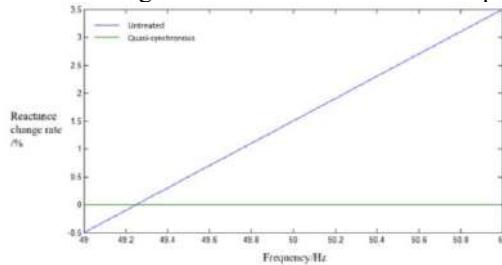


Figure 4. Rate of change for reactor at different frequencies.

As can be seen from Figure 3, the equivalent resistance value changes with the increase of frequency, showing a trend of first decreasing and then increasing. After introducing the quasi synchronous algorithm, the resistance value is almost unchanged. It can be seen from Figure 4 that with the increase of frequency, the equivalent reactance value also gradually increases, showing a positive correlation. After the introduction of quasi synchronous algorithm, the reactance value is almost unchanged.

4.2. Effect of temperature rise

When the reactor works under the rated working condition, the winding temperature will increase and the equivalent reactance and resistance of the reactor will change. In order to further study its stable value, after the voltage, current and other parameters are stable, the equivalent reactance and resistance values will be summarized every 0.5h. The calculation process of the equivalent resistance change rate dR/dt and the equivalent reactance change rate dX/dt is shown in equation (15)

$$\begin{aligned} dR/dt &= \frac{\Delta R}{\Delta t} = \frac{R_1 - R_2}{t_1 - t_2} \\ dX/dt &= \frac{\Delta X}{\Delta t} = \frac{X_1 - X_2}{t_1 - t_2} \end{aligned} \quad (15)$$

Where, R_1 and X_1 are the equivalent resistance value and reactance value at time t_1 , R_2 and X_2 are the equivalent resistance value and reactance value at time t_2 . ΔR , ΔX and Δt are equivalent resistance variation, equivalent reactance variation and time variation respectively. Draw the curve of equivalent resistance change rate and reactance change rate, as shown in Figure 5.

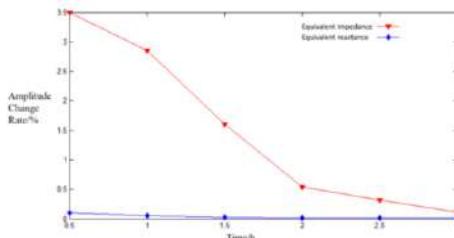


Figure 5. Change rate of equivalent resistance and reactance.

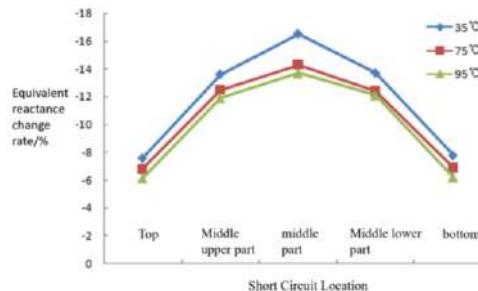


Figure 6. Relationship between the change rate of equivalent reactance and temperature rise.

As can be seen from Fig. 5, with the passage of time, the change rate gradually decreases, the equivalent resistance first increases and then tends to be stable, and the change rate is only 0.5% after 2h; Similarly, with the passage of time, the equivalent reactance value increases slightly, and the change rate of equivalent reactance is basically 0. Therefore, the equivalent reactance and resistance can be measured after 2h.

The reactor is cylindrical and the winding is wound from inside to outside. According to the axial direction, the winding is divided into five equal parts from

top to bottom. The short-circuit fault location can be divided into top, middle upper part, middle lower part and bottom, with a total of 5 parts, each accounting for 20%. Introducing inter turn short-circuit simulation leads at different positions of the reactor and manually closing the short-circuit switch can simulate short-circuit faults with different turns at different positions [16-19]. R' and X' are the equivalent resistance change rate and reactance change rate at the current short-circuit position respectively. The calculation process is shown in equation (16)

$$R' = \frac{R_s - R_0}{R_0} \times 100\% \quad (16)$$

$$X' = \frac{X_s - X_0}{X_0} \times 100\%$$

Among them, R_0 and X_0 are the equivalent resistance and reactance values under normal conditions, and R_s and X_s are the equivalent resistance and reactance values under current short circuit position. In order to further study the relationship between equivalent reactance of reactor and resistance and temperature under different turn to turn short circuit positions, common temperatures such as winding temperature 35 °C, 75 °C and 95 °C are selected for testing, and the curves are drawn as shown in figures 6 and 7.

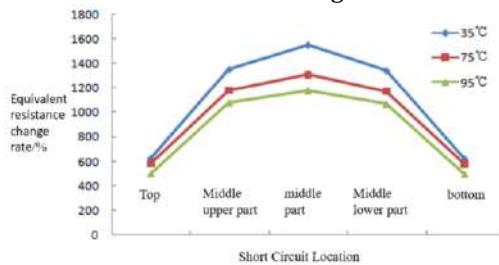


Figure 7. Relationship between the change rate of equivalent resistance and temperature rise.

It can be seen from Fig. 6 that under the symmetrical position, the equivalent reactance rate of inter turn short circuit of reactor is consistent, the absolute value of change rate in the middle is the largest, the absolute value of change rate at both ends is the smallest, and the overall change rate is between - 5.8% and - 16.2%. Under the same short-circuit position, the absolute value of the change rate of equivalent reactance decreases with the increase of temperature. It can be seen from Figure 7 that the change rate of equivalent resistance of reactor turn to turn short circuit also has the symmetry of short circuit position. The overall change rate is between 4.12 ~ 15.3 times. At the same short circuit position, the change rate of equivalent resistance decreases with the increase of temperature.

4.3. influence of different short circuit degree

Define fault degree st as the ratio of short-circuit winding height h_{st} to total reactor winding height h_t , i.e

$$st = \frac{h_{st}}{h_t} \times 100\% \quad (17)$$

Obviously, according to the short circuit position, there is $st \leq 20\%$, for further quantification, it is divided into three intervals [0% ~ 7%), [7% ~ 14%) and

[14% ~ 20%], corresponding to mild fault, moderate fault and severe fault respectively.

The curves of equivalent reactance and equivalent resistance with short-circuit position under different fault degrees are drawn, as shown in Fig. 8 and Fig. 9 respectively, in which mild fault, moderate fault and severe fault are 6%, 12% and 18% respectively.

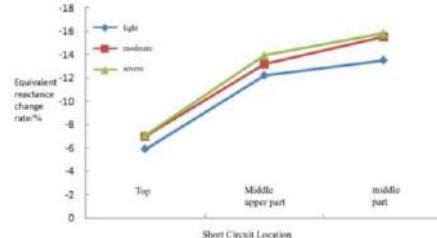


Figure 8. Relation between degree of short circuit and change rate of reactance.

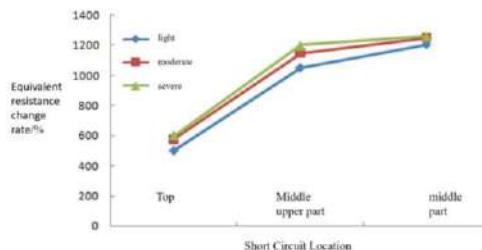


Figure 9. Relation between degree of short circuit and change rate of resistance.

The winding temperature is 95°C. Due to the symmetry of the short circuit position, only the top, middle, upper and middle short circuits are studied here. As can be seen from Fig. 8, with the deepening of short-circuit fault, the equivalent reactance decreases, and the decreasing range increases. Similarly, it can be seen from Figure 9 that with the deepening of short-circuit fault, the equivalent resistance tends to increase, and the increasing range increases.

5. Conclusion

In this paper, a set of on-line monitoring system for reactor short-circuit fault is built, and the influencing factors of equivalent reactance and resistance are studied. The main conclusions are as follows: 1) The equivalent reactance and resistance change rate of inter turn short circuit of reactor have the symmetry of short circuit position. The absolute value of change rate in the middle is the largest and the absolute value of change rate at both ends is the smallest. Under the same short-circuit position, the absolute values of the change rate of equivalent reactance and equivalent resistance decrease with the increase of temperature. 2) With the deepening of short-circuit fault, the equivalent reactance decreases and the decreasing range increases; The equivalent resistance shows an increasing trend, and the increasing range increases.

References

- [1] Zhang L, Jia-Sheng L, Wang Y H. Field test on the turn-to-turn insulation for 35kV dry-type air-core reac-tors[J]. Electric Machines & Control, 2014, 18(06): 66-71.
- [2] Wang Ping, Zhao Yingyu, Lv Fangcheng et al. Effect of rain dis-tribution on surface electric field of ± 100 kv smoothing reactor de-sign[J]. High Voltage Engineering, 2017(10):75-81.
- [3] Gao xiaodong, Qu Wentao, Chen Rengang. Fault analysis and pre-ventive measures on 35 kv air-core dry-type reactor[J]. Power Ca-pacitor & Reactive Power Compensation, 2015, 36(2): 85-88.
- [4] Fu Weiping, Zhao Jingwu, Huo Chunyan. Reason Analysis for a 35kV Dry Type Reactor Fault[J]. Power Capacitor & Reactive Power Compensation, 2011, 32(01): 59-62.
- [5] Zhang Bo. Application of Reflux Reactor in Solving Problem of In-sulation Burning at Station Rail End[J]. Railway Signalling & Communication Engineering, 2019, 16(S1): 16-19.
- [6] Wu Dongwen. Temperature field distribution and infrared tempera-ture measurement method research of 35kv dry-type reactor[J]. Transformer, 2013, 50(9): 62-65.
- [7] PAN Zhen, JI Donghong, ZHANG Wenqing, et al. Explosion analysis on 330 kV current transformer [J]. Power Capacitor& Reactive Pow-er Compensation, 2017, 38(1): 90-92.
- [8] JIANG Zhipeng, WEN Xishan, WANG Yu, et al. Test and coupling calculation of temperature field for uhv dry-type air-core smoothing reactor[J]. Proceedings of the CSEE, 2015, 35(20): 5344-5350.
- [9] Zhang Jie,Tan Xiangyu,Wang Ke et al. Application of distributed fi-ber bragg grating sensor system in dry-type air-core reactor[J]. Sci-ence Technology and Engineering, 2015(16): 160-164.
- [10] Sun Zhiyong. Design of on-line monitoring and fire warning system of dry-type reactor[D]. ShangDong: Shangdong University, 2016: 16-18.
- [11] Jiang Zhipeng, Wen Xishan,Wang Yu,et al.Test and Coupling Calcu-lation of Temperature Field for UHV Dry-Type Air-Core Smoothing Reactor[J]. Proceedings of the CSEE, 2015, 35(20): 5344-5350. [23]
- [12] LI Ying, YANG Songwei, ZHANG Jianhuan, et al. Study of temperature monitoring method of dry-type air-core reactor based on distributed optical fiber[J]. Instrument Technique and Sensor, 2017(2): 43-46,50.
- [13] WANG Xiaowen, LIU Yubo. The study based on detecting coil for dry-type shunt reactors fault detection[J]. Shenyang Institute of Engineering (Natural Science), 2016, 12(4): 336-339.
- [14] LI Xiaoxiao,YU Xiangyang. Characteristic analysis of three-phase four-leg inverter based load unbalance compensator for mi-crogrid[J]. Power System and Clean Energy, 2017,33(5): 13-17.
- [15] Lin Shenghong. Research on the Protection between turns of high voltage shunt reactors[D]. Nanjing:Nanjing Normal University, 2016.
- [16] Huang Xinbo,Zhang Long,Zhu Yongchan, et al. Online inter-turn in-sulation monitoring technology based on power angle characteristics for dry-type air-core reactor[J]. Electric Power Automation Equipment, 2019, 39(02): 143-148.
- [17] LIU Chuanghua,LIU Mei. Detection technology of the interturn insulation of dry-type air-core reactor J. Proceedings of the CSUEPSA, 2016, 28 (Supplement): 115-118.
- [18] GUO Xiangfu, MENG Bo. Interturn short circuit fault current and overcurrent protective system of shunt reactor with air core[J]. Transformer, 2008, 45(11): 20-22.
- [19] ZHANG Bin, HUANG Wenwu, LING Yun, et al. Simulation on transient characteristics of inter-turn short circuit fault of dry-type aircore reactor based on ANSYS Maxwell[C] // International Conference on Electrical Engineering and Automation Con-trol (ICEEAC2017). Nanjing,China:[s. n.],2017: 416-422.

This page intentionally left blank

Subject Index

1oo2 redundancy architecture	431	carbon reduction	188
2D grid maps	450	cellular automata	707
300MW subcritical boiler	148	charging pile	745
3S technology	71	clean energy	188
5G	515	cluster	35
5G NR	268	clustering	640
accelerometer	724	CMA	420
accountable authority	440	CNC	35
action recognition	467	CNNs	306
adaptive equalization	420	CO ₂	178
agglomerative clustering	716	coil offset	199
alerting	659	colored Petri net	707
all zero blocks	493	communication	335
Alzheimer's disease	63	communication architecture	102
antenna	335	complex coefficient	570
AP cluster	321	component materials	211
artificial intelligence algorithms	379	compressed sensing	321
artificial intelligence systems	7	compressor efficiency	431
artificial load	745	consensus control	118
artificial neural networks (ANNs)	239	constellation	335
asymmetric algorithm	387	Constrained Shortest Path First	
auto regression	14	(CSPF)	597
automatic registration	349	construction technology method	132
automatic target recognition	402	content delivery	544
autonomic nervous system	731	contextual anomaly detection	659
autonomy economical control	118	contrast enhancement	581
auxiliary equipment	171	convolutional neural network	54, 63
availability and confidentiality		corporate bond default risk	
of data	379	prediction	42
backward-assisted positioning	621	correlation analysis	220
Bayesian inference	159	countermeasures	484
Bayesian threshold	684	CPE	268
beam effect	500	cross attention	677
beyond 52.6 GHz	268	cubesat	335
bi-plateau histogram equalization	581	cyber physical security	379
big data	7	data augmentation	603
bilateral filter	581	data dimension reduction	631
bilinear group	440	data gathering	659
BiLSTM	14	data interaction	387
blind equalization	420	data processing	356
bluetooth technology	349	de-ICI filtering	268
bottleneck	515	decision directed algorithm	420
C-band	335	deep learning	54, 63, 369

deep space	335	GAN model	603
degree of short circuit	753	gateway	515
demand response	102	generalized hybrid carrier (GHC)	414
depthwise convolution	677	generative adversarial network	262
DETR	677	genetic algorithm	321, 737
difference operator method	402	green transformations	188
digital filters	292	HAD	394
distributed control	118	handoff	544
distribution network planning	220	harmonic current	745
distribution reactor	753	heart rate asymmetry	731
double-layer model	737	heterogeneous resource	
DPMZM	459	management	536
driver health	379	HEVC	493
dual mode algorithm	420	high availability SoC	431
dynamic simulation	745	high dynamic range	581
Ebola virus	707	high-performance	631
edge loss	262	high-speed data streams	550
EEG	54	highly integrated	525
electric vehicle wireless charging	199	Hobq	71
electronic system	525	home-based rehabilitation	724
emergency management	343	human-human interface (HHI)	724
emergency search and rescue	343	hybrid deployment	563
EMG	724	hyper-parameter optimization	283
empirical research	477	identity-based encryption	440
energy consumption structure	188	IEC61850	171
ensemble of timeseries forecasts	659	image classification	369
entity recognition	667	image segmentation	306, 716
environment	283	image sharpening	369
epilepsy	54	image super-resolution	262
equivalent reactance	753	IMDD-OOFDM	550
equivalent resistance	753	imitation learning	536
event trigger	118	immune response	707
evolutionary algorithms (EAs)	84, 239	improved grey wolf optimization	
expensive functions	84	(IGWO)	650
explainable	306	infrared imaging system	581
fault diagnosis	650	instrument device	141
FDTD	300	instrumentation selection	141
feature importance	283	integrated circuit assembly	239
federal learning in multi-domain	612	integrity	379
fingerprint extraction	640	intelligent question and answer	667
fingerprint localization	631	interactivity	1
flight control system	507	Internet of Things (IoTs)	90
forest swamp	77	interpretable	306
fractal	335	investment estimation	220
frequency multiplication circuit	459	isentropic efficiency	431
function design	102	Jaya algorithm	239
functional safety	362, 394	joint adjacency graph	467
fuzzy mathematics theory	199	JSON	349
FWA	515	junior high school mathematics	356

knowledge graph	667	network culture data	233
Ku-band	335	network culture development	233
LabVIEW	178	neuromuscular electrical stimulation	
last mile access	515	(NMES)	724
layout SLAM	450	node importance	575
LCL	745	non-Lambertian optical beams	500
learning fossilization	477	numerical analysis	26
left ventricle	306	OFC generation	459
life cycle management	90	OFDM	268
life forecasting	90	on-load tap changer (OLTC)	650
Link Plan Table (LPT)	597	optical data center network	591
lithium-ion batteries	14	optical networks	575, 612
load peak-valley	737	optical transmission networks	603
long short-term memory (LSTM)		optical vortices	300
network	650	Optimized OSPF with Link Plan	
machine learning	659	(OOWLP)	597
machine learning metric	26	OSNR prediction	612
machine load forecasting	255	outdoor terminal box	211
manufacturing	247	overcomplete dictionary	321
mapping of remote sensing	77	OVO-SMOTE-Adaboost ensemble	
massive MIMO	631, 640	model	42
master station	387	parabolic dune	71
MEC	544	parameter optimization	650
metamodels	84	parametric process optimization	239
metro optical network	544	particle swarm optimization	255
microservice	102	path loss	283
microstrip	335	peak to average power ratio	
microwave photonic filter	570	(PAPR)	414
microwave photonic phase shifter	570	pedestrian detection	394
miniaturized safety systems	362	personnel detection	677
mixed-grid optical network	557	phase noise	268
mobile device	306	phase tracking	268
monitoring of publication activity	7	phasors	292
morphological	71	physical network functions	563
motion	724	PMSE	26
multi condition constraint	603	PN	716
multi-class imbalanced		point-to-multipoint	557
classification	42	PolM	459
multi-layer satellite network		pose fusion	450
(MLSN)	597	position of short circuit	753
multi-point fusion positioning	621	power cable	90
multi-stage	35	power grid load	737
multimedia network	477	power grid weakness	159
Mumford-Shah	716	power system data	693
navigation	450	power grid	575
network autonomous learning		PTRS design	268
platform	484	Python	356
network communication	233	QoE	515
network culture	233	radar range profile	402

radar signal processing	402	spectrum availability	557
random forest	77	speech enhancement	684
RBF	84	SQLite	349
RDO quantization	493	statistical analysis	693
real-time	581	status assessment	90
reconfigurable bandwidth	591	streaming	515
recurrent neural network	90	structure design	211
redundant frames reduction	467	substation area model	171
refrigeration	178	subwavelength ring gratings	300
regulators	292	support vector machine	255
relays type selection	211	surrogate modeling	26
reliability	220, 394	suspension roll steer	315
remote sensing image	262	swarm	335
residual connection	667	symbol synchronization algorithm	550
resource allocation	557	symmetric algorithm	387
resource shortages	575	synchronization	507
RGB-D SLAM	450	systemic review	343
Ridgelet	321	task migration	536
RISC-V	362	teaching environment	477
RMS	292	teaching-learning based optimization	
robust prediction	35	(TLBO) algorithm	239
routing	557	terminal	387
RV1108	581	text mining	693
S-band	335	thermal system simulation	148
safety SoC	362	three detection coils	199
salt storage tank	132	three term weighted type fractional	
SATD	493	Fourier transform (3-WFRFT)	414
SCD	171	three-dimensional	621
scientometric indexing	7	threshold function	684
SCR inlet flue gas temperature	148	time series	35
security encryption	102	time slots	591
security level	507	total variance approach	315
seizure detection	54	tower-shaped solar thermal power	
self-driving	394	generation	132
Sentinel	77	trading platform	233
side-slip angle unified transfer		training sequence	550
function	315	transfer learning	306
silicon photonics	300	transformer	677
simulation	707	transmission distribution factor	159
simulation experiment	178	ultra-HD	515
sleep	731	unidirectional network device	379
SM9	440	user income	737
small-size	525	user stickiness	1
smart compressed air compressor	431	using efficiency	484
social Q & A community	1	variable-length data	603
software-defined network	536	vehicle handling dynamics	315
solar thermal power generation		video	515
system	141	virtual classes	467
spacecraft	525	virtual power plant	102

virtual simulation teaching		wavelet transform	684
platform	178	wide-load denitrification technology	
virtualized network functions	563	transformation	148
visible light communications	500	window-based method	402
visualization	26	wireless data center	500
vocabulary enhancement	667	wireless link characteristics	500
vocational college English		worker and artificial intelligence	247
learners	484	X-band	335
volume flow measurement	431	Zhihu website	1

This page intentionally left blank

Author Index

Abdelawwad, M.	362, 431	Feng, R.	102
Abi Zeid Daou, R.	379	Fränti, P.	716
Abou-Abbas, L.	54	Fu, M.	597
Abramonte, R.	292	Fu, Q.	387
Aguirre, Y.	292	Gao, B.	693
Al Shahadat, T.	431	Gao, C.	220
Aliaga, A.	292	Gao, L.	14
Amstutz, S.	659	Gao, X.	570, 612
Ang, Z.	387	Garg, A.	14
Arce, N.	292	Gelenbe, E.	343
Arellano Ramirez, C.E.	335	Geng, J.	525
Ba, Z.	71	Ghaffari, P.	707
Bao, Y.	349, 356	Gheorghe, M.	659
Bi, H.	343	Gong, C.	459
Boercsoek, J.	362, 379, 431	Gonzales Palacios, O.F.	335
Cai, B.	148	Goodridge, W.	515
Cao, H.	220	Gu, J.	557
Cao, W.	171	Gu, T.	745, 753
Carmi, S.	420	Guan, Z.	440
Carpentieri, B.	26	Guo, B.	591, 597
Chai, Y.	544	Guo, L.	536
Chen, H.	536	Guo, Y.	315, 591
Chen, J.	387	Gutierrez, F.	292
Chen, J.-x.	1	Ha, X.	575
Chen, L.	603	Hahn, E.	362, 431
Chen, X.	621, 631, 640	Han, T.	102
Chen, Z.	612	Han, Y.	148
Choi, S.	268	He, F.	387
Christos, P.	684	He, H.-m.	677
Collins, J.	659	He, X.	159
Cui, Z.	450	Henni, K.	54
Deng, F.	306, 581	Holzapfel, G.A.	26
Diaz Vargas, R.E.	335	Hu, J.	159
Ding, J.	315, 500	Huang, J.	603, 612
Dong, Z.	612	Huang, L.	477
Du, H.	102	Huang, S.	544, 563, 570, 591, 597
Duan, Z.	493	I, C.-L.	500
Ducq, Y.	379	Ingham, C.	659
El Hajal, G.	379	Irina, R.	7
Elena, P.	7	Jemal, I.	54
Fairbrother, J.	431	Jiang, G.	102
Fang, H.	188	Jiang, L.	484, 544
Feng, G.	575	Kalinowski, D.	362

Khan, K.	515	Miao, S.	118
Kholisakhon, I.	7	Miao, Y.	459
Kim, S.	268	Mitiche, A.	54
Konur, S.	659	Mueller, W.	362
Krini, O.	394	Müller, N.	394
Lai, D.T.C.	283	Myung, S.	268
Lan, F.	90	Na, Q.-l.	677
Lei, C.	550	Nguyen, M.	14
Lei, J.S.	667	Ning, Z.	63, 536
Li, G.	525	Niu, C.	387
Li, H.	220, 321, 745, 753	Olga, G.	7
Li, J.	199, 262, 467, 731	Otto, P.	431
Li, P.	731	Ouertani, A.	394
Li, R.	188	Pan, J.	507
Li, W.	159	Pang, C.	597
Li, X.	557, 563, 575	Patel, D.	716
Li, Y.	199, 414, 603, 731	Pavel, M.	7
Liang, G.M.	667	Pham, T.	14
Liang, L.	550	Pinchas, M.	420
Liang, Y.	603	Pudeev, A.	268
Liang, Z.	387	Qi, Z.	171
Liao, X.	63	Qian, J.	570
Lin, T.	612	Qian, K.	199
Lin, X.	306	Qin, T.	148
Liu, B.	262	Qin, W.	745, 753
Liu, C.	171, 467, 731	Quan, T.	14
Liu, H.	35	Ren, Y.	621, 631, 640
Liu, J.	306, 387, 581	Rong, J.	148
Liu, L.	467, 544	Sajjadinia, S.S.	26
Liu, Q.	132	Sanchez, V.S.	707
Liu, S.	459	Sani, U.S.	283
Liu, W.	159, 262, 525	Savelyev, D.	300
Liu, Y.	199, 233	Shah, N.	716
Liu, Z.	118, 141, 199	Shan, N.	493
Lu, F.	90	Shen, L.	35
Lu, L.	575	Si, P.	394
Luo, J.	321	Si, W.	63
Luo, S.	102, 440	Sia, T.T.L.	724
Luo, W.	737	Sibalija, T.	239
Ma, H.	650	Song, C.	90
Ma, P.	570	Song, H.	306
Ma, Q.	188	Song, R.	233
Ma, R.	575	Song, W.	667
Malik, O.A.	283	Song, Z.	414
Maltsev, A.	268	Stakem, P.H.	335
Marina, P.	7	Su, D.	677
Mendoza-Valencia, J.	247	Su, Z.	581
Meng, Q.	255	Sun, J.	42, 570
Mezghani, N.	54	Sun, T.	414

Sun, W.	621, 640	Xu, P.	387
Sun, Y.	178	Xu, T.	171
Sun, Z.	349, 356	Xuan, S.	467
Tadmor, S.	420	Xue, B.	90
Tallón-Ballesteros, A.J.	v	Xue, X.	591
Tan, L.	199	Yan, C.	731
Tang, F.	621, 631, 640	Yan, X.	211, 525
Tang, Y.	63, 563, 631	Yan, Y.	650
Tenne, Y.	84	Yan, Z.	557
Thanh Son, L.	402	Yang, B.	132, 141
Trung Kien, T.	402	Yang, C.-C.	369
Truong, L.	14	Yang, G.	148, 387
Van Loi, N.	402	Yang, H.	450, 597
Van Truong, T.	402	Yang, Q.	171
Vu Hop, T.	402	Yang, S.	268
Wan, Y.	450	Yang, S.Y.	667
Wang, B.	35, 414	Yang, W.	118
Wang, C.	220	Yang, X.	603
Wang, F.	557	Yang, Y.-x.	677
Wang, G.	188, 356	Yang, Z.	90
Wang, J.	77, 178	Ye, S.C.	667
Wang, K.	440, 467	Ye, T.	450
Wang, L.	500	Yi, L.	387
Wang, N.	575	Yin, S.	544
Wang, R.	581	Yong, C.Y.	724
Wang, W.	233	Yu, L.	737
Wang, X.	459, 459, 536, 621, 631, 640	Yu, X.	557, 575
Wang, Y.	148, 306, 581	Yu, Y.	148
Wang, Z.	211, 220	Yuan, G.	306
Wei, H.	493	Yuan, P.	90
Wei, J.	563	Yuan, S.-j.	1
Wei, W.	575	Yue, Y.	321
Wei, Y.	459	Zadeh, E.	659
Wei, Z.	693	Zaitsev, D.A.	707
Wu, C.	178	Zhai, Y.	171
Wu, J.	211, 550	Zhang, C.	90, 507
Wu, Q.	132, 141	Zhang, D.	550
Wu, X.	102	Zhang, G.	550
Wu, Y.	220	Zhang, H.	550, 603
Xi, H.	132, 141	Zhang, J.	102, 450
Xia, C.	621	Zhang, K.	262, 575
Xiao, Y.	550	Zhang, L.	563, 684
Xiao, Z.	631	Zhang, R.	132, 141, 684
Xie, J.	178	Zhang, S.	693
Xie, P.	603	Zhang, X.	591
Xie, X.	132, 141	Zhang, Y.	159, 603
Xie, Z.	102	Zhao, Yadi	693
Xin, J.	563	Zhao, Yisong	591
Xiong, Y.	118	Zhao, Yongli	557, 575, 603, 612

Zhao, Yuan	440	Zhou, W.	493
Zhao, Yuanyuan	612	Zhu, B.	612
Zhao, Z.	591	Zhu, J.	42
Zheng, K.	603	Zhu, Q.	557
Zheng, Q.	612	Zhu, X.	737
Zhou, H.	612		